1 Introduction

Researchers have long discussed the importance of transport policies in tackling social exclusion and improving access to out-of-home activities (Kwan, 2000; Lucas, 2012; van Wee & Geurs, 2011). In recent years, transport equity has been increasingly recognized as a crucial component of the sustainable mobility paradigm (Banister, 2008; UN HABITAT, 2013), attracting wider attention to the social
impacts of transport policies. As a result, there has been a growing interest in the distributional effects of transport investments on socio-spatial inequalities in access to opportunities (Ciommo & Shiftan, 2017; Forth, Manaugh, & El-Geneidy, 2013). Nonetheless, most studies measure accessibility using cumulative opportunity measures applied at a single spatial scale and zoning system. Although various studies have demonstrated that such accessibility estimates can be sensitive to the modifiable areal unit problem (MAUP), this does not seem to have changed practices (Apparicio, Abdelmajid, Riva, & Shearmur, 2008; Kwan & Weber, 2008; Ortega, López & Mónzon, 2012). The MAUP is a source of statistical bias that can significantly impact the results of spatial analyses because of the ad-hoc ways in which data are aggregated in space according to different spatial scales and zonal schemes, such as census tracts or traffic analysis zones. The effects of MAUP have been widely documented in urban studies and, researchers have sometimes found that the conclusions of spatial analyses and subsequent policy recommendations can be significantly different depending on the spatial scale and shape of areal units employed (Omer, 2006; Tan & Samsudin, 2017). No study has yet investigated whether the effects of MAUP could influence the equity assessment of transportation projects.

Moreover, while many researchers and transport planners in developed countries are concerned with improving accessibility (Boisjoly & El-Geneidy, 2017; Papa, Silva, te Brömmelstroet, & Hull, 2015) and equity (Karner & Niemeier, 2013; Manaugh, Badami, & El-Geneidy, 2015), these issues have received much less attention in the Global South (Blanco, Lucas, Schafran, Verlinghieri, & Apaolaza, 2018; Vasconcellos, 2001). The literature on justice in transport policy is still predominately focused on developed countries in North America and Europe, with a few notable exceptions (Campbell, Rising, Klopp, & Mbilo, 2019; Delmelle & Casas, 2012; Maia, Lucas, Marinho, Santos, & de Lima, 2016; Vermeiren et al., 2015).

This paper investigates how investments in public transport can reshape socio-spatial inequalities in access to opportunities and it shows how the equity assessment of transport policies based on cumulative-opportunity measures can be sensitive to MAUP. Using Rio de Janeiro (Brazil) as a case study, this paper looks at the transport policies adopted in preparation to host the 2014 Football World Cup and the 2016 Olympic Games. It analyzes how those policies impacted access to schools and job opportunities for different income groups considering different spatial scales and zoning schemes.

Recent sports mega-events triggered substantial investment in Rio’s public transport infrastructure. These investments played a central role in the justification used by public authorities to host those mega-events (Kassens-Noor, Gaffney, Messina, & Phillips, 2016). According to various official documents, those infrastructure projects would improve accessibility to major sports venues in the city but they would also leave a positive legacy that would help improve transport conditions in the city, particularly for low-income people living in deprived areas (BOC, 2009; Brazil, 2009; Rio de Janeiro, 2008). Nonetheless, it is not clear which income groups and areas of the city have benefited from that promised legacy, particularly because those projects in Rio were followed by a reorganization of many bus lines and by cuts in service levels in response to a drop-in passenger demand (see Section 2). Overall, these policies have brought significant changes to the city’s transport system in a relatively short period of time, giving a unique opportunity to assess its equity impacts on urban accessibility to various types of activities for different population groups.

This study estimates how access to public high-schools and formal jobs have changed for different income groups before and after the transport policies driven by mega-events held in Rio de Janeiro in 2014 and 2017. A quasi-counterfactual scenario was also used to separate the effects of newly added infrastructure from the reorganization of transport services. Employment and educational opportunities were chosen because of their central role for human development and in the reproduction of socio-economic inequalities in society. Accessibility levels are calculated using geolocated timetables of Rio’s public transport organized in General Transit Feed Specification (GTFS) format combined with fine-
grained sociodemographic data from the population census and geolocated data on schools and jobs. The spatial association between accessibility changes and household income per capita is tested using spatial regression and cluster analysis. The investigation is conducted using multiple geographical scales and zoning schemes to test whether results are robust to the modifiable area unit problem (MAUP), i.e., whether the conclusions of the evaluation remain the same when the data is aggregated in space using different geographical scales and areal units.

The paper is organized as follows: Section 2 presents a brief review of the literature on transportation equity. Section 3 presents the study area of Rio de Janeiro and how its public transport network evolved between 2014 and 2017. Data and methods are presented in Section 4 and results presented in Section 5. Finally, Section 6 brings a discussion of the main conclusions about the case study of Rio and some lessons that can be drawn for future studies on the equity impacts of transport policies in other cities.

2 Literature review—transportation equity

Equity is a multidimensional concept and there is no standard definition of equity across studies that evaluate the social impacts of transport policies (Martens, Golib, & Robinson, 2012; van Wee & Roeser, 2013). There is, however, a growing consensus in both transport planning practice (Boisjoly & El-Geineidy, 2017; Manaugh et al., 2015) and in the academic literature (Martens, 2016; Pereira, Schwanen, & Banister, 2017; van Wee & Geurs, 2011) that improving access to key destinations such as employment, healthcare, and educational services should be among the primary goals of equitable transport policies; and moreover, that such policies should prioritize improving accessibility for disadvantaged groups such as the elderly, disabled, and low-income people who are typically more dependent on public transport.

This convergence in the literature has been largely influenced by or is in line with the theory of justice proposed by John Rawls (Rawls, 1999, 2001). A core idea in this theory is that morally arbitrary factors beyond a persons’ choice such as being born with a physical disability, in a poor family or ethnic group should have no bearing on life chances and opportunities. Rawls is particularly concerned with how institutions in society should be arranged in order to mitigate the negative effects those factors may have on life chances and thus minimize inequalities of opportunities that arise from arbitrary circumstances to the benefit of the better-off groups and at the expense of the worse off. While Rawls’ theory relates more directly to the definition of society’s basic institutions, Rawls proposes that his high-level theory of justice ultimately should serve as a reference against which the fairness of proposed laws and policies should be assessed (Rawls, 1999, pp. 175, 253, 314–317; 2001, pp. 89-90).

Previous studies have proposed to use Rawls’ principles of justice to guide and assess transport policies (Pereira et al., 2017; van Wee & Geurs, 2011; van Wee & Roeser, 2013) and urban planning more generally (Basta, 2015; Fainstein, 2010; McKay, Murray, & Macintyre, 2012). In a recent paper, Pereira, et al. (2017) put forward an interpretation of how Rawls’s theory of justice can be applied to evaluate the fairness of transport policies. According to the authors, governmental investments in transport infrastructure and services can only be considered just if they give priority to benefit those disadvantaged groups whose transport experience is systematically undermined by morally arbitrary circumstances. The concern with justice thus requires researchers to move beyond descriptive cross-sectional analysis of how different groups have different levels of access to opportunities towards a better understanding of what role governmental policies can play in reducing such inequalities (Pereira et al., 2017). Another core idea in Rawls’ theory of justice is the absolute priority and inviolability of one’s basic rights and liberties such as the physical and psychological liberty and integrity of the person (Rawls, 2001, p. 44). Accordingly, no governmental policy can be considered fair if it violates these basic rights and liberties,
even for the sake of general welfare. Although this idea might be more often embedded in the principles that guide transport policies in the Global North, it is much less so in countries in the Global South, as the case of Rio will illustrate.

3 Study area: Rio de Janeiro

Rio de Janeiro is one of the largest and richest cities in the Global South and yet one of the most unequal (UN-HABITAT, 2010). In different ways, Rio is an exemplar of what more cities may become in the future, the result of rapid population growth coupled with stark socio-spatial inequalities and poor transport infrastructure where there is a strong increase in automobile use with all the negative externalities it creates (Ribeiro, 2014; UN HABITAT, 2013). Like many other cities in the Global South, Rio has a history of uneven urban development marked by spatial segregation (Préteceille & Cardoso, 2008; Ribeiro, Rodrigues, & Corrêa, 2010) and an unequal distribution of public transport infrastructure (Câmara & Banister, 1993). Moreover, traffic conditions have been deteriorating over the past decade, as reflected by the rise in average one-way commute time by all transport modes from 41.5 minutes in 2001 to 49.6 minutes in 2014 (Pereira & Schwanen, 2013). In 2014, one in every four commuting trips took longer than one hour.

When the city of Rio bid to host the 2014 Football World Cup and the 2016 Olympic Games, one of the key motivations of public authorities was to use those mega-events to leverage urban development and improve the city’s transport conditions (Gaffney, 2010; Kassens-Noor et al., 2016). The city invested more than US$4.5 billion in its public transport system in preparation for the events (Castro, Gaffney, Rodrigues, dos Santos e Orlando Alves dos Santos, Jr. 2015). Some of the most significant investments included the extension of a subway line, the construction of a light-rail system and two BRT corridors that together stretch approximately 108 kilometers across the city (Figure 1). This figure also shows the main clusters of sports venues of recent mega-events, which were generally located in consolidated urban areas and close to existing transport hubs. Alongside this new infrastructure, local authorities have also reorganized many bus lines to accommodate the newly added infrastructure and to streamline the transport system. More recently, however, a severe economic crisis led to a 70% cut in the budget of the Secretary of Transport (Magalhães & Rodrigues, 2017), which coupled with a drop in the number of passengers in the public transport system (França, 2016; Rodrigues, 2017) has led the government to adopt fiscal austerity measures that have reduced transport services. In total, 70 bus lines out of 485 were eliminated, 41 lines were rerouted or shortened and 16 new bus lines were created (G1, 2017). These measures can potentially undermine the positive effects of the new transport investments and cast doubt on which socioeconomic groups and areas in the city have benefited from the accessibility changes brought about by this cycle of investment and disinvestment.
According to Rio’s candidature files to host the World Cup and the Olympics, the aim of those investments was to create a high-capacity transport ring connecting key areas in the city where most of the sports competitions would take place and, above all, to benefit “low-income workers, who live in the most distant neighborhoods and spend more time in traffic.” (Brazil, 2009, p.54). Nonetheless, one year after the Olympics, those investments have already been widely criticized for being over-budget and the subject of corruption investigations (Cuadros, 2016; Guimarães & Leitão, 2017; Sandy, 2016). Academics and grassroots movements have also claimed that those projects were elaborated with little social participation and transparency (Legroux, 2014; Sánchez & Broudehoux, 2013) and that many BRT stations present barriers to those with physical disabilities and poor integration with other transport modes (ITDP Brasil, 2014, 2015). Moreover, official figures indicate that those investments led to the eviction of 2,125 families between 2009 and 2015 to create space for transport infrastructure (Rio de Janeiro, 2015). According to various accounts, many those evictions involved physical and psychological violence by the police, characterizing the violation of human rights (Barbara, 2016; CPCORJ, 2015; Gaffney, 2016; Watts, 2015). Despite the utilitarian discourse that those evictions were justifiable because the transport investments would bring large benefits to a great number of people, the violation of rights and the other issues mentioned are of central importance from a social justice point of view and go explicitly against a Rawlsian conception of justice (Pereira et al., 2017). These issues are not addressed in this study but should be considered in a comprehensive equity assessment of the mega-events in Rio.

Another important equity issue to consider involves the distributional effects those investments have had in improving the transport conditions of different social groups in the city. While the official rhetoric proposed prioritizing low-income groups in line with the design of equitable transport policies, the extent to which the investments have effectively improved the accessibility of low-income neighborhoods is a question that remains largely unanswered (Kassens-Noor et al., 2016), and which we address in the next sections.
4 Methodology

Accessibility is understood as the ease with which a person can reach places and opportunities from a given location and it results from the interaction between the transport system, land-use patterns, and the constraints of individuals (van Wee & Geurs, 2011). This paper estimates how transport policies implemented in Rio changed accessibility levels considering different spatial scales and zoning schemes, and it examines whether wealthier areas have gained more accessibility than poorer ones. The method used in this study involved five secondary data sources and two main steps as detailed below.

4.1 Data sources

Population count data comes from the 2010 Brazilian Census, organized in a regular grid of 200x200 meters (IBGE, 2016). Resident population in each grid cell was categorized according to income decile based on the average household income per capita of each grid cell. This was imputed from the 2010 census data organized in 1,136 relatively homogeneous socioeconomic polygons known as Human Development Units (Ipea, UNDP, & FJP, 2015). This method needs to be used with caution because it incurs ecological fallacies by disregarding heterogeneity within Human Development Units.

Data on schools come from the School Census conducted by the Brazilian Ministry of Education, covering all of the 278 public high schools in the municipality of Rio de Janeiro in 2015. Data on formal jobs come from RAIS, a dataset organized by the Ministry of Labor and Employment. RAIS is a national register that brings the full address of all public and private institutions and the socioeconomic characteristics of their employees working in the formal labor market. In 2015, there were 2,914,238 formal workers employed in 227,362 institutions in Rio. Among the 50 largest employers, 17 institutions were found to misreport their address and were removed from the analysis. The location of 83,589 employees of the secretary of education could be recovered from the school census. In the end, a database covering 92.3% of all formal workers in the city was used. Schools and jobs were geolocated using the full address information available in their respective datasets.

A limitation of this study is that it does not cover the informal labor market because there is no data source with the addresses of informal jobs. Approximately 36% of Rio’s workers were employed informally in January 2016. Nonetheless, informal jobs are relatively more accessible with shorter commute times than formal jobs (Morte-Baumvol, Aguilera, Bonin, & Nassi, 2016), and it is reasonable to assume that formal jobs are generally preferable given the associated labor rights and social benefits. Moreover, the 2003 household travel survey of Rio shows that the numbers of formal and informal jobs in each traffic zone are correlated at 0.78 (Pearson correlation statistically significant at the 0.001 level), suggesting that the spatial distribution of formal and informal jobs in the city are not radically different.

Spatial information on road networks and pedestrian infrastructure comes from OpenStreetMap. Finally, data on the public transport network organized in General Transit Feed Specification (GTFS) format was provided by Fetranspor (Federation of Passenger Transport Companies in Rio de Janeiro) for April 2014 and March 2017. GTFS data provides detailed geolocated information of routes, stops and timetables of the public transport system. Because the GTFS data available only covers the transport network within the municipality of Rio, schools and jobs outside of the study area are not considered. As a result, absolute accessibility levels close to the northern border of Rio are likely to be underestimated, which may slightly inflate the relative change in accessibility ($R_{io}$ in eq. 2 below) in areas close to that border. However, the effects of excluding schools and jobs outside Rio on the analysis will diminish rapidly with increasing distance from the border and be smaller in border areas that contain greater numbers of schools and jobs. The overall effects on the analysis will therefore be limited.
In the before-and-after comparison of Rio’s transport network, the spatial distribution of the population, as well as the location of schools and jobs, were kept constant. This assumption allowed us to isolate the effect of the recent transport policies on the variation in accessibility levels between 2014 and 2017. This assumption disregards changes in the spatial allocation of schools/jobs and redistribution of population and socio-economic groups that may have occurred over the period due to the processes of displacement and gentrification. Nonetheless, previous research has shown that such phenomena are fairly stable over time in Rio, particularly over short timescales as the one used in this study (Lago, 2000; Ribeiro, 2014); changes during the 2010–2017 period are thus unlikely to have significantly affected the overall results of the current analysis.

4.2 Estimating accessibility levels

A cumulative opportunity measure was used to estimate accessibility in terms of how many schools and jobs could be reached from people’s residences via public transport and walking under 60 minutes. This type of measure was chosen because there has as yet been no exploration of the extent to which transport equity appraisals based on this metric are sensitive to methodological choices regarding spatial scales and zoning schemes of accessibility modeling. Moreover, this metric is the most commonly used by transport agencies in North America and Europe to compare the benefits of transportation investments and evaluate their social impacts (Boisjoly & El-Geneidy, 2017). In most cases, those agencies use a single time threshold that vary between 30 and 60 minutes (ibid.). A threshold of 60 minutes was chosen given that Rio has relatively higher commute times when compared to international metropolitan areas (Pereira & Schwanen, 2013). This threshold is also very close to the average commute time by public transport and walking (57 minutes) observed in Rio in its latest household travel survey of 2013 (Central, 2016). Some advantages of this accessibility measure are that it does not require prior information about travel behavior, it is computationally inexpensive to calculate, and it produces results that are easy to communicate to policy makers and stakeholders. This measure has limitations, though, since it assumes that all destinations are equally desirable, regardless of the time spent on travelling; and it does not take monetary costs or competition effects into account (Geurs & van Wee, 2004). Moreover, this metric uses an arbitrary time threshold and thus ignores activity points just outside the threshold.\(^1\)

Following the recommendation of previous studies (Horner & Murray, 2004; Páez & Scott, 2005), the analysis was conducted using multiple spatial scales and zoning schemes in order to test whether the results are robust to MAUP and not a simple artefact of the ways in which data are arranged in space. The modifiable areal unit problem has two components. The scale effect, which relates to how the result of spatial analyses differ when they are conducted at different geographical scales, and the zoning effect, which relates to when those results are influenced by the shape of the aggregation units used in spatial analyses. To account for both effects, accessibility analysis considered the municipality of Rio divided into the official traffic zones of Rio’s latest household travel survey of 2013, and two hexagonal grids of 2 km and 0.5 km. The number of polygons in these spatial schemes is, respectively, 390, 313 and 3485. The traffic zones and the 2 km grid have a similar number of polygons (similar scale) but vastly different shapes, while the two hexagonal grids allow us to estimate accessibility levels at markedly different spatial scales but using the same aerial unit. The sizes of the grid cells were chosen because they allow us to generate granular estimates of accessibility and capture even subtle differences between adjacent cells without compromising computational tractability.

\(^1\)See Pereira (2019) for a discussion on the how the time threshold choice can influence the equity appraisal of transport projects.
In the first step of the analysis, OpenTripPlanner\textsuperscript{2} was used to estimate various travel-time matrices by public transport and/or walking between every pair of zone centroids for each spatial grid. Following the recommendation of Stępniak and Jacobs-Crisioni (2017), population-weighted centroids were used in order to minimize aggregation errors. To account for temporal variations in public transport services across different times of the day, 36 travel-time matrices were calculated departing every 20 minutes between 7am and 7pm. These travel-time matrices consider door-to-door estimates that incorporate walking time from the point of origin to the transit stop, waiting time for the vehicle, actual travel time through the transport network and eventual transfers, and the walking time from the transit stop to the destination.

Next, travel-time matrices were combined with land-use data (population, schools and jobs). Based on equation (1), we calculated the median number of opportunities that can be accessed from each grid cell via public transport in under 60 minutes across the 36 travel-time matrices for each year (2014 and 2017). In order to account for a qualitative match between socioeconomic levels of workers and jobs, estimates of access to jobs were based on the assumption of a proxy match between household income and educational qualification of jobs. For residents in households above the 5th income decile, accessibility estimates only considered jobs that required high or secondary education, while for households below the 5th income decile only jobs that required secondary or primary education were considered.

\begin{equation}
A_{o,i,T} = \text{median} \left( \sum_{d=1}^{n} P_d f(t_{odr}) \right)
\end{equation}

\begin{align*}
  f(t_{odr}) &= \begin{cases} 
  1 & \text{if } t_{odr} \leq T \\
  0 & \text{if } t_{odr} > T
\end{cases}
\end{align*}

Where:
- $A_{o,i,T}$ is the accessibility level at origin $o$ for population of income $i$ within time threshold $T$.
- $P_d$ is the number of opportunities (e.g., jobs or schools) in location $d$.
- $t_{odr}$ is the travel time in minutes between origin $o$ and destination $d$ at departure time $r$.
- $f(t_{odr})$ is a binary function of time threshold and takes values of 1 or 0, depending on whether travel time to $d$ $r$ is larger or smaller than time threshold $T$.

The positive accessibility gains brought about by the new infrastructure investments in Rio may have been offset by the reorganization of the bus lines and by the austerity measures adopted, including the reduction of service levels in some areas of the city. In order to measure what would have been the sole effect of the new transport investments, accessibility changes were also estimated under a quasi-counterfactual scenario. This scenario assumes that all public transport services provided in April 2014 would have been kept constant, so the only changes to Rio’s transport system would have been the addition of the new infrastructure, i.e., the BRTs Transcarioca and Transolímpica and the new subway and light-rail lines, highlighted in Figure 1.

Due to data availability constraints, the accessibility analysis deployed in this study has not considered issues related to affordability, safety, age, gender or disability. It is widely documented in the literature how transport disadvantages related to these issues can hinder one’s ability to use public transportation (Casas, 2007; El-Geneidy et al., 2016; Rya, Wretstrand, & Schmidt, 2015). By not considering these issues, this study is likely to underestimate accessibility inequalities between high- and low-income groups (Neutens, Schwanen, Witlox, & Maeyer, 2010).

\textsuperscript{2}OpenTripPlanner is an open-source multimodal trip planner available at https://github.com/opentripplanner/OpenTripPlanner
4.3 Analyzing association between income and accessibility change

The final step was to test whether there is any association between household income per capita and the accessibility changes between 2014 and 2017, and whether this association is robust to MAUP. While this could be done simply by comparing the average accessibility variation for different income groups, regression models allow us to test whether these differences are statistically significant while controlling for other confounding factors. One important factor in this case is spatial autocorrelation. It is well-known that many social, transport and land-use processes are spatially dependent in the sense that the characteristics of a location are related to its neighboring areas (Anselin, 2010). Similarly, the accessibility level of one location is affected by the characteristics of the transport system and land use in nearby areas. The presence of spatial dependence in the data violates basic assumptions of conventional statistical analysis, such as t-tests or Pearson correlation and OLS regressions, and leads to “information loss, biased and/or inefficient parameters and the possibility of seriously flawed conclusions and policy prescriptions” (Páez & Scott, 2005, p.55). To overcome these limitations, spatial regression models have been commonly used by studies investigating the relationship between transport accessibility and land-use prices (Efthymiou & Antoniou, 2013; Mitra & Saphores, 2016; Mulley, Ma, Clifton, Yen, & Burke, 2016), but these models have thus far found little uptake in the literature on transportation equity.

In order to control for spatial autocorrelation, a spatial regression model was used to estimate whether wealthier areas were able to attract more accessibility gains from the transport policies recently implemented in Rio. Other variables can influence where new transport infrastructure is likely to be built and consequently which locations have higher accessibility gains from new investments. For example, areas with greater densities of population and economic activities are likely to have higher transport demand and make transport projects more economically viable (Kutz, 2003). Topography is another relevant variable that puts physical barriers for construction, impacting construction costs and influencing where surface transportation modes are likely to go (ibid.). In order to control for these factors, the regression model includes local level control variables for topography, population and job or school density. These variables were chosen because they are known to influence transport policy decisions and accessibility levels (ibid.), but the authors acknowledge that the specification of the regression is not informed by a full theoretical framework listing of all relevant variables that could affect the dependent variable. The baseline model without spatial effects is written as:

\[ \log (R_{i,o}) = \beta_0 + \beta_1 \log (I_i) + \beta_2 \log (P_i) + \beta_3 \log (O_{i,o}) + \beta_4 \log (E_i) + \varepsilon_i \]  

Where:  

- \( R_{i,o} \) is the relative change in accessibility level from polygon \( i \) to opportunities of type \( o \) between 2014 and 2017.  
- \( A_{i,o}^{2017} \) is the number of opportunities of type \( o \) that could be accessed from origin \( i \) under travel-time threshold \( T \) in the year 2017.  
- \( A_{i,o}^{2014} \) is the number of opportunities of type \( o \) that could be accessed from origin \( i \) under travel-time threshold \( T \) in the year 2014.  
- \( I_i \) is the average household income per capita in polygon \( i \).  
- \( P_i \) is the population density in polygon \( i \).  
- \( O_{i,o} \) is the density of jobs or number of schools in polygon \( i \).  
- \( E_i \) is the average terrain elevation in meters of polygon \( i \).
Variation in accessibility was measured as the ratio between accessibility levels in 2017 and 2014. This approach means that changes in accessibility have proportionally lower impact in those areas that already had higher accessibility in 2014, in line with decreasing marginal returns. In other words, it implies that for a person living in an area with access to 100 jobs, an increase of, for example, 50 jobs add more to this person’s utility than for a person who already had access to 1000 jobs.

The spatial dependence underlying the baseline model was evaluated using the Lagrange multiplier test following standard methods used in the literature (Elhorst, 2010; LeSage, 2008). This suggested that the spatial Durbin model (SDM) was the most appropriate model. The SDM specification controls for spatial autocorrelation in both the dependent and the independent variables. This specification allows estimation of how strongly the accessibility gain in polygon i is associated not only with the local characteristics (including income) of i but also with the characteristics and accessibility gains in the neighboring polygons of i. Because the Durbin model includes a spatial lag in both set of variables, one advantage of this model is that it produces coefficient estimates that are unbiased even if there is one or more relevant omitted variables in the regression equation, and thus it minimizes the issue that the regression is not informed by a full theoretical framework (Elhorst, 2010; LeSage, 2008). Polygons were defined as neighbors if they share at least one boundary point using a standard queen contiguity matrix (ibid.).

This spatial Durbin model calculates the global average spatial association between income and changes in accessibility for the entire city. To complement the global analysis, local correlation between those two variables was estimated using the local indicator of spatial association (LISA). The bivariate LISA is an extension of its univariate version (Anselin, Syabri, & Smirnov, 2002) and it measures the correlation between one variable in an area and the average of another variable in its neighbors. In practice, this is a useful technique to identify clusters of areas with high and low income that have gained or lost access to opportunities. For this local correlation measure, a queen contiguity matrix was used to define neighboring areas. Pseudo p-values were estimated with a thousand random simulations in each round.

5 Results

5.1 Descriptive results

Figure 2 presents the spatial distribution of income groups in Rio according to different spatial scales and zoning schemes. It shows how lower-income families are mainly located in the north and particularly northwest areas of Rio while higher-income groups live mainly along the coast in the southwest close to the city center where most of the economic activity is concentrated. This figure also illustrates how spatial analysis at coarser resolutions or using different zoning schemes can lead to information loss. With the aggregation of household income per capita one loses valuable detail regarding the spatial heterogeneity of the data and, in the most extreme cases, the income decile classification of some areas changes from poor to rich and vice-versa.
Figure 2. Spatial distribution of population classified by deciles of household income per capita, Rio de Janeiro, 2010

Figure 3 illustrates for the year 2017 how median access to formal jobs and public schools by public transport varies substantially across space. For both types of opportunities, it becomes clear that accessibility levels are greater along the train and subway lines and part of the new Transcarioca BRT line. Figure 3 also draws attention to the way spatial inequalities in accessibility are largely shaped by the unequal distribution of land-use activities. The historical development of Rio, with strong concentration of jobs close to the city center helps explain, for example, why residents in the west region of the city have such low levels of access to opportunities. To some extent, this issue has been minimized by the spatial planning of public schools, which has been relatively successful in allocating public high schools more evenly across Rio. Moreover, these figures also show how spatial analyses at a higher resolution allows for accessibility estimates with more spatial detail, making the influence of transport corridors on spatial accessibility more visible.
The summary results suggest that average access to jobs and schools by public transport dropped by approximately -4.5% and -6.1% between 2014 and 2017 in the city of Rio. Using a quasi-counterfactual scenario to isolate what would have been the sole effect of the new infrastructure investments, the analysis shows that between 2014 and 2017 average access to jobs and schools would have increased by approximately 13.4% and 11.7% respectively. These findings suggest that the rationalization of bus lines and recent cuts in service levels have played an important role in offsetting the positive impacts of the new transport infrastructure.

Nevertheless, these are only average results that hide heterogeneity in the data, and it is important to analyze how accessibility changes varied by income groups and areas of the city. A crucial issue to consider from a social justice point of view is not only the level of inequality in access to opportunities but more importantly to what extent new policies contribute to reducing such inequalities, particularly by improving the access of lower-income groups. In 2014, the level of access to jobs was 84% higher for the richest 20% than for the poorest 20%. This inequality increased to 116% in 2017 as a result of the new infrastructure expansion combined with the reorganization of many bus lines and cuts in service.
levels. The analysis of the counter-factual scenario indicates that the infrastructure investments alone would still have increased this inequality to 92% in the year 2017. These results provide evidence that the accessibility benefits from the recent cycle of investments and disinvestments in Rio generally accrued to middle- and higher-income groups and they have reinforced rather than reduced socio-spatial inequalities in access to opportunities.

Figures 4 and 5 compare for each zoning scheme and scale of analysis how the number of jobs and schools accessible have changed between 2014 (horizontal axis) and 2017 (vertical axis) in both implemented and quasi-counterfactual scenarios. In these figures, each dot represents a grid cell where dot size varies according to population size and color indicates its income classification. Dots located above the diagonal line have seen an improvement in accessibility by public transport. From a simple visual analysis, these results look consistent across different spatial scales and zoning schemes. Both figures suggest that those areas that gained accessibility are mostly wealthier areas (colored in blue) while most of the areas with accessibility loss are poorer ones. Regarding access to schools, there are substantially more dots below the line, reflecting that access to schools has been affected more adversely than job access by the policies implemented (Figure 5). The income composition and distribution of dots above the diagonal line is very similar in the implemented and quasi-counterfactual scenarios. There is a marked contrast, however, in the distribution of dots below the diagonal line, which gives a clearer picture of how the reduction in services levels implemented between 2014 and 2017 undermined access to jobs and schools, offsetting the benefits of recent transport investments.

Figure 4. Proportion of formal jobs accessible from each grid cell via public transport and walking under 1 hour in the years 2014 (X axis) and 2017 (Y axis) in implemented and counterfactual scenarios, Rio de Janeiro
An important component of an equity evaluation of the policies implemented in Rio is the identification of those income groups in the city that have benefited the most from the recent cycle of transport investments and disinvestments. Although a simple visual analysis of Figures 4 and 5 suggests that most gains in access to jobs and schools have accrued to richer areas (shown in blue), the answer to this question demands a more robust statistical analysis.

5.2 Global association between income and accessibility change

A spatial regression model provided a more robust analysis, indicating whether the differences in accessibility gains by income levels are statistically significant while controlling for other confounding factors. Moreover, it allowed us to test whether the distributive effects of recent transport policies are sensitive to the modifiable areal unit problem (MAUP), which cannot be captured only by looking at the descriptive analysis presented thus far.

Tables 2 and 3 report the results of the spatial Durbin model, measuring how changes in access to jobs/schools between 2014 and 2017 are associated with household income per capita in 2010 in both implemented and counterfactual scenarios. In the scenario of implemented policies in the period, the results show a small but positive association between income and variation in accessibility. However, both the strength and statistical significance of this association vary across multiple spatial scales and zonings.
The difference in the results between 0.5 and 2 km grids reflect the scale effect of MAUP. Meanwhile, the difference between 2 km grid and traffic zones reflect the zoning effect, and the difference between 0.5 km and traffic zones result from the interaction between both the scale and zoning effects of MAUP. Using 0.5 km hexagonal grid, for example, a 1% higher income is associated with an increase of 0.08% in the number of schools accessible by public transport (Table 1). This effect is twice as large when the analysis is conducted based on a 2 km grid or on traffic zones (0.16%) but the result is only statistically significant when traffic zones are used, a consequence of the interaction between both scale and zoning effects of MAUP. When analyzing the changes in access to jobs (Table 2), the results only showed to be statistically significant for the analysis based on traffic zones. These results suggest a moderate association where a 1% higher income is associated with an increase of 0.21% in the number of jobs accessible by public transport. The statistical significance of the control variables used in the model also varied in an inconsistent way. Nonetheless, the direction of the associations shows that recent policies in Rio had on average larger accessibility benefits in those areas with higher population and job densities, and lower accessibility gains in communities located on hilly areas.

Under a quasi-counterfactual scenario, the results of the spatial Durbin model (Tables 2 and 3) indicate that the direction of the association between income and accessibility gains remains positive, but the magnitude of the effects is generally smaller, and the significance of some spatial scales also change. Overall, it also shows that even if there had been no reorganization of bus lines and austerity measures, wealthy areas in the city would still have benefited more from the new transport infra-structure than poor areas. Furthermore, these tables confirm that the statistical relationship between income level and accessibility gain is affected by both the scale and zoning effects of MAUP.

Table 1. Total effects of spatial Durbin model showing association with gains in access to public high schools under various spatial scales and zoning schemes in implemented and counterfactual scenarios

<table>
<thead>
<tr>
<th>Dependent Variable: log(ratio)</th>
<th>Implemented Scenario</th>
<th>Counterfactual scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5 Km</td>
<td>2 Km</td>
</tr>
<tr>
<td>log(RDPC)</td>
<td>0.08** (0.04)</td>
<td>0.16 (0.09)</td>
</tr>
<tr>
<td>log(popd)</td>
<td>0.03*** (0.02)</td>
<td>0.08 (0.06)</td>
</tr>
<tr>
<td>log(jobd)</td>
<td>0.03 (0.02)</td>
<td>0.08 (0.07)</td>
</tr>
<tr>
<td>log(elevation)</td>
<td>-0.08** (0.03)</td>
<td>-0.10 (0.05)</td>
</tr>
<tr>
<td>rho</td>
<td>0.88*** (0.01)</td>
<td>0.54*** (0.05)</td>
</tr>
<tr>
<td>Pseudo R2</td>
<td>0.76</td>
<td>0.28</td>
</tr>
<tr>
<td>Num. obs.</td>
<td>3485</td>
<td>313</td>
</tr>
<tr>
<td>Parameters</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-280.63</td>
<td>-157.25</td>
</tr>
<tr>
<td>AIC (Linear model)</td>
<td>4124.87</td>
<td>432.85</td>
</tr>
<tr>
<td>AIC (Spatial model)</td>
<td>-543.26</td>
<td>377.01</td>
</tr>
<tr>
<td>LR test: statistic</td>
<td>4670.13</td>
<td>57.84</td>
</tr>
<tr>
<td>LR test: p-value</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Notes: AIC, Akaike information criterion. Standard errors are in parentheses: ***p < 0.001, **p < 0.01, *p < 0.05.
Table 2. Total effects of spatial Durbin model showing association with gains in access to formal jobs under various spatial scales and zoning schemes in implemented and counterfactual scenarios

<table>
<thead>
<tr>
<th>Dependent Variable: log(ratio)</th>
<th>Implemented Scenario</th>
<th>Counterfactual scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5 Km</td>
<td>2 Km</td>
</tr>
<tr>
<td>log(RDPC)</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>log(popdens)</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>log(jobdens)</td>
<td>0.03*</td>
<td>0.08*</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>log(elevation)</td>
<td>-0.06</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>rho</td>
<td>0.89***</td>
<td>0.47***</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.06)</td>
</tr>
</tbody>
</table>

Notes: AIC, Akaike information criterion. Standard errors are in parentheses; ***p < 0.001, **p < 0.01, *p < 0.05

Finally, spatial regression analyses show that the direction of the relationship between income and accessibility change remains the same across spatial scale and zoning schemes. However, the magnitude and statistical significance of results are sensitive to the spatial scale and zoning scheme of choice, which indicates that the evaluation of the distributional effects of recent transport policies in Rio is sensitive to the modifiable areal unit problem (MAUP). Ideally, the conclusions of policy evaluations should not be dependent on the ad-hoc ways in which the data are spatially aggregated (Kwan & Weber, 2008). This stability in the results is not observed in Rio and this case study illustrates the importance of using sensitivity analysis in the evaluation of transport policies. It is worth mentioning, though, that according to Log Likelihood and AIC tests (Tables 2 and 3), the regressions conducted using hexagonal grids of 0.5 km had the best fit to the data, showing to be more appropriate for the spatial analysis.

5.3 Local association between income and accessibility change

While the spatial Durbin model gives average global results for the entire city, it is important to identify which areas in Rio gained or lost access to opportunities, and how the association between accessibility gains and income varies across space. Figure 6 shows bivariate cluster maps of the local association between household income per capita and the variation in access to jobs and schools observed between 2014 and 2017. Spatial clusters colored light-green are the ‘sweet spot’ from a transport equity point of view as they represent areas of relatively lower income population that had relatively higher accessibility gains. In contrast, red clusters show locations where higher accessibility gains accrued to higher-income groups. Darker green areas are also problematic as they represent clusters of low-income areas which lost accessibility in the period. This high/low classification of clusters is based on the mean of each variable.
The cluster analysis shows for example how the new light-rail system in the city center had little effect on access to schools and jobs. This is partially explained by its small network which is still not properly connected to the rest of the public transport system. On the other hand, the new subway line in the south has mostly improved access to opportunities for higher-income groups (red clusters). The analyses at 0.5 km and traffic zones, nonetheless, indicate that the new subway line has also improved accessibility levels of two poor favelas, Vidigal and Rocinha shown in light green. These favelas, however, are not captured at 2 km resolution. Again, the results demonstrate how analyses at coarse scales lead to information loss and potential issues of ecological fallacy and aggregation error.

Figure 6. Bivariate LISA based spatial clusters showing the local association between average household income per capita and gains in access to jobs and schools by public transport between 2014 and 2017, Rio de Janeiro
Furthermore, there are a few clusters of both high- and low-income people that have gained access due to the new Transcarioca BRT line (Figure 1) and its connection to the subway line. The Transolímpica BRT, on the other hand, has had no significant effect on access to either schools or employment opportunities. This is closely related to the fact that the passenger demand in Transolímpica is only at 41% of the demand originally projected for the investment, which has led the company that runs this corridor to close some stations (Magalhães & Rodrigues, 2017). Finally, there are various clusters of low-income areas that experienced a decrease in accessibility in the urban fringes in the west region of the city. This likely the result of worsening traffic conditions combined with service cuts in those areas.

6 Conclusions

Recently implemented public transport policies in Rio de Janeiro have been evaluated at multiple spatial scales to assess their distributional effects on access to schools and employment opportunities. Overall, the results show that infrastructure investments related to the 2014 World Cup and the 2016 Olympic Games, combined with subsequent cuts in service levels were followed by a small loss in average accessibility levels in Rio. A quasi-counterfactual analysis indicates that the reorganization of bus lines and reduction in services have offset the potential benefits of new transport infrastructure and were just as important as the new investments in reshaping accessibility levels. Moreover, the results show that wealthier areas had on average small but statistically significant higher gains in access to opportunities between 2014 and 2017 than poorer areas. Contrary to the official discourses of transport legacy, there has been little accessibility improvement in the most deprived areas, which are doubly disadvantaged with low income and low accessibility.

It is unrealistic to demand that new transport investments equally benefit every neighborhood in a city. From a social justice point of view, however, the least one would expect is that such governmental policies would improve the transport conditions of those in the worst-off position and not reinforce inequalities (Pereira et al., 2017). Nonetheless, the findings of this study indicate that the transport legacy of recent mega-events in Rio are in direct conflict with the Rawlsian principles of justice and the stated aims in the justification of the investment program. If anything, those policies have violated basic rights with the forced eviction of families and further exacerbated socio-spatial inequalities in access to employment and educational opportunities.

The case of Rio exemplifies how the uneven nature of urban and transport development tends to create place-based advantages for higher income groups that are hard to tackle even with large infrastructure investments. Reducing socio-spatial inequalities in access to opportunities is a challenging task which can only have limited results without a full integration between transport and land-use planning. The austerity measures adopted in Rio’s public transport systems resonate well with the experience of local authorities elsewhere, particularly in developing countries, which often have to invest and manage their transport systems under severe budgetary constraints. The method used in this paper relies mostly on datasets that are organized in standardized formats and commonly available for different cities in the world, allowing for the replicability of this research method in other urban contexts.

This paper advances previous studies that analyze the social effects of transport policies by showing that equity assessment of such policies are sensitive to MAUP. The study takes seriously the issue of spatial frame dependence in the evaluation of the impacts of infrastructure change on accessibility and it shows that the consideration of MAUP is integral to assessing the accessibility impact of a transport policy and how it is distributed. The results of this type of analyses are dependent on the spatial frame that is chosen. While the direction of the association between accessibility gains and income remains positive in every analysis, the magnitude and statistical significance of this association changed considerably when using different spatial disaggregation. Furthermore, the paper has shown that MAUP also affects the spatial identification of which areas have gained and lost accessibility. It also shows how the analysis based on traffic zones has captured the geography of poverty and wealth and its relationship...
with uneven access to opportunities. The statistical tests indicated that the 0.5 km grid was best attuned
to capture the socio-spatial differences of the changes in accessibility brought about by recent policies.
However, it is unclear whether these spatial scale/zonings would still be appropriate if one used different
regression specifications or measured accessibility differently. It may well be possible that some scales and
spatial units could be more appropriate for certain types of activities than others.

The MAUP is a longstanding challenge in spatial analysis and there is yet no simple solution to it.
A potential alternative to circumvent MAUP would be to use space-time accessibility measures, which
are not sensitive to spatial aggregation effects (Kwan & Weber, 2008). Nonetheless, these measures also
have their own limitations, such as being data-hungry and less transparent/easy to communicate to
policymakers. While space-time accessibility could be a promising approach to deal with MAUP, further
research is needed to explore the potentials and limits of this type of measures in the equity assessment
of transport policies. This reinforces the importance of sensitivity analysis in this type of policy evalua-
tion. This type of analysis can be more easily conducted when spatial information of land-use activities is
available as point data, which can be aggregated in various ways to test different spatial scales and zoning
schemes. Even when data is only available in a single ad-hoc spatial aggregation, as is often the case, there
are still ways to run sensitivity analyses using regionalization methods that regroup existing areal units to
produce new sets of spatial aggregation schemes (Xu, Huang, Dong, & Abdel-Aty, 2014).

This paper has focused on physical accessibility and its limitations need careful attention in future
studies. This study has not considered the influence of personal characteristics such as physical dis-
abilities, age or gender and how they can hinder one's ability to use public transportation. As noted
in Section 3, the question of affordability is particularly relevant in Rio, where public transport costs
can be prohibitive for low-income families. Moreover, the results presented are based on timetables of
services, which do not consider non-recurrent congestion levels in Rio. Future research would benefit
from instead using GPS data of vehicles, which would render more accurate travel times for accessibility
estimates (Wessel, Allen, & Farber, 2017). Finally, future research should examine the long-term eco-
nomic and social consequences of such transport developments for Rio's metropolitan area, for example
in terms of air quality, reallocation of jobs, real estate appreciation and gentrification.

A lesson that can be drawn from the case of Rio de Janeiro is that policy makers should incorporate
accessibility scenario analysis in the early phases of transport planning to give a more complete picture
of who benefits from new investments and subsequent changes in service levels. Furthermore, the case
of Rio is not an exception when it comes to the way in which official rhetoric seeks to justify transport
projects based on promised benefits to society as a whole and to low-income groups in particular. It
serves as a cautionary tale about how the potential benefits of new infrastructure projects can be under-
mined when these projects are seen in isolation from rest of the city and its transport system, and when
governments do not sustain the level of spending necessary to maintain and run services at affordable
prices. The recent experience of Rio with the transport investments related to mega-events illustrates
how governments have significant capacity to reshape inequalities of access and opportunities in cities
through the provision of public transport services and investments, but they may not deliver what has
been promised.

Acknowledgements

The authors thank Fetranspor for providing GTFS datasets. This work was funded by Capes Founda-
tion, Ministry of Education, Brazil [Grant Number BEX 1397/13-3]; the Institute for Applied Eco-
nomic Research (Ipea - Brazil) and by the WRI Ross Center for Sustainable Cities [2017 Lee Schipper
Memorial Scholarship].
References


