Pontifícia Universidade Católica do Rio de Janeiro

Maína Celidonio de Campos

Urban Mobility, Inequality and Welfare in Developing Countries: Evidence from 2016 Olympics in Rio de Janeiro

Tese de Doutorado

Thesis presented to the Programa de Pós-graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Economia.

Advisor : Prof. Gabriel Lopes de Ulyssea
Co-advisor: Prof. Juliano Junqueira Assunção

Rio de Janeiro
February 2019

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

Maína Celidonio de Campos

Urban Mobility, Inequality and Welfare in Developing Countries: Evidence from 2016 Olympics in Rio de Janeiro

Thesis presented to the Programa de Pós-graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Economia. Approved by the undersigned Examination Committee.

Prof. Gabriel Lopes de Ulyssea
Advisor
Departamento de Economia – PUC-Rio

Prof. Juliano Junqueira Assunção
Co-advisor
Departamento de Economia – PUC-Rio

Prof. Claudio Ferraz
Departamento de Economia – PUC-Rio

Prof. João Paulo Pessoa
Departamento de Economia – Fundação Getúlio Vargas - São Paulo

Prof. Daniel Da Mata
Departamento de Economia – Fundação Getúlio Vargas - São Paulo

Prof. Ciro Biderman
Escola de Administração Pública e Economia – Fundação Getúlio Vargas - São Paulo

Prof. Augusto Cesar Pinheiro da Silva
Vice Dean of Graduate Studies
Centro de Ciências Sociais – PUC-Rio

Rio de Janeiro, February 25th, 2019

Maína Celidonio de Campos

Completed her Bachelor of Arts degree in Economics from Pontifícia Universidade Católica do Rio de Janeiro in 2009 and obtained her Master of Science degree in Economics from Pontifícia Universidade Católica do Rio de Janeiro in 2011.

Bibliographic data

Celidonio de Campos, Maína

Urban Mobility, Inequality and Welfare in Developing Countries: Evidence from 2016 Olympics in Rio de Janeiro / Maína Celidonio de Campos; advisor: Gabriel Lopes de Ulyssea; co-advisor: Juliano Junqueira Assunção. – Rio de Janeiro: PUC-Rio, Departamento de Economia, 2019.

98 f: il. color. ; 29,7 cm

Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, 2019.

Inclui bibliografia.

1. Economia – Teses. 2. Mobilidade urbana. 3. Desigualdade. 4. Tempo de deslocamento. 5. Transporte público. 6. Avaliação de impacto. I. Ulyssea, Gabriel Lopes de. II. Assunção, Juliano Junqueira. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 330

To my family: Miguel, Diego and Artur.

Acknowledgments

To my advisors Gabriel Ulyssea and Juliano Assunção for their great generosity and encouragement during these five years;

To the members of the Examination Committee, and specially to Claudio Ferraz, for their extremely valuable contributions;

To CNPq and CAPES Foundation Program, since this study was financed in part by the Coordenação de Aperfeiçoamento Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, and CNPq;

To Instituto de Estudos de Política Econômica (Casa das Garças) for hosting me in the last years and for giving me the opportunity to engage in valuable debates;

To my friends and colleagues from the CPI team, Dimitri Szerman and Vitor Pereira, for the valuable contributions;

To Davi Bhering, Gabriel Monteiro and Helena Silberman, for the remarkable dedication and impeccable assistance;

To Bernardo Ribeiro for being an amazing friend and assistant;

To my dear friends, Fernanda Guardado and Amanda Schultz, for all the support and wise critics;

To Clarissa Gandour for all the help, kindness and extremely valuable contributions from the first to the very last day of the PhD;

To Lívia Gouvea and Tassia Cruz, my "sisters in arms", for being there when I needed with kind words and meaningful advices;

To my family, for being an example of love, courage and resilience, and specially to my mother, for the unwavering support;

To Miguel Foguel, my husband and love of my life, for the companionship and joy that help me every step of the way;

For my kids, Diego and Artur, for showing me the sweetest side of life.

Abstract

Celidonio de Campos, Maína; Ulyssea, Gabriel Lopes de (Advisor); Assunção, Juliano Junqueira (Co-Advisor). **Urban Mobility, Inequality and Welfare in Developing Countries: Evidence from 2016 Olympics in Rio de Janeiro**. Rio de Janeiro, 2019. 98p. Tese de doutorado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This dissertation assesses the aggregate and distributional effects of the recent transport infrastructure expansion in Rio de Janeiro (Brazil) triggered by 2014 Football World Cup and 2016 Olympic Games. In preparation for the sports events, the city invested more than 4.5 billion dollars in its public transport system, which included the extension of a subway line, the construction of a light-rail system and two BRT corridors that stretch approximately 108 kilometers. Chapter 1 provides a description of new transport infrastructure and estimates its potential effects on commuting times. I compute travel times in the absence of the investments using random forest regression methodology and data from 2011 and 2018 travel times. Estimates suggest that the new infrastructure significantly reduced travel times. The remaining chapters explore two different methodologies to account for the impacts of the transport investments. Chapter 2 explores the timing of announcement and inauguration of new BRT and subway stations in Rio de Janeiro City to investigate the effects of the expansion of transport infrastructure on growth and reorganization of economic activity. Firm's addresses were geocoded to construct a panel data set that contains information on number of firms and jobs per 100 meter's grid cell from 2006 to 2016. Applying a difference-in-differences methodology on this novel data set, I estimate the heterogeneous effects of the transport expansion according to workers' characteristics and industry. All effects are obtained for eight different distance rings ranging from 250m to 2km. Chapter 3 aims to measure the effects of transportation infrastructure on the city's wages, productivity and welfare, investigating heterogeneous impacts for high and low skilled workers. To answer these questions, I construct an extensive database for the Rio de Janeiro Metropolitan Area that combines information on residence and employment for each skill group inside each city block. In order to measure general equilibrium effects, I develop a model of internal city structure that features heterogeneous workers and production externalities across worker's skill levels. I estimate structural parameters using generalized method of moments. Finally, I perform contrafactual exercises to assess the impacts of the recent transport infrastructure expansion in Rio de Janeiro using 2018 travel times collected from Google Maps API and travel times computed in the first chapter. Results show that connecting new areas to the central business district results in lower residential concentration and higher employment concentration. The improvement of transportation services allows citizens to work in high productivity locations and live in high amenity locations, which leads to higher overall welfare. Nevertheless, benefits are not evenly split. High-skilled workers benefit twice since they have higher benefits from agglomeration and, consequently, they are able to pay for higher residential prices from lower commuting costs. Moreover, areas in the vicinity of the new transport stations saw an increase in economic activity. The bulk of the impact is characterized by small firms, from the commerce and service sectors. Additionally, most of the workforce employed by these firms are low-skilled.

Keywords

Urban mobility; Inequality; Commuting time; Public transportation; Impact evaluation

Resumo

Celidonio de Campos, Maína; Ulyssea, Gabriel Lopes de; Assunção, Juliano Junqueira. **Mobilidade Urbana, Desigualdade e Bem-Estar nos Países em Desenvolvimento: Evidências das Olimpíadas 2016 no Rio de Janeiro**. Rio de Janeiro, 2019. 98p. Tese de Doutorado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação estima os efeitos agregados e distributivos da recente expansão da infraestrutura de transporte no Rio de Janeiro (Brasil), desencadeada pela Copa do Mundo de 2014 e pelos Jogos Olímpicos de 2016. Em preparação para os eventos esportivos, a cidade investiu mais de 4,5 bilhões de dólares em seu sistema de transporte público, que incluiu a extensão de uma linha de metrô, a construção de um VLT e dois corredores de BRT que se estendem por aproximadamente 108 quilômetros. O Capítulo 1 fornece uma descrição da nova infraestrutura de transporte e seus potenciais efeitos nos tempos de deslocamento. Os tempos de deslocamento (na ausência dos investimentos) são calculados usando metodologia de regressão random forest e dados de tempos de deslocamento de 2011 e 2018. As estimativas sugerem que a nova infraestrutura reduziu significativamente os tempos de viagem. Os capítulos restantes exploram duas metodologias diferentes para estimar os impactos dos investimentos em transporte. O Capítulo 2 utiliza as datas de anúncio e inauguração das novas estações de BRT e metrô na cidade do Rio de Janeiro para investigar os efeitos da expansão da infraestrutura de transportes no crescimento e reorganização da atividade econômica. Os endereços das empresas foram georeferenciados para construir um painel com informações sobre número de empresas e empregos por célula de 100 metros quadrados de 2006 a 2016. Aplicando uma metodologia de diferenças em diferenças, eu estimo os efeitos heterogêneos da expansão do transporte de acordo com as características dos trabalhadores e da indústria. Todos os efeitos são obtidos para oito diferentes anéis de distância de 250m a 2km. O Capítulo 3 tem como objetivo medir os efeitos da infraestrutura de transporte sobre os salários, a produtividade e o bem-estar da cidade, investigando impactos heterogêneos para trabalhadores com alto e baixo nível de qualificação. Para responder a essas perguntas, eu construo um extensa base de dados para a Região Metropolitana do Rio de Janeiro, que combina informações sobre residência e emprego para cada grupo de trabalhadores dentro de cada área de ponderação do Censo 2010. Para medir os efeitos de equilíbrio geral, eu desenvolvo um modelo de estrutura interna de cidade que possui trabalhadores heterogêneos e diferentes externalidades de produção para cada grupo de trabalhador. Eu estimo os parâmetros estruturais usando o método de momentos. Por fim, realizo exercícios contrafactuais para avaliar os impactos da recente expansão da infraestrutura de transporte no Rio de Janeiro usando os tempos de viagem de 2018 coletados do API do Google Maps e os tempos de viagem na ausência dos investimentos (computados no primeiro capítulo). Resultados mostram que os investimentos de transporte levaram a menor concentração residencial e maior concentração de empregos. Melhores serviços de transporte permitem que os cidadãos trabalhem em locais de alta produtividade e morem em locais de alta amenidade, o que aumenta o bem-estar de todos os trabalhadores. Entretanto, os benefícios não são divididos igualmente. Os trabalhadores altamente qualificados se beneficiam duplamente, uma vez que têm maiores benefícios de economias de aglomeração e, consequentemente, são capazes de pagar por custos mais altos de moradia. Ademais, as áreas no entorno das novas estações tiveram um aumento na atividade econômica. A maior parte do impacto é caracterizada por pequenas empresas, dos setores de comércio e serviços. Além disso, a maior parte da força de trabalho empregada por essas empresas é pouco qualificada.

Palavras-chave

Mobilidade urbana; Desigualdade; Tempo de deslocamento; Transporte público; Avaliação de impacto;

Table of Contents

List of Figures

List of Tables

Introduction

Urbanization has been a major driving force of recent global development; today half of the world's population lives in cities, representing 70-80% of global GDP. Cities' high densities foster economic growth through reducing transaction costs, diffusing knowledge, and enabling investments in infrastructure and services (World Bank, 2015). In this context, public transportation plays a key role in shaping the consequences of urbanization. Transportation connects individuals with jobs and services, reducing spatial frictions and promoting socioeconomic inclusion. In particular, it facilitates the separation of workplace and residence, allowing people to work in high-productivity locations and live in high-amenity locations. This means that individuals can choose their place of work optimally, matching with high productivity jobs and taking advantage of positive production externalities via agglomeration effects. At the same time, individuals have more degrees of freedom when choosing their place of residence, which leads to higher overall welfare.

Although urbanization has the potential to foster economic growth and generate prosperity, increased urbanization has been closely associated with rising inequality and exclusion within cities. Urban poverty has been increasing rapidly, especially since most of this urbanization process has been taking place in poor countries in Africa, Asia and Latin America. In highly segregated cities, urban dwellers are denied access to basic rights, such as education and health, which creates further barriers to human capital accumulation. Additionally, a large segment of the lower income population faces long and costly commutes to work. As a result, workers with limited access to job centers may opt for lower quality employment to reduce their commuting costs. Therefore, commuting costs not only prevent cities from fully seizing the benefits of agglomeration economies, but can also reinforce inequality.

By 2050, with an expected 2.5 billion people moving to cities, mostly in developing countries, demand for transport services will increase dramatically. To meet these demands and curb congestion, governments must spend vast sums investing in public transportation systems. Although the literature provides robust estimates on the impacts of transport infrastructure, little is known about distributional effects of such investments in developing countries.

Thus, it is paramount to estimate the impacts of such investments and to uncover heterogeneous and distributional effects.

This dissertation sheds light on these issues by estimating the effects of a major transport infrastructure expansion in Rio de Janeiro (Brazil) across several dimensions of heterogeneity. Rio de Janeiro constitutes a unique empirical setting to analyze the impacts of transport infrastructure and how they affect segregation and inequality. First, the city underwent a major expansion of public transport infrastructure, which is a rare episode in a developing country. In preparation for the 2014 World Cup and 2016 Olympic Games, Rio invested more than 4.5 billion dollars in its public transportation system. The investments included the extension of a subway line, the construction of a light rail system (LRT) and two Bus Rapid Transit (BRT) corridors that stretch approximately 118 kilometers. Second, Rio, with more than 11 million inhabitants, is the second largest metropolitan area in Brazil and the third in Latin America. Third, the city is marked by high inequality: the income of the top 20% is 17 times the income of the bottom 20%. Fourth, Rio has the highest average round trip commuting time in the country: 1 hour and 40 minutes. Finally, there is a unique availability of high quality and detailed data.

In order to estimate distributional effects, I construct a unique data set that combines different sources of micro data at the individual, firm, and city-block levels. At the individual level, I use 2010 Brazilian Demographic Census, which contains information on socio-demographic characteristics, labor market outcomes, and place of residence. On the firm side, I use restricted access, administrative data from the Ministry of Labor, the RAIS data set. This data set contains all formal firms and workers in Brazil and firm-level information on address, number of workers, and industry, in addition to employee educational level. Baseline travel time information comes from a restricted access, origin-destination survey of the state of Rio de Janeiro (2011). Finally, I collect endline travel time information using Google Maps API (2018).

This comprehensive and original data set allows me to estimate the transport investments impacts on a range of outcomes. In Chapter 1, I assess the effects of the new BRT, LRT and subway stations on travel times. I combine information from origin-destination survey (2011) and Google Maps API (2018) to build a travel time panel data set. Using random forest regression, I investigate how much of the travel time difference is explained by the transport stations opened between 2011 and 2018. Then I predict counterfactual travel times in five scenarios - without BRT, without the subway extension, without LRT, without BRT and the subway extension, and no investments - and compare each with the 2018 travel times. Results indicate that transport investments had a relevant impact on travel times. In the absence of all of these investments, average commuting times would have increased by 45 minutes[1]. Comparing each transportation mode, BRT has the highest effect on reducing travel time (20 minutes), followed by LRT (8 minutes) and the subway extension (3 minutes). More important, these results provide evidence of investment complementarities: the total effect is higher than the sum of partial effects. This is expected since the goal of the Olympics Plan's was to create a transportation ring connecting different parts of the city.

Chapter 2 explores the timing of inauguration of new BRT, subway, and LRT stations to investigate the effects of the expansion of transport infrastructure on the number of firms, jobs and average wage in a 2-kilometer radius from stations. I also estimate the impacts of the announcement that Rio de Janeiro was selected as Olympics' host city. The firms' addresses were geocoded to construct a panel data set that contains information on the number of firms, jobs, and average wages per 100-meter grid cell from 2006 to 2016. Then, I use a difference-in-differences methodology to estimate the impacts of each transportation mode across eight rings of distance rings of distance from the stations, beginning at 250m up to 2km (in 250m intervals).

Results show evidence of relevant heterogeneous effects. First, the announcement and inauguration of subway and BRT stations have positive and significant impacts on the number of firms, jobs, and average wage; while LRT has null or negative impacts. Since LRT was constructed in the central business district, and I only observe the first year since inauguration, these results may reflect short-term displacement effects. Second, the announcement impacts are larger than inauguration impacts, evincing that firms anticipate the effects of new transport infrastructure. Third, subway stations inauguration effects are larger and have greater geographical reach. It is important to highlight that these heterogeneous effects between BRT and subway may be due to an ex-ante difference among the treatment groups and not because of the transport technologies alone. If transportation infrastructure is complementary to other urban infrastructure, the effects might be non-linear in the initial density of economic activity. Since urban infrastructure and employment density are correlated, effects might be higher in denser areas. In fact, the subway treatment group has much higher employment density in baseline. Additionally, new subway and BRT stations had impact in grids with zero employment in baseline, which indicates that investments led to city sprawl, specially for BRT stations.

[1] Unweighted in-sample average

Concerning heterogeneous effects according to workers' characteristics, impacts on number of jobs are stronger for workers with up to high school level education in comparison with workers with a college degree. Besides, all workers experience the same proportional change in wages, which may be a result of selection or agglomeration externalities. Finally, the bulk of the positive effects on the number of firms come from small firms (2 to 10 employees) and from the service and commerce sectors.

In Chapter 3, I estimate the general equilibrium effects of transportation infrastructure expansion on wages, employment, inequality, productivity and welfare. Even though the first approach uncovered important heterogeneous results, it does not allow me to infer overall effects on inequality and welfare. Hence, I develop a model of internal city structure that features high- and low-skilled workers, production and residential externalities, and heterogeneous city blocks. The model builds on a recent quantitative urban model Ahlfeldt et al. (2015) and extends it to include heterogeneous workers. Although the model remains fairly simple after introducing heterogeneous workers, it can account for important features of city structure linking inequality and spatial segregation.

The source of inequality in the model is the existence of agglomeration externalities specific to high-skilled workers. Due to this additional agglomeration force, high-skilled workers yield higher wages due to larger productivity gains from agglomeration. In turn, this agglomeration force can impact segregation through two mechanisms. First, high-skill jobs will be more geographically concentrated. Thus, depending on the transport infrastructure, residence decision may also be more geographically concentrated around workplace in order to diminish commuting costs. Second, higher agglomeration force leads to higher wage inequality. Since high- and low-skilled workers bid for floor space, high-skilled workers will concentrate in high amenities residence locations. Higher prices will push low-skilled workers out of these locations, a phenomenon known as gentrification. In the presence of poor transport infrastructure, commuting costs increase rapidly with distance. This exacerbates both mechanisms and gives rise to highly-segregated cities, where high-skilled workers agglomerate close to the city center and low-skilled workers live in the outskirts of the metropolitan area. This configuration is common among developing world metropolises, including Rio de Janeiro.

To estimate the structural model, I combine information on residence and employment for each skill group inside each city block in 2010. Additionally, commuting times between all city blocks (57,122 combinations) are computed using random forest regression and data from restricted access origin-destination survey (2011). Structural parameters are determined according to a three-step estimation procedure that involves calibration, generalized method of moments and grid search. Estimated parameters indicate large productivity gains from agglomeration for all workers and even larger for high-skilled workers. Besides results suggest that exogenous characteristics, such as proximity to the beach, are much more relevant for total amenities than agglomeration forces, represented by residential density.

The estimated model is then used to perform counterfactual exercises to assess the impacts of the recent transport infrastructure expansion in Rio de Janeiro. In particular, I compare equilibrium outcomes using 2018 Google Maps and counterfactual travel times (estimated in Chapter 1) to infer the overall and distributional effects of transport investments. Results point that transport investments increased the welfare of high- and low-skilled workers. Nevertheless, high-skilled workers experienced a larger increase, raising inequality. The expansion of the transport infrastructure connected new locations with Rio's central business district, which increased the number of residential options with lower commuting costs. This led to a reduction in the concentration of residents and increased the concentration of jobs. Both effects are stronger for high-skilled workers, which raises residential and employment segregation. In particular, a gentrification process took place: higher demand and prices in the newly connected area led to an increase in residential segregation. This process was exacerbated by the fact that, among the newly connected areas, some had high amenities due to their proximity to the beach.

The evidence drawn from the three chapters suggests that connecting new areas to the central business district results in lower residential concentration and higher employment concentration. As mentioned earlier, the improvement of transportation services allows citizens to work in high productivity locations and live in high amenity locations, which leads to higher overall welfare. Although the pie grows,the benefits are not evenly split. High-skilled workers benefit twice since they have higher benefits from agglomeration and, consequently, they are able to pay for higher prices from lower commuting costs. Moreover, due to the sprawl of residents, newly connected areas saw an increase in economic activity. The bulk of the impact is characterized by small firms, from the commerce and service sectors. Additionally, most of the workforce employed by these firms are low-skilled.

These results are related to the literature strand that examines the intra-city effects of urban transport interventions. Specifically, Chapter 2 relates to the reduced forms approaches. Perdomo (2011), Rodríguez and Mojica (2009) and Martinez et al. (2018) find evidence of positive impact in residential prices in the vicinities of BRT systems. Bocarejo et al. (2014) show that areas served by BRT have higher population growth than areas without access to the system and Scholl et al. (2018a) estimate positive effects on employment outcomes for individuals living in the most vulnerable areas in baseline. Concerning LRT and the subway, Gibbons and Machin (2005), Billings (2011) and Dorna G (2017) use difference-in-differences approaches and find positive impacts in housing prices in areas close to the system. Recently, Gonzalez-Navarro and Turner (2018) investigated the relationship between the extent of a city's subway network, its population, and its spatial configuration for the 632 largest cities in the world. Although they find that subways have an economically insignificant effect on population growth, they also show subways cause cities to decentralize. In relation to firm size, Atack et al. (2008) find that the introduction of the railroad in the 1850's led to an increase in establishment size in manufacturing. I contribute to this literature by investigating the effects of different transportation modes in the same framework, which allows for direct comparison between modes. Besides, I estimate effects across several dimensions of heterogeneity at a fine grid.

In turn, my dissertation dialogs with the stream of literature that uses general equilibrium models to asses aggregate and distributional effects of transportation improvements (Ahlfeldt et al. (2015), Redding and Rossi-Hansberg (2017)). With a similar approach to Chapter 3, Tsivanidis (2018) looks at the aggregate and distributional effects of TransMilenio, Bogota's BRT system. Also based on Ahlfeldt et al. (2015), the model introduces multiple types of workers by incorporating multiple types of firms with different demand for worker groups. The author finds that while the system caused increases in welfare and output larger than its cost, gains accrued slightly more to high-skilled workers. Results suggest an increase in residential segregation by skills. As already mentioned, I contribute to this literature by estimation the effects of different transportation technologies in the same framework. Besides, differently from Tsivanidis (2018), I observe workers' educational level, which enables me to directly introduce heterogeneous workers in the model and estimate how agglomeration forces contribute to wage inequality. Thus my model incorporates the main mechanism that describes how city growth results in higher inequality and spatial segregation. Hence, my dissertation is also related to the broader literature on the nature and sources of agglomeration economies, as reviewed in Behrens and Robert-Nicoud (2015).

Chapter 1
Rio de Janeiro Metropolitan Area: the transport intervention and its consequences in commuting times

1.1
Introduction

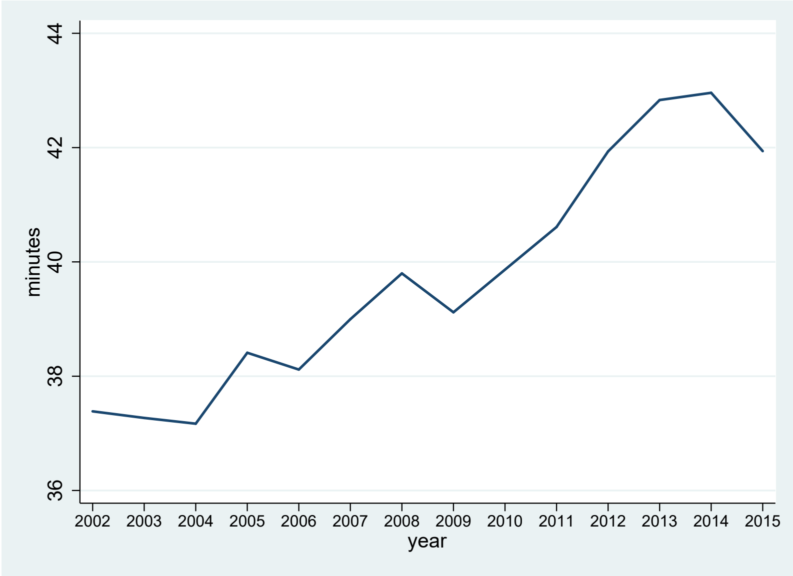
Rio de Janeiro is the second largest metropolitan area in Brazil and the third largest in Latin America. With more than 12 million inhabitants (PNAD 2015), Rio is marked by high income inequality. Between 2000 and 2010, despite the significant reduction in income inequality in the country overall, the Gini index increased in Rio and São Paulo (Ipeadata). The income of the top 20% of inhabitants in Rio is 17 times larger than that of the bottom 20%; in São Paulo, the difference is 14 times larger (Datasus 2012). In 2018, Rio de Janeiro Metropolitan Area ranked as the 18th urban agglomeration with the highest income inequality in the world (UN 2018).

Additionally, the metropolitan area has the highest commuting time in Brazil: average one-way commute reached 49 minutes in 2015 (Figure 1.1 and Figure 1.2). [1] Consequently, to host the 2016 Olympics Games, Rio de Janeiro underwent a major expansion of public transport infrastructure, investing more than 4.5 billion dollars in its public transport system. Thus Rio de Janeiro constitutes a unique empirical setting to analyze the impacts of transport infrastructure and how they affect segregation and inequality.

To understand how the transportation investments translates into travel times, I propose a methodology to predict counterfactual travel times. First, I combine information from origin-destination survey (2011) and Google Maps API (2018) to build a travel time panel data set. Second, I use random forest regression to estimate how much of the travel time difference is explained by the transport stations inaugurated between 2011 and 2018. Then I predict counterfactual travel times in five scenarios - without BRT, without the subway extension, without LRT, without BRT and the subway extension, and no investments - and compare it with the 2018 travel times. Results indicate that transport investments had a relevant impact. Comparing each transportation

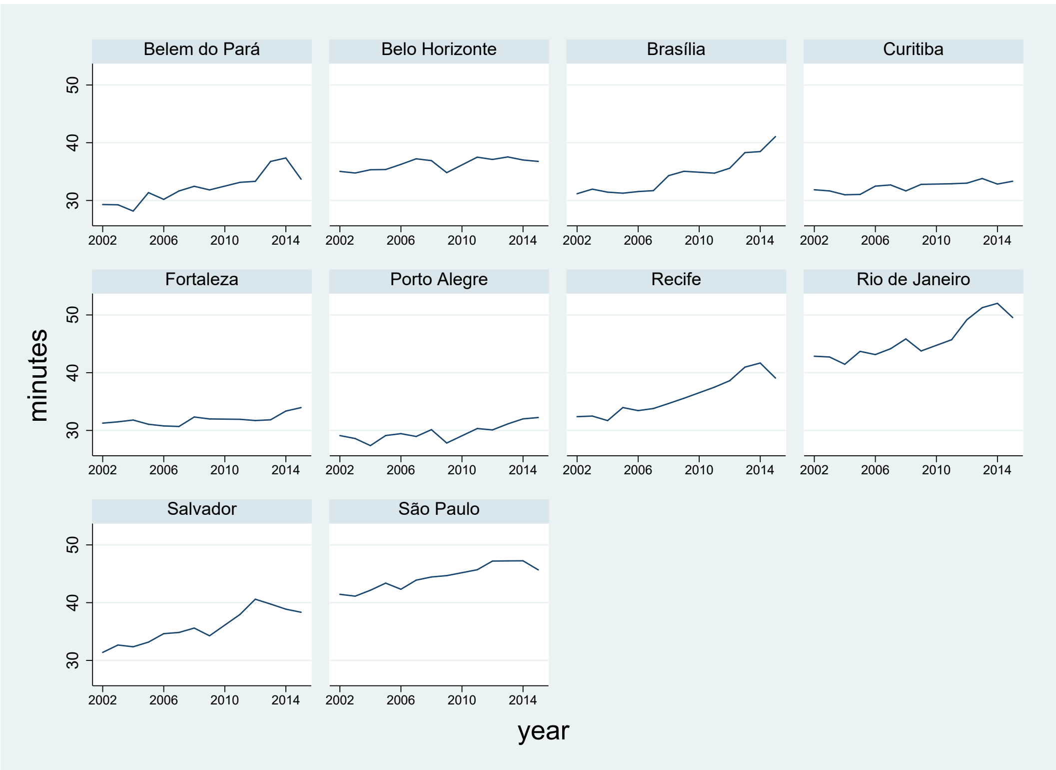
[1] Comparing commuting time in 74 major metropolitan areas around the globe, Rio de Janeiro came in last according to Moovit data App.

Figure 1.1: Commuting time: Metropolitan Brazil (2002-2015)



Notes: Data is from 2002 to 2015 PNAD (Pesquisa Nacional de Amostra de Domicílios), IBGE. I exclude observations in rural tracks.

Figure 1.2: Commuting time: Metropolitan areas (2002-2015)



Notes: Data is from 2002 to 2015 PNAD (Pesquisa Nacional de Amostra de Domicílios), IBGE.

mode, BRT has the highest effect, followed by LRT and the subway extension. More important, results provide evidence of investments complementarities:

Figure 1.3: Rio de Janeiro Metropolitan Area: municipalities

Notes: Map shows Rio de Janeiro Metropolitan Area municipalities. The Metropolitan Area is defined according to 2010 Census.

the total effect is higher than the sum of partial effects.

The chapter proceeds as following. Next section gives some background on the metropolitan area and present key stylized facts. Section 3 describes the new transport infrastructure. Section 4 presents the estimation method and counterfactual analyses.

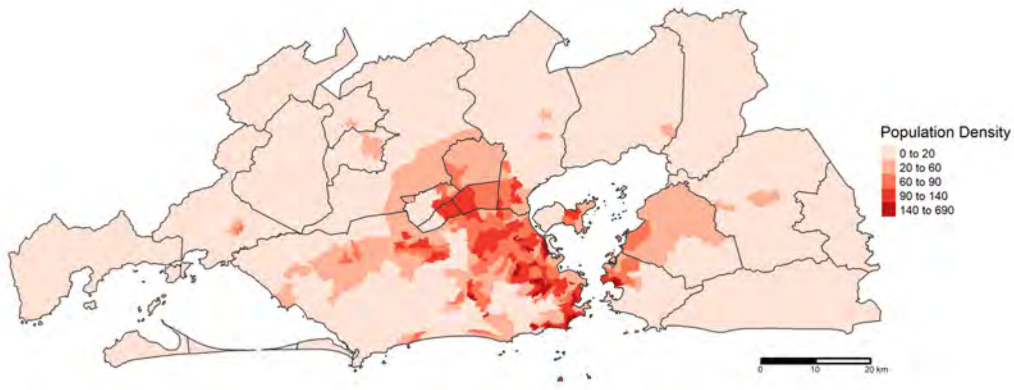
1.2 Background

Rio de Janeiro Metropolitan Area (RJMA) is comprised of 20 municipalities[2] (Figure 1.3), where the city of Rio de Janeiro represents more than half of the total population. Figure 1.4 and Figure 1.5 present the residence and employment distribution. The maps show evidence of the residence and employment high-density areas. In particular, Figure 1.5 identifies the central business district, which is the downtown neighborhood in Rio city. The neighborhood is home to 665,000 formal jobs, 18% of formal employment in the Metropolitan Area.

These concentrations arise due to agglomeration forces: the high density of workers and residents generates positive externalities. More specifically, high densities result in productivity gains for firms, wage gains for workers, and higher amenities for the residents. The mechanisms that explain this agglomeration effect can be classified into three categories: sharing, firm-worker matching, and learning. The high density of people and firms generates gains through the sharing of high fixed-cost indivisible goods. For example, sharing enables the construction of public goods such as universities, parks, hospitals

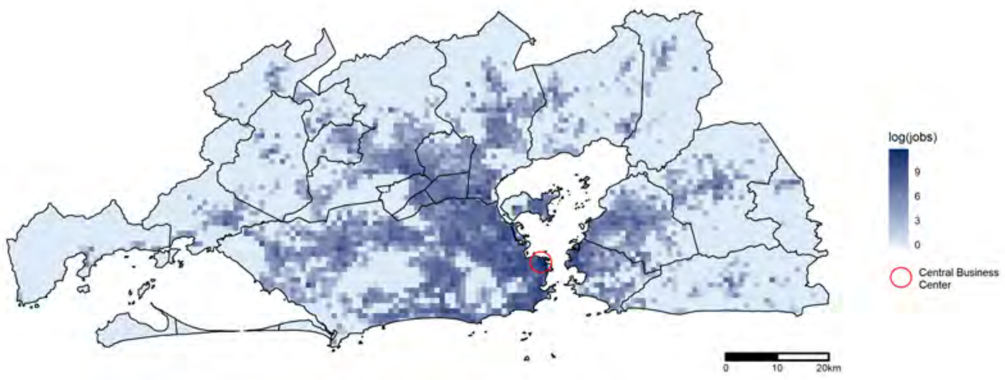
[2] I adopt the Census 2010 definition of the Rio de Janeiro Metropolitan Area.

Figure 1.4: Rio de Janeiro Metropolitan Area: population density



Notes: Data is from 2010 Census. The information is presented by 2010 census statistical areas. The lines represent the municipalities borders.

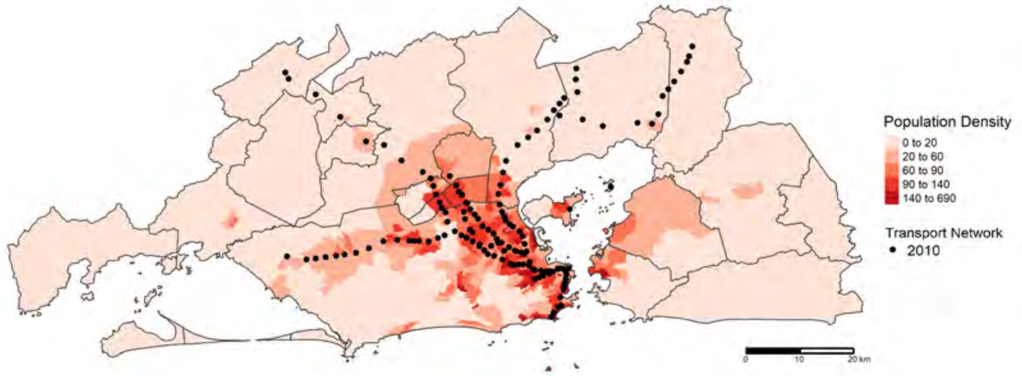
Figure 1.5: Rio de Janeiro Metropolitan Area: jobs



Notes: Data is from 2010 RAIS. Each firm adress was geocoded and matched to 1 kilometer square grid shapefile. In the original data set, firms' adresses were geocoded and matched to 100 meters grid distance cell. Nevertheless, due to vizualization dificulties, I plot data per 1 kilometer grid distance cell.

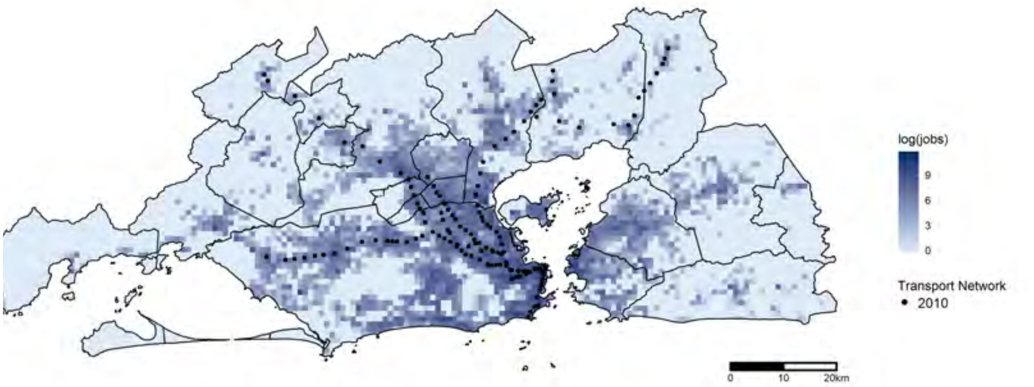
and transportation infrastructure. The concentration of economic activity also facilitates the matching between firms and workers. A higher concentration of jobs increases the likelihood of matching and the quality of it. Finally, agglomeration facilitates learning since it allows for more interactions among workers and, consequently, the diffusion of innovations, ideas, and exchange of experiences Duranton and Puga (2004). In relation to the sharing mechanism, Figure 1.6 and Figure 1.7 highlight the overlay between residents and jobs location, and transport infrastructure in 2010. Causality goes both ways. Firms and residents locate near transport stations to diminish commuting costs. At the same time, transport infrastructure is placed in areas of higher demand. In Chapter 2, I disentangle the two mechanisms and estimate the causal effect of transport infrastructure on economic activity in the surroundings of the new

Figure 1.6: Rio de Janeiro Metropolitan Area: population density and transport infrastructure



Notes: Population data is from 2010 Census. The information is presented by 2010 census statistical areas. The lines represent the municipalities borders. Transport stations shapefile is from Data Rio, Instituto Pereira Passos, Rio de Janeiro Municipality.

Figure 1.7: Rio de Janeiro Metropolitan Area: jobs and transport infrastructure

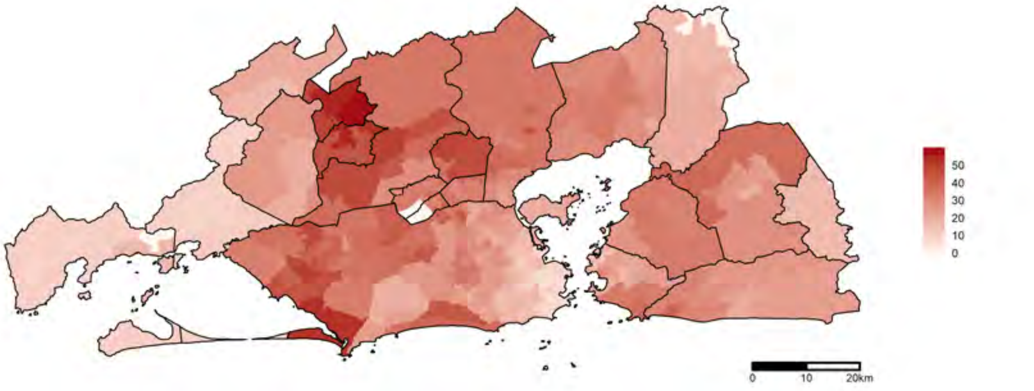


Notes: Emploment data is from 2010 RAIS. Each firm adress was geocoded and matched to 1 kilometer square grid shapefile. In the original data set, firms' adresses were geocoded and matched to 100 meters grid distance cell. Nevertheless, due to vizualization dificulties, I plot data per 1 kilometer grid distance cell.Transport stations shapefile is from Data Rio, Instituto Pereira Passos, Rio de Janeiro Municipality.

stations.

As a result of this spatial distribution, 55% of all workers between 18 and 70 years old live in the city of Rio de Janeiro, and 66% work in it (Censo 2010). This distribution means that each day 3.5 million workers commute within the city, and 18% come from outside the city. Figure 1.8 shows the percentage of workers that commute for more than an hour per residence area (one-way trip). Data shows a clear pattern: commuting time increases with distance from the central business district. In the fringes of the metropolitan area, the percentage of workers with long commutes diminishes, which indicates different commuting behavior after a certain distance threshold.

Figure 1.8: Percentage of workers that commute for more than na hour by residence location

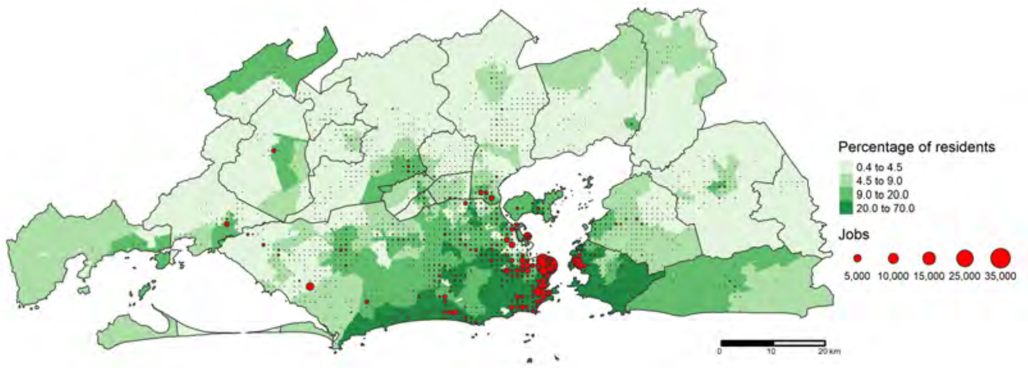


Notes: Data is from 2010 Census. The information is presented by 2010 census statistical areas. The lines represent the municipalities borders.

To demonstrate the link between inequality and commuting costs, Figures 1.9 and 1.10 combine data on residence and employment location separately for high- and low-skilled workers. Figure 1.9 presents the proportion of the working-age population, up to high school education, per residence location and the number of formal jobs occupied by workers at this educational level. Figure 1.10 plots the same variables for workers with a college educational level. The residence location represents the distribution of the potential labor supply in the metropolitan area, while the employment location represents the formal labor demand. Three facts stand out. First, the metropolitan area is characterized by high segregation levels. The high-skilled population is concentrated around the central business district and along the seafront. The location suggests that the optimal choice for high-skilled workers is to live close to the workplace (business district) and in high amenity neighborhoods, such as the seafront.

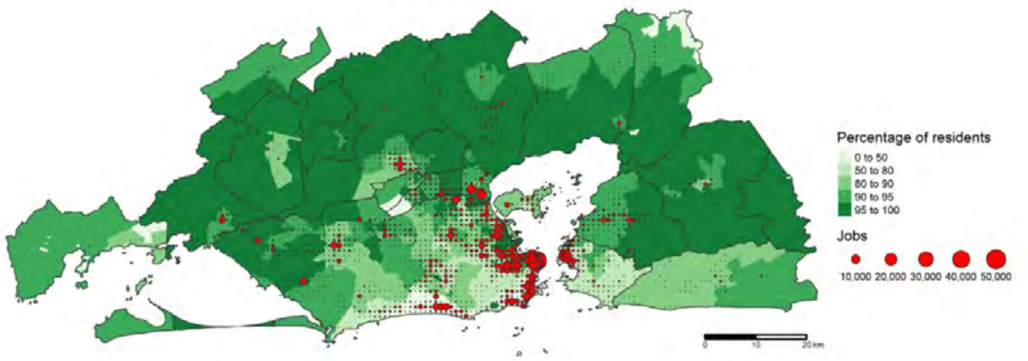
Second, the low-skilled population is subject to spatial mismatch. The majority of jobs is located in the eastern part of the Rio de Janeiro City, while the population concentrates in the opposite zone. It is important to highlight that this points to larger potential commuting costs for the low-skilled population, not actual commuting times. The potential distance between the place of residence and place of formal employment transforms into effective commuting time if the resident decides to participate in the labor market, is employed and has a formal job. In this regard, Gutiérrez-I-Puigarnau and van Ommeren (2010), Gershenson (2013) and Black et al. (2014) provide evidence that commuting times negatively affect labor supply.

Figure 1.9: Workers up to high school educational level



Notes: Residents data is from 2010 Census and Jobs data is from 2010 RAIS.

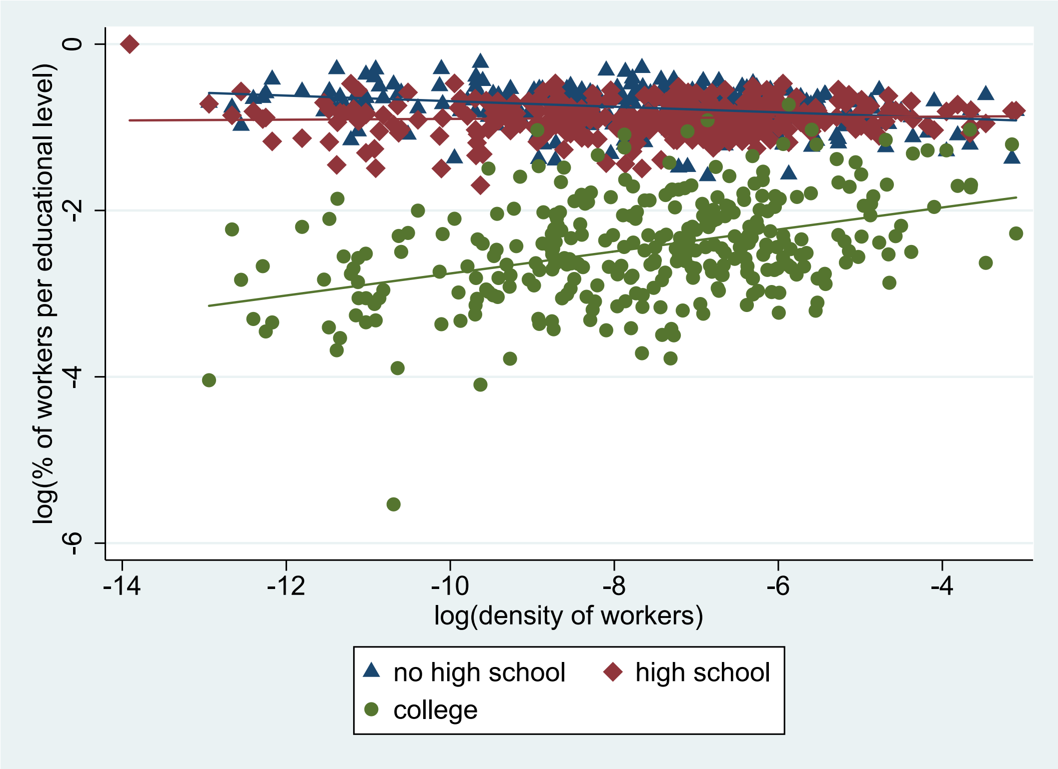
Figure 1.10: Workers with college educational level



Notes: Residents data is from 2010 Census and Jobs data is from 2010 RAIS.

Third, high-skilled jobs are more geographically concentrated than low-skilled jobs. There is evidence that agglomeration forces are stronger for high-skilled workers because high density generates more pronounced learning effects when worker skill level is higher Behrens and Robert-Nicoud (2015),Baum-Snow and Kahn (2000). In this sense, Figure 1.11 shows that the percentage of workers with a college degree increases with the density of jobs. In particular, a 1% increase in the density of jobs is associated with a 13% increase in the percentage of workers with full college education. The other educational groups do not follow the same pattern. Additionally, Figure 1.12 indicates that wage gains from agglomeration are higher for workers with college education.

Figure 1.11: Employment density and porcentage of workers by educational level



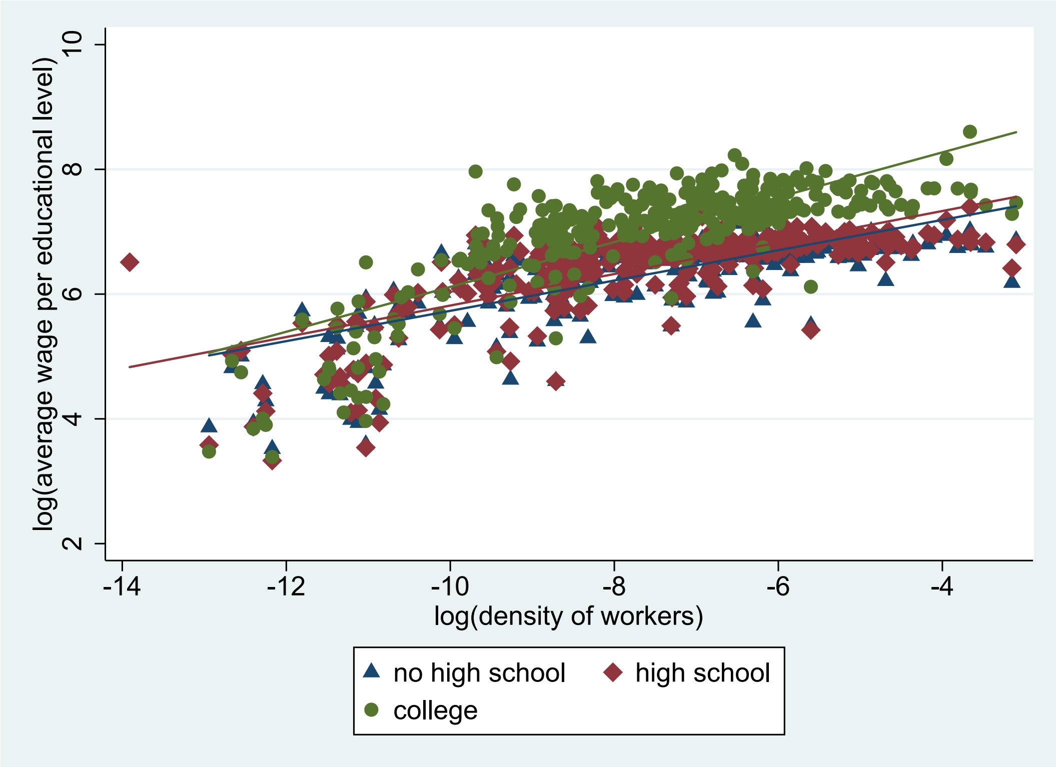
Notes: Data is from 2010 RAIS.

1.3 Transportation Investments

In preparation for the sports events, Rio invested more than 4.5 billion dollars in its public transportation system. The investments included the extension of a subway line, the construction of a light rail transit (LRT) system and three Bus Rapid Transit (BRT) corridors that stretch approximately 122 kilometers. The Transport Strategic Plan for the Rio 2016 Olympic and Paraolympic Games was delivered in 2009, as part of the bid documents that the candidate cities had to submit. In October 2009, Rio was announced as the host city. The official transportation plan was updated in 2012 to include the subway extension and LRT. Nevertheless, updated plans were already known in 2010 and, except for the LRT, construction began in the same year.

The transportation plans had two goals. First, they aimed to provide the city with a High Performance Transport Ring that connected all of the Olympic zones (Deodoro, Barra, Copacabana, Maracanã, Port area) with public and accessible transport. Second, the goals aimed to assure the legacy of the Games through the public transport infrastructure. According to the authorities, the City of Rio de Janeiro took the opportunity of this mega-event to create a long-

Figure 1.12: Employment density and wages by educational level

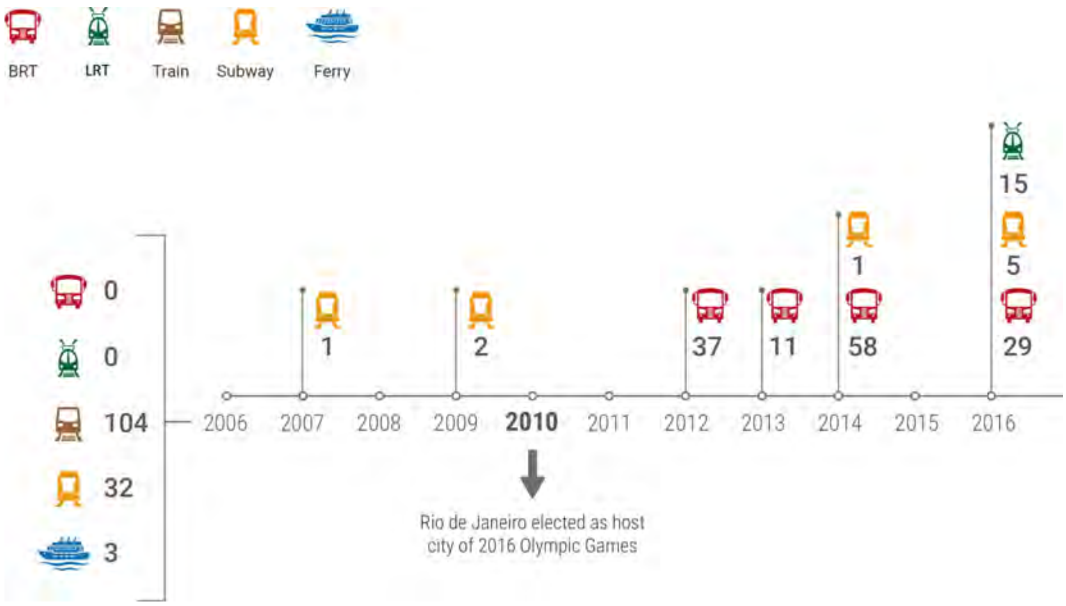


Notes: Data is from 2010 RAIS.

term public transport infrastructure for its citizens Silva and Torres (2013) .

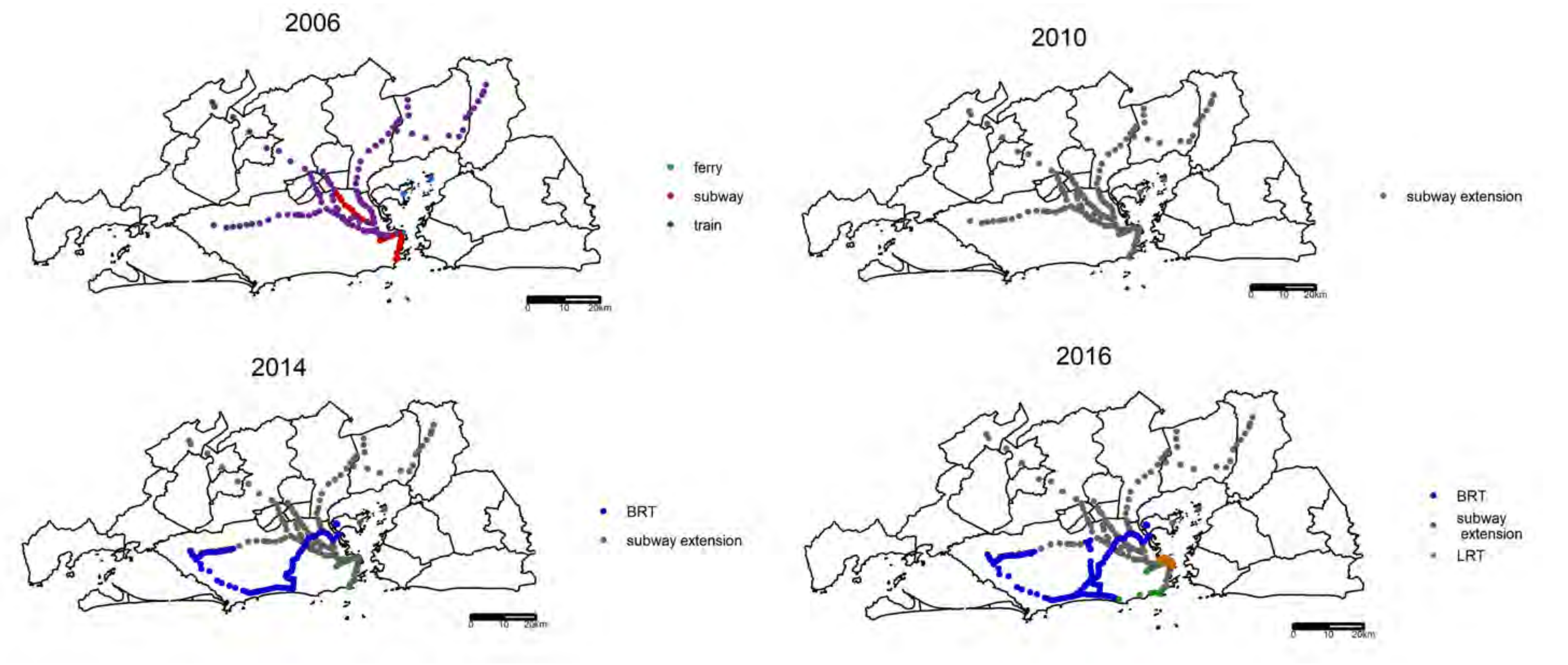
Figures 1.13 and 1.14 describe the evolution of the transportation network. In 2006, the Metropolitan Area had 104 train stations, 32 subway stations and 3 ferry stations. In the period before the Olympic plans (2006-2010), 3 subway stations were opened: one in 2007 and two in 2009. Construction of these stations began in the 1980's, exemplifying the stagnant state of transportation investments. The subway station that opened in 2014 also dates back to the 1980's and cannot be attributed to the Olympic Games.

Figure 1.13: Transport infrastructure timeline



Notes: Data is from Data Rio, Instituto Pereira Passos, Rio de Janeiro Municipality.

Figure 1.14: Transport infrastructure: 2006, 2010, 2014, 2016



Notes: Data is from Data Rio, Instituto Pereira Passos, Rio de Janeiro Municipality.

Directly linked to the Olympic plans, Subway Line 4 was inaugurated in 2016. This highly criticized line is composed of five stations, connecting the south zone to the Barra neighborhood. In the first Olympics plan, a BRT line would make the connection, which is a much cheaper technology. Besides, the construction of the subway line connecting downtown Rio and Niteroi, which wasannounced in 2011, was never started even though it was considered a higher priority. In fact, Rio-Niteroi is the route with the highest number of commuters per day in the Metropolitan Area (Origin-destination Survey 2011).

The infrastructure expansion of the mass transit introduced two new modes of transportation: BRT and LRT. There are four essential features that define BRT: dedicated bus lanes, off-board fare collection, prohibited turns for traffic across the bus lane and platform level-boarding (ITDP 2018). These features make BRT as similar as possible to a subway line, significantly reducing travel times and increasing transport reliability. From 2012 to 2016, 135 stations opened [3].

The construction of the LRT started in September 2014. The system is composed of two lines: one started its operation right before the Olympic Games, and the other was only inaugurated in 2017. This new transportation mode was part of the urban revitalization process that took place in the Port Area. The project called "Porto Maravilha" included the removal of an elevated highway, construction of a new tunnel, public subsidies to induce residential building projects, the construction of new museums and an aquarium. Due to these numerous initiatives, it not possible to link any change in the area exclusively with the LRT.

1.4
Quantifying the effects of transport investments on travel times

Induced by the sports events, the investments brought about a massive expansion in the transport infrastructure. The aim of this section is to measure how this new infrastructure affected travel times. I propose a methodology that uses travel time data from before and after the intervention and machine learning techniques to estimate counterfactual travel times in the absence of the investments. Next, I describe the data sets, estimation procedure and results.

In order to construct a panel data set of trips within the Rio de Janeiro Metropolitan Area, I combine two sources of travel time information: 2011 origin-destination survey and Google maps API. Baseline information

[3] The construction of the Transbrasil corridor was interrupted and the corresponding 26 stations never opened.

Figure 1.15: Origin-destination survey sampling frame

Notes: Data is from 2011 Origin-Destination Survey, Rio de Janeiro State.

comes from restricted access 2011 origin-destination survey. This household survey collected information on trips made within the Metropolitan Area by individuals with 10 years old or more. The sample consists of 3,600 households in the Rio de Janeiro Metropolitan Area, distributed within 730 traffic zones[4] (Figure 1.15). In each household, individuals listed all trips made in the previous day, and provided information on origin and destination zone, time of departure, transportation mode, commuting time and costs. Additionally, individual and household characteristics were collected, such as age, sex, income and car ownership. The final data set contains approximately 13,000 trips. Table 8 shows the sample distribution by origin and destination municipality.

Endline information was collected using Google Maps API. The service allows for 2.5 thousands free searches per day. The search parameters are: origin and destination addresses (or geographical coordinates), transportation mode, departure or arrival time. There are four types of transportation mode: driving, public, walking, and bicycling. I restrict this exercise for trips made by public transportation or driving since I suppose that travel time by non-motorized modes (walk and bike) were not affected by transportation investments. To make the baseline and endline information compatible, I define Google maps search parameters according to trip characteristics from origin-destination survey. I use the coordinates of origin and destination traffic zones centroids as a proxy for origin and destination coordinates. Departure time in Google maps is the same as declared in the survey. Trips made by bus, train, subway and ferry are set as public transportation mode in Google search. And trips

[4]Traffic zones were defined according to 2000 and 2010 Census zoning; and the sampling frame from 2002 origin-destination surveys for the metropolitan area.

Figure 1.16: Ilustrative example of measurement error

Notes: Author's elaboration.

made by car, taxi and mototaxi are set as driving mode. The final data set consists of 9,206 observations with commuting time for 2011 and 2018.

The time difference between 2011 and 2018 can be a result of several factors. First, information comes from different sources. In particular, the origin-destination survey collects the declared commuting time, which may be subject to response bias. Second, origin and destination coordinates used in Google Maps search are proxies for the origin and destination addresses, which may introduce measurement error. Figure 1.16 illustrates the two possible cases of measurement error.

For example, suppose I observe a trip from zone 1 to zone 2 in the origin-destination survey. Then trip length is set to equal the distance between zone's centroids: points C1 and C2. If the true origin and destination points are A and B, the distance C1-C2 is an underestimate. On the other hand, if origin is the point C and destination is point D, my proxy distance overestimates the trip's length. In both cases, the magnitude of the measurement error is positively correlated with traffic zones' size. Consequently, measurement error is correlated with the straight distance between origin and destination centroids.

Third, difference may be due to idiosyncratic reasons connected to the exact day that the interview was held in the origin-destination survey. Since interviewers asked households to list the trips made in the previous day, declared commuting time are subject to idiosyncratic fluctuations. In the case of Google maps, this is not a concern because the API reports the best prediction based on historical averages.

Finally, the time difference may be a result of transportation investments: new BRT, LRT and subway stations may have reduced commuting times by public transportation modes. At the same time, driving commuting times may also have changed due to new commuter behavior. In particular, if public commuting times were reduced, some commuters may have decided to change transportation mode from car to BRT, which diminishes road congestion and affects transit. Thus I model the time difference as a function of the distance between origin and destination centroid's zones, the distance from origin and destination to the new stations, and an idiosyncratic term, as presented below.

$$Y_i = (T_i^{2018} - E[T_i^{2018}]) - (T_i^{2011} - E[T_i^{2011}]) \tag{1-1}$$

$$Y_i = f(dist_i, O_{BRT,i}, O_{SUB,i}, O_{LRT,i}, D_{BRT,i}, D_{SUB,i}, D_{LRT,i}, \varepsilon_i) \tag{1-2}$$

The term $(T_i^{2018} - E(T_i^{2018})) - (T_i^{2011} - E(T_i^{2011})$ corresponds to the demeaned travel time difference between 2011 and 2018 for route i. I subtract the mean for each year to account for any systematic differences between the two data sources. The variable *distance* is the straight line distance between origin and destination. With the inclusion of this variable, I address the correlation between measurement error and the size of the origin and destination traffic zones. The variables O_{BRT}, O_{SUB} and O_{LRT} represent the straight line distance from the origin to the closest BRT, LRT and new subway station, respectively. Analogously, D_{BRT}, D_{SUB} and D_{LRT} represent the straight line distance from the destination to the closest BRT, LRT and new subway station.

Since the objective of this exercise is to estimate travel times in the absence of transport investment, I rewrite equation (1) using categorical variables to describe the distance between origin and destination points and stations.

$$Y_i = f(dist_i, \sum_r \Gamma^r_{BRT,i}, \sum_r \Gamma^r_{SUB,i}, \sum_r \Gamma^r_{LRT,i}, \sum_r \Lambda^r_{BRT,i}, \sum_r \Lambda^r_{SUB,i}, \sum_r \Lambda^r_{LRT,i}, \varepsilon_i), \tag{1-3}$$

where the variables $\sum_r \Gamma^r_{BRT,i}$, $\sum_r \Gamma^r_{SUB,i}$ and $\sum_r \Gamma^r_{LRT,i}$ correspond to dummy variables that indicate if origin lies within straight distance grid cell r from a subway, BRT or LRT station, respectively. Analogously, $\sum_r \Lambda^r_{BRT}$, $\sum_r \Lambda^r_{BRT}$ and $\sum_r \Lambda^r_{BRT}$ indicate if destination lies within straight distance grid cell

Table 1.1: Machine Learning results

Method	Prediction Performance (R2)		Relative improvement over OLS, by quantile of commuting time				
	Training sample	Hold-out sample	1st	2nd	3rd	4th	5th
OLS	29.2%	21.1% [18.8%,23.3%]	-	-	-	-	-
Regression Tree	27.2%	21.8% [19.7%,23.9%]	-2.3%	-20.1%	19.5%	31.0%	-7.9%
Lasso	28.1%	22.2% [20.0%,24.4%]	-0.6%	16.7%	10.5%	4.8%	-0.8%
Random Forest	68.1%	32.8% [30.4%,35.3%]	10.9%	18.8%	27.0%	28.9%	13.6%
Ensemble	59.4%	30.8% [28.4%,33.3%]	7.7%	14.8%	23.6%	23.9%	14.8%
Boosted Tree	93.3%	24.7% [21.3%,28.0%]	14.9%	-56.6%	-33.0%	-6.8%	6.6%

Notes: The dependent variable is the difference between travel time declared in the 2011 origin-destination survey and travel time collected in 2018 Google Maps API for the same route. Covariates include: a dummy variable that indicated if the trip is made by public transportation; the straight line distance betwen origin and destination; a set of dummy variables that indicate the distance from origin to the closest BRT, LRT and subway; a set of dummy variables that indicate the distance from destination to the closest BRT, LRT and subway station. I consider the 15 distance grid cells: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16 18, and 20 kilometers. All algorithms are fitted on the same, randomly drawn training sample of 1,600 units and evaluated on the remaining 6,524 held-out units. The numbers in brakets in the hold-out sample column are 95 percent bootstrap confidence intervals for hold-out prediction performancce, and represent measurement variation for a fixed prediction function.

straight r from a subway, BRT or LRT stations. I define 15 distance grid cells. Up to 10 kilometers, I consider distance grid cells of 1 kilometer intervals. Between 10 and 20 kilometers, I consider 2 kilometers intervals.

This grid cells specification allows me to estimate counterfactual travel times in the absence of transport investment by setting dummy variables equal to zero. For example, to simulate the no BRT scenario I set the following variable equal to zero: $\sum_r \Gamma^r_{BRT,i} = \sum_r \Lambda^r_{BRT,i} = 0$. Nevertheless, in order to estimate this counterfactual travel time, it is necessary to know the function $f(.)$. Since the objective of this exercise is to predict counterfactual travel times, and not to recover the causal parameters, I use machine learning techniques. In particular, I follow Mullainathan and Spiess (2017) and use five different methods: regression tree, lasso, random forest, ensemble and boosted tree. In order to compare the methods' performance, I also report OLS estimates. Results in table 1.1 show that the random forest method obtains the best fit out of sample.

Now I present the estimated counterfactuals using random forest method. I predict travel times in five scenarios: no investments; no BRT; no LRT; no subway extension; and no BRT and subway stations. Figure 1.17 plots the travel time distribution under all 5 scenarios and the true distribution in 2018.

Table 1.2: 2018 Google Maps travel time and estimated counterfactuals

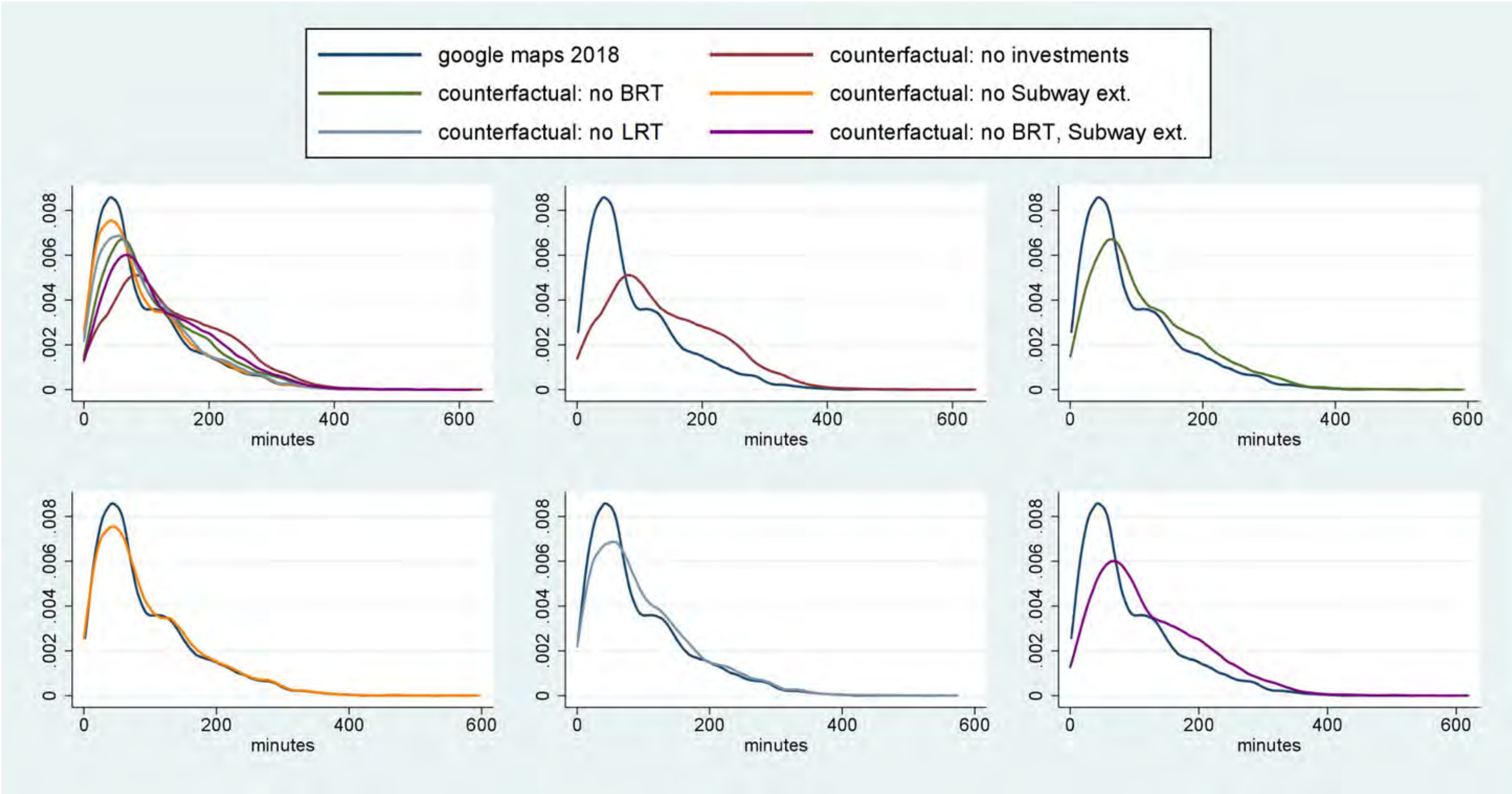
	mean	sd	Differences in means (contrafactual - 2018)
2018	93.77	75.03	
No investments	139.55	87.78	45.77
No BRT	113.93	80.31	20.16
No subway extension	96.53	76.17	2.76
No LRT	101.90	76.25	8.12
No BRT and subway extension	123.28	82.64	29.51

Notes: Counterfactual travel time were estimated using random forest regression. The dependent variable is the difference between travel time declared in the 2011 origin-destination survey and travel time collected in 2018 Google Maps API for the same route. Covariates include: a dummy variable that indicated if the trip is made by public transportation; the straight line distance betwen origin and destination; a set of dummy variables that indicate the distance from origin to the closest BRT, LRT and subway; a set of dummy variables that indicate the distance from destination to the closest BRT, LRT and subway station. I consider the 15 distance grid cells: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16 18, and 20 kilometers. All algorithms are fitted on the same, randomly drawn training sample of 1,600 units and evaluated on the remaining 6,524 held-out units. First, I estimate the difference in travel times in the absense of the stations. For BRT contrafactual, all dummy variables concerning the distance between origin and destination and BRT stations are equal to zero. Second, I add the difference to the 2018 travel times collected from Google Maps.

Results indicate that transport investments had a relevant impact on travel times. In the absence of all of these investments, Table 1.2 shows that the unweighted in-sample average commuting times would have increased by 45 minutes. Comparing each transportation mode, BRT has the highest effect on reducing travel time (20 minutes), followed by LRT (8 minutes) and the subway extension (3 minutes).

It is important to highlight three facts. First, in all scenarios, the distribution shifts to the right, indicating that, in the absence of investments, travel times would be higher. Second, estimates show the relative magnitude of each transportation mode effect. As expected, larger effects occur under the scenario of no investments. Then effects follow the decreasing order: no subway and BRT, no BRT, no LRT, and no subway. More important, these results provide evidence of investment complementarities: the total effect is higher than the sum of partial effects.

Figure 1.17: 2018 and counterfactual travel times



Notes: Counterfactual travel time were estimated using random forest regression. The dependent variable is the difference between travel time declared in the 2011 origin-destination survey and travel time collected in 2018 Google Maps API for the same route. Covariates include: a dummy variable that indicated if the trip is made by public transportation; the straight line distance betwen origin and destination; a set of dummy variables that indicate the distance from origin to the closest BRT, LRT and subway; a set of dummy variables that indicate the distance from destination to the closest BRT, LRT and subway station. I consider the 15 distance grid cells: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16 18, and 20 kilometers. All algorithms are fitted on the same, randomly drawn training sample of 1,600 units and evaluated on the remaining 6,524 held-out units. First, I estimate the difference in travel times in the absense of the stations. For BRT contrafactual, all dummy variables concerning the distance between origin and destination and BRT stations are equal to zero. Second, I add the difference to the 2018 travel times collected from Google Maps.

Chapter 2
Estimating the effects of transport stations on economic activity: a reduced-form approach

2.1
Introduction

The objective of this chapter is to establish a causal link between the inauguration of a new transport station and the organization of economic activity on its vicinity. Evidence from Chapter 1 suggests that the new infrastructure caused a reduction in travel times. Lower commuting costs can impact the stations' surroundings by two mechanisms. First, residents of the treated area may increase their accessibility to jobs ((Tyndall, 2017)). In the case of Rio, BRT and subway stations connected new areas to the central business district, which can raise the probability of employment and its quality. Thus newly connected areas attract more residents and its existent residents may experience a positive income shock. In response, economic activity grows and/or changes to attend this new demand. On the other hand, lower commuting costs can attract firms. For example, firms may reallocate from more expensive consolidated areas to newly connected ones. The two mechanisms are not excludent, but their relative strength determines land use patterns.

Although these mechanisms explain the causal link between new infrastructure and economic activity, causality goes both ways. New stations are often built to attend existing demand for transportation services in these areas. This is specially true for developing countries, where rapid urbanization leads to cities sprawl without the adequate urban planning. Consequently, estimating causal parameters is a empirical challenge. Additionally, even if treatment was randomly allocated, it is hard to define a pure control group due to the possibility of reallocation of economic activity. If firms or residents reallocate from control groups to treatment groups, estimated parameters will be biased.

In order to tackle this issues, I construct a highly detailed data set: the metropolitan area is divided in 100 meters square grids and, for each grid, I record information on number of firms, number of jobs, and average wage from 2006 to 2016. With this detailed intra-city data, I explore the timing of announcement and inauguration of new BRT, subway and LRT to estimate a fixed effects specification. In particular, I estimate effects across eight distance rings from stations: up to 250m, 250 to 500m, 500 to 750m, 750 to 1 kilometer, 1 to 1.25 km, 1.25 to 1.5 km, 1.5 to 1.75 km and 1.75 to 2 kilometers. Consequently, control grids are 2 kilometers apart from stations. I show that pre-trends assumption holds, indicating that I recover causal parameters.

Additionally, I shed light on the specific characteristics of the intervention by estimating effects across several dimensions of heterogeneity. First, I distinguish effects between different types of transportation modes. Since BRT is a much cheaper technology than subway, it is paramount to evaluate its relative benefit. Since I estimate the impacts of the three types of transportation technology in the same framework, results are directly comparable. I also estimate effects across workers' educational level, firms' sector of activity and size.

In relation to heterogeneous effects by transportation mode, I have four main results. First, announcement and inauguration of subway and BRT stations have positive and significant impacts on number of firms, jobs and average wage; while LRT has null or negative impacts. Since LRT was constructed in the central business district and I only observe the first year since inauguration, results may reflect short-term displacement effects. Second, announcement impacts are higher than inauguration impacts, evincing that firms anticipate the effects of new transport infrastructure. Third, subway stations inauguration effects are larger and have higher geographical reach. It is important to highlight that these heterogeneous effects between BRT and subway may be due to ex-ante difference between the treatment groups and not because of the transport technologies by itself. Forth, new subway and BRT stations had impact in grids with zero employment in baseline, which indicates that investments led to city sprawl, specially for BRT stations.

Concerning heterogeneous impacts according to workers' characteristics, impacts on number of jobs are stronger for workers up to high school level in comparison with workers with a college degree. Besides, all workers experience the same proportional change in wages, which may be a result of selection or agglomeration externalities. Finally, the bulk of effects come from small firms (2 to 10 employess) and from the service and commerce sectors.

This chapter relates to two streams of literature that attempt to estimate intra-city impacts of transportation investments. Concerning BRT stations, Perdomo (2011) finds a positive impact on property prices in areas in the vicinity of TransMilenium, Bogota's BRT. Bocarejo et al. (2016) show that areas served by TransMilenio have higher population growth than areas without access to the system. On the other hand, Rodríguez and Mojica (2009) find increases in property asking price in areas already served by TransMilenio that benefited from an extension, but estimate null impacts for newly connected areas. Regarding the Metropolitano BRT in Lima, Peru, Scholl et al. (2018a) find evidence of increases in residential rent-prices in feeder lines connected to the system, but not on the line itself. Also related to Metropolitano feeder lines, Scholl et al. (2018b) estimate positive effects on employment outcomes for individuals living in the most vulnerable areas in baseline. Concerning LRT and subway, there is a large literature, but only a few papers deal with causality issues. Gibbons and Machin (2005), Billings (2011) and Dorna G (2017) use difference-in-differences approaches and find positive impacts in housing prices in areas close to the system. Recently, Gonzalez-Navarro and Turner (2018) investigated the relationship between the extent of cities' subway network, its population and its spatial configuration in the 632 largest cities in the world. Although they find that subways have an economically insignificant effect on population growth, subways cause cities to decentralize. I contribute to both literatures by estimating impacts for both transportation technologies in the same framework, which allows for direct comparison. Additionally, I uncover heterogeneous effects in several dimensions.

Next section describe the data and present summary statistics. Section 3 presents the empirical strategy. Section 4 discusses identification and presents results.

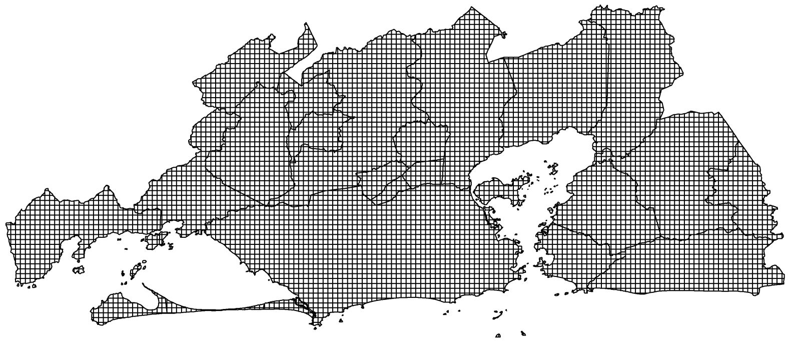
2.2 Data

2.2.1 Data sources and definitions

To investigate the effects of the expansion of transport infrastructure on economic activity, I combine three types of information about the Rio de Janeiro Metropolitan Area: firm location; timing and location of new transport stations; and RJMA administrative division. First, I divide the Rio de Janeiro Metropolitan Area in 100 meters square grids, which sum to about 580 thousand grids. Due to the challenge of visualizing such a fine grid, Figure 2.1 plots a 1-kilometer square grid.

Then, I geocoded firms' addresses from RAIS 2006 to 2016 and matched them to grid shapefile. As a result, the panel data set has information on

Figure 2.1: Rio de Janeiro Metropolitan Area: 1 km square grid with municipalities division



Notes: Author's elaboration

the number of firms, the number of jobs, and the average wage per year and grid. Additionally, since RAIS discloses information on firms' and workers' characteristics, I calculated the number of jobs and average wage per worker‘s educational level as well as the number of firms per firm size and sector of activity. I consider three education levels: incomplete high school, complete high school, and complete college. Firms are divided into five size categories: 0, 1, 2 to 10, 11 to 20, and more than 20 employees. The firms are also divided into six sectors: agriculture, public administration, commerce, industry, services, and construction. Agriculture and public administration are treated as separate categories and are excluded from the aggregate variables. As mentioned in Chapter 1, the addresses of public administration firms are known to be misreported. For example, all public education teachers are allocated formally to the municipality headquarters and not to their school's address. So, to avoid measurement error, public administration jobs are not included in the total jobs in this analysis. Concerning agriculture, I expect negative or null effects since the city is notably urban. Most of agricultural firms are located out of the Metropolitan Area. I use these sectors to perform robustness exercises.

Grids are classified as a treated unit if they lie within a 2 kilometers distance from a functioning station. And, inside this 2-kilometer radius, I differentiate between eight treatment intensities[1], according to the straight line distance between the grid and the station: up to 250m, 250 to 500m, 500m to 750m, 750m to1km, 1 to 1.25km, 1.25 to 1.5km, 1.5 to 1.75km and 1.75 to

[1]The same grid can be classified as treated in two different transportation modes, but not in two different distance rings. In case of a grid that lies within 400m from BRT station A and 1100m from BRT station B, it will be classified as treated by the closest distance, between 250 and 500m.

Figure 2.2: Rio de Janeiro City: Administrative Division

Notes: Rio de Janeiro City administrative division is from Data Rio, Instituto Pereira Passos, Rio de Janeiro Municipality.

2 km. Additionally, I define four types of treatment. First, grids are treated independently of the transportation technology. Second, I define treatments specifically for each transportation mode: BRT, LRT and subway. In total, there are 32 treatment variables (4 types x 8 intensities). Consequently, control grids are more than 2 km apart from any subway, LRT, and BRT station.

The RJMA administrative division is displayed in Figures 1.3 and 2.2. The metropolitan area is composed of 20 municipalities. Representing more than half of the total population and jobs, the city of Rio de Janeiro is divided into seven zones. Each zone is divided into districts (RA), which sums to 33 units. The RAs are the smallest unit of city planning. To account for differences among these regions, each grid is attributed to a municipality. And, for the grids that belong to the city of Rio de Janeiro, I record information on each zone and district.

2.2.2 Summary Statistics

Figure 2.3 shows the location of the transport infrastructure in 2006 and 2016. Since all of the new stations lie inside the city of Rio de Janeiro, treated units are mostly located in the city. As a consequence, all other municipalities in the Metropolitan Area belong to the control group. To guarantee that control and treatment are as similar as possible and diminish the possibility of selection bias, I chose to estimate results using only grids inside the Rio de Janeiro city[2]. Therefore, the following statistics concern only Rio de Janeiro city (~123 thousand grids).

[2]Results are similar using the full data set

Figure 2.3: Transport Infrastructure

Notes: Data is from Data Rio, Instituto Pereira Passos, Rio de Janeiro Municipality.

The number of grids by treatment status and year are reported in Table 2.1. As detailed in Chapter 1, LRT and BRT are new transportation technologies and there were no stations at baseline. In this period, 135 BRT and 15 LRT stations opened. The subway was extended with nine new stations, amounting to a total of 41 stations. Two comments are in order concerning the timing of their inauguration. First, all LRT stations opened in 2016, which means that I only observe one year of treatment. Second, the construction of subway stations that opened between 2007 and 2014 started in the 1980s. Consequently, only the stations that opened in 2016 can be linked directly to the Olympic‘s plans.

Table 2.2 shows the evolution of the main outcomes between 2006 and 2016. Two facts stand out: a strong growth in 2010/2011; and a recession between 2014 and 2016. Although this evolution is largely explained by the country’s performance, the 2010 growth may also be a consequence of Rio‘s election as host city of the 2016 Olympics Games. Even more important, since transportation plans were part of Rio application as host city, the election probably had differential impacts in the control and treatments groups. In this sense, the Rio’s election made transportation plans credible. Figure 2.4 plots the mean and standard deviation of the logarithm of total jobs, firms and average wage by treatment status. I present information on three treatment groups: grids within 250m, between 250 and 500m and between 500 and 750m from a new station, regardless of the transportation mode. Figures 2.5, 2.6 and 2.7 replicate Figure 2.4 for specific treatment groups: BRT, LRT and subway.

It is important to highlight three facts. First, at baseline, grids closer to stations have on average more firms, jobs and higher wages than grids 2 km away. This is true for all treatment groups and outcomes (Table 2.3). Second, between 2006 and 2009, treatment and control groups seem to have

Table 2.1: Number of treated grids

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Panel A: Subway											
up to 250m	957	987	987	1043	1043	1043	1043	1043	1075	1075	1227
250 to 500m	1817	1885	1885	1954	1954	1954	1954	1954	2020	2020	2314
500 to 750m	2041	2121	2121	2151	2151	2151	2151	2151	2222	2222	2527
750 to 1000m	1776	1824	1824	1821	1821	1821	1821	1821	1879	1879	2225
1000 to 1250m	1634	1686	1686	1660	1660	1660	1660	1660	1728	1728	2107
1250 to 1500m	1532	1572	1572	1526	1526	1526	1526	1526	1577	1577	1982
1500 to 1750m	1409	1419	1419	1399	1399	1399	1399	1399	1465	1465	1877
1750 to 2000m	1397	1395	1395	1391	1391	1391	1391	1391	1439	1439	1808
Panel B: BRT											
up to 250m	0	0	0	0	0	0	1080	1391	2967	2967	3731
250 to 500m	0	0	0	0	0	0	1832	2187	4523	4523	5586
500 to 750m	0	0	0	0	0	0	2164	2354	4631	4631	5681
750 to 1000m	0	0	0	0	0	0	2389	2546	4644	4644	5529
1000 to 1250m	0	0	0	0	0	0	2513	2637	4632	4632	5381
1250 to 1500m	0	0	0	0	0	0	2480	2572	4458	4458	4984
1500 to 1750m	0	0	0	0	0	0	2264	2382	4054	4054	4580
1750 to 2000m	0	0	0	0	0	0	2135	2254	3895	3895	4379
Panel C: LRT											
up to 250m	0	0	0	0	0	0	0	0	0	0	301
250 to 500m	0	0	0	0	0	0	0	0	0	0	238
500 to 750m	0	0	0	0	0	0	0	0	0	0	240
750 to 1000m	0	0	0	0	0	0	0	0	0	0	230
1000 to 1250m	0	0	0	0	0	0	0	0	0	0	225
1250 to 1500m	0	0	0	0	0	0	0	0	0	0	230
1500 to 1750m	0	0	0	0	0	0	0	0	0	0	243
1750 to 2000m	0	0	0	0	0	0	0	0	0	0	265

Notes: I define treated grids by the following procedure. First, I define buffers from stations' location with distances: 250 m, 500 m, 750 m, 1 km, 1.25 km, 1.5 km, 1.75 km and 2 km. The grid is considered treated if any part of the grid lies inside these buffers. Then grids are classified by the smallest distance buffer.

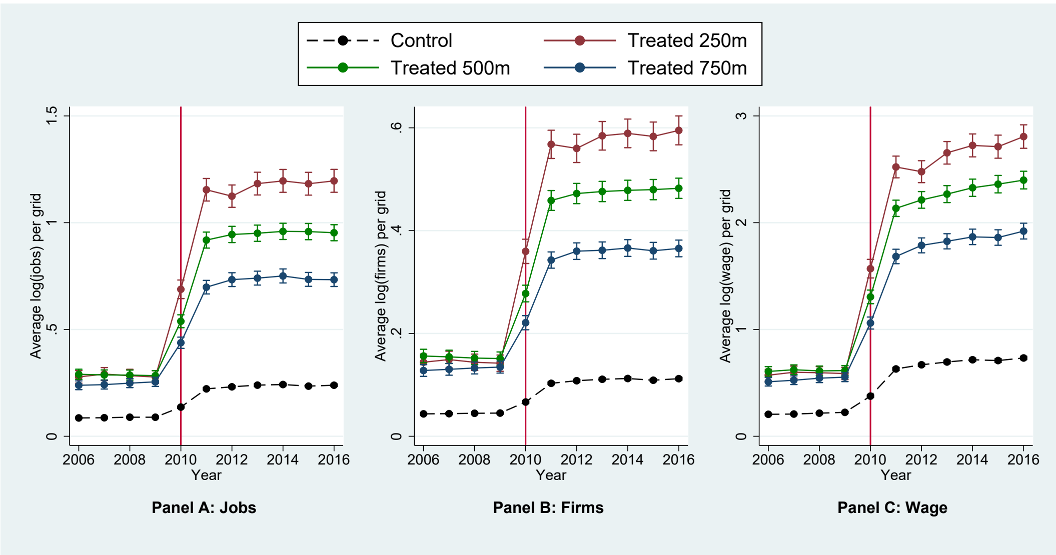
similar trends. Third, in 2010 and 2011, the whole city experienced growth, but the trend is much stronger for treatment groups. And the effect is larger for treatment groups closer to the new stations. This pattern is compatible with the timing of Rio's announcement as 2016 Olympics city host and point to a correlation between new transportation infrastructure and economic activity growth. In this direction, Figure 2.8 shows the spatial distribution of employment in baseline (2006) and endline (2016). As expected, it is notable the overlap between the location of jobs and the transport infrastructure.

Table 2.2: Rio de Janeiro Metropolitan Area: total number of formal jobs and firms from 2006 to 2016

	Firms	Jobs
2006	103,480	1,362,710
2007	104,862	1,404,753
2008	105,640	1,445,293
2009	107,938	1,564,298
2010	110,078	1,610,570
2011	119,087	1,929,472
2012	124,168	2,025,448
2013	126,144	2,062,004
2014	127,256	2,143,115
2015	125,109	2,020,842
2016	124,202	1,886,164

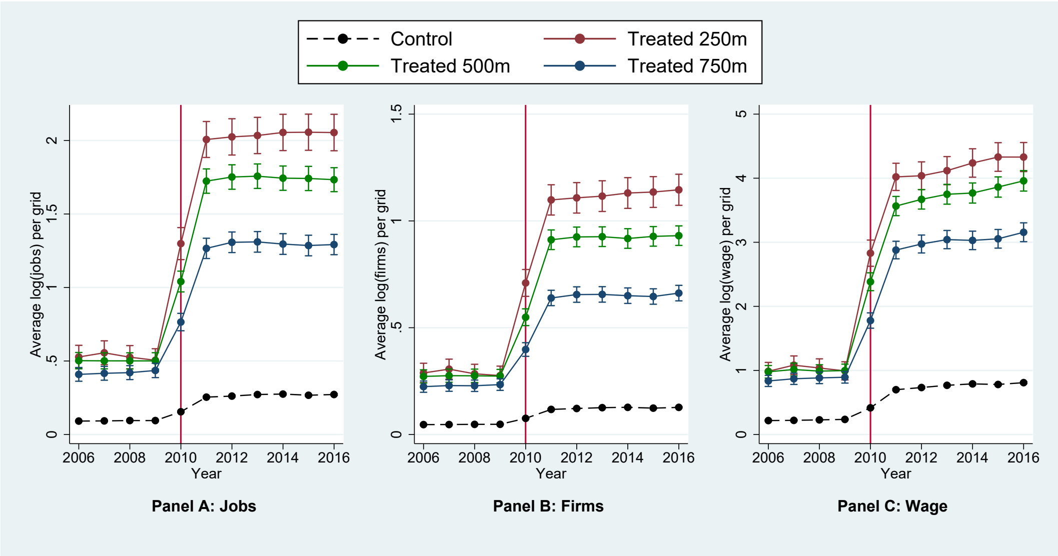
Notes: Data is from RAIS 2006 to 2016.

Figure 2.4: Treatment and Control groups



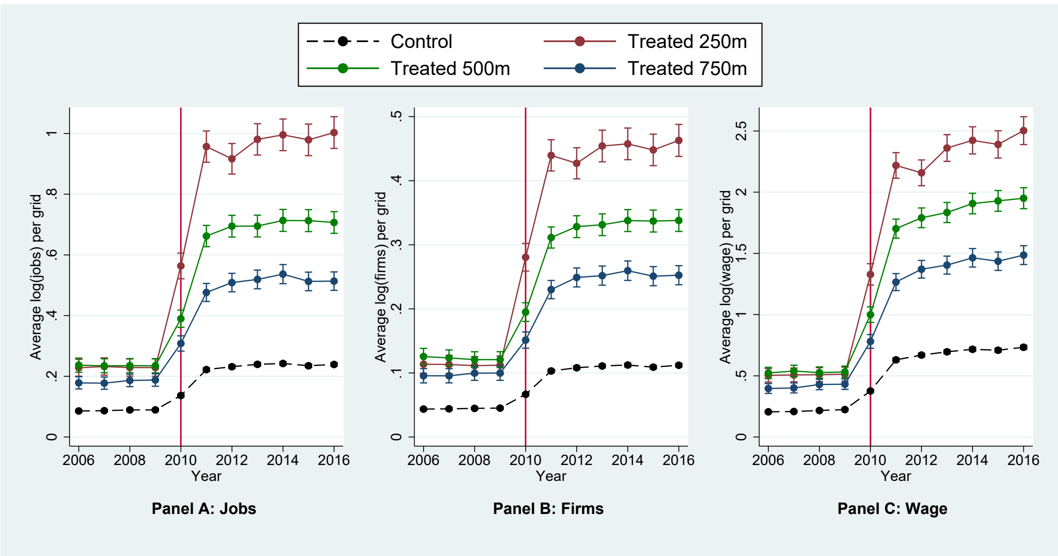
Notes: Data is from RAIS, 2006 to 2016. In the y axis, I plot average log(jobs) per group. Treatment groups are composed by grid at a determined distance ring from any BRT, LRT or new subway station. Control group are composed by grids at least 2 kilometers apart from any BRT, LRT and subway station.

Figure 2.5: Treatment and Control groups: Subway



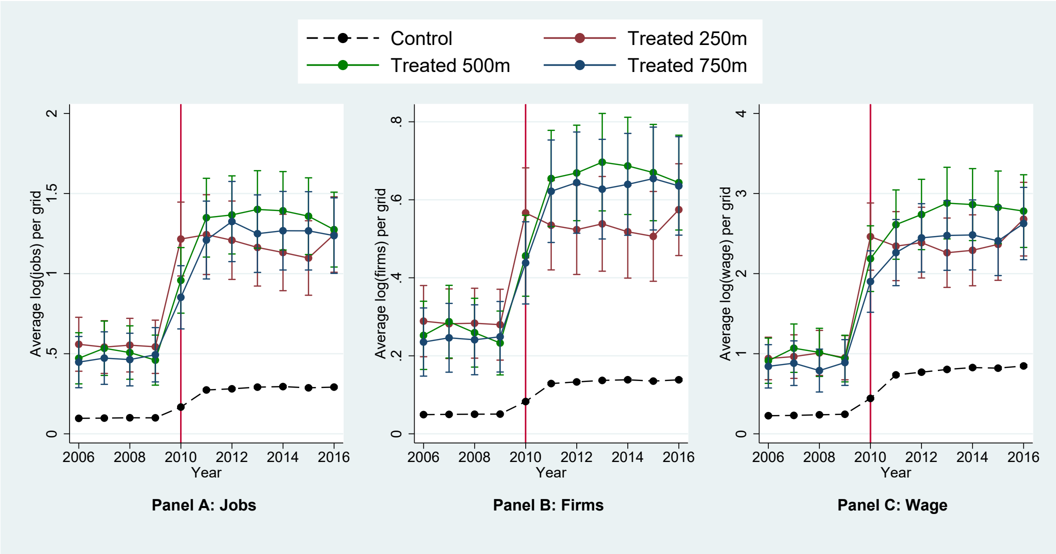
Notes: Data is from RAIS, 2006 to 2016. In the y axis, I plot average log(jobs) per group. Treatment groups are composed by grid at a determined distance ring from any subway station. Control group are composed by grids at least 2 kilometers apart from any BRT, LRT and subway station.

Figure 2.6: Treatment and Control groups: BRT



Notes: Data is from RAIS, 2006 to 2016. In the y axis, I plot average log(jobs) per group. Treatment groups are composed by grid at a determined distance ring from any BRT station. Control group are composed by grids at least 2 kilometers apart from any BRT, LRT and subway station.

Figure 2.7: Treatment and Control groups: LRT



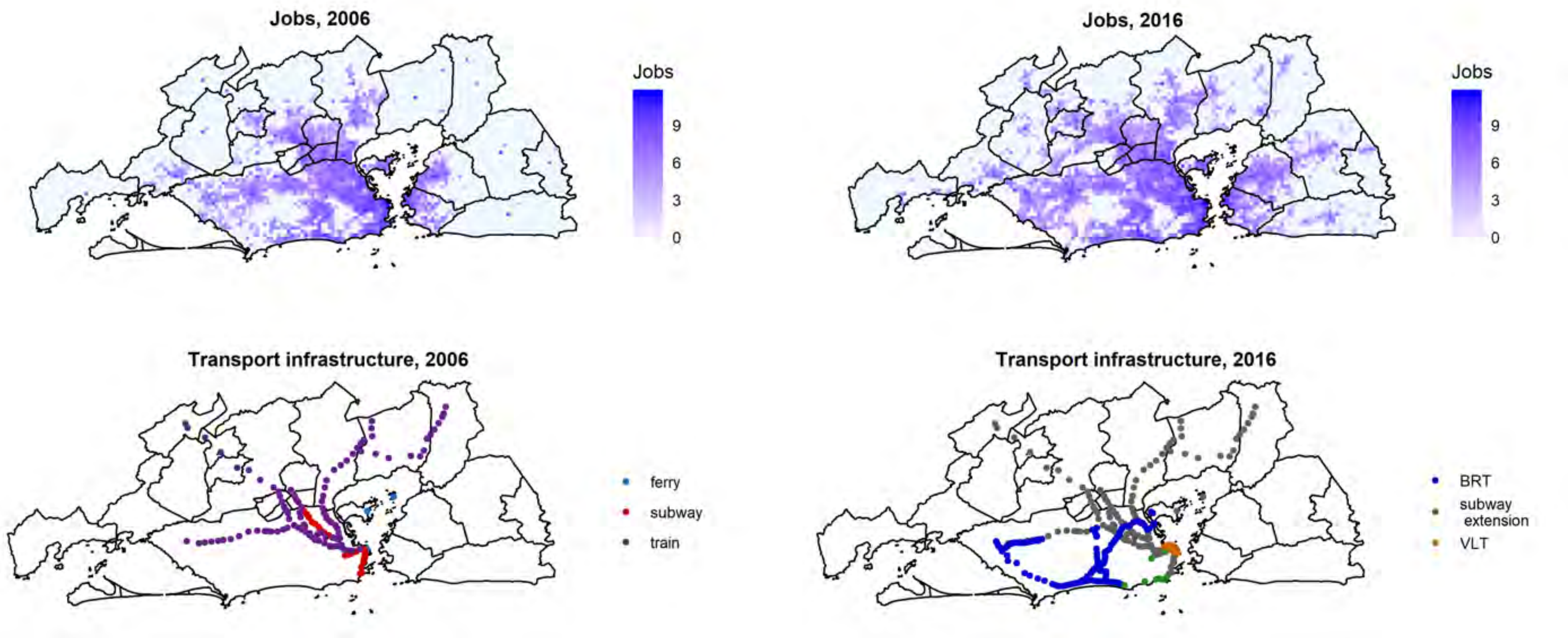
Notes: Data is from RAIS, 2006 to 2016. In the y axis, I plot average log(jobs) per group. Treatment groups are composed by grid at a determined distance ring from any BRT, LRT or new subway station. Control group are composed by grids at least 2 kilometers apart from any BRT, LRT and subway station.

Table 2.3: Mean outcomes in 2006

	Control	Treatment (meters from station)							
		up to 250	250 to 500	500 to 750	750 to 1000	1000 to 1250	1250 to 1500	1500 to 1750	1750 to 2000
Panel A: Jobs									
Total	3,1	83,5	41,4	23,6	14,3	13,1	9,7	9,3	10,6
No high school	1,7	31,0	18,5	10,4	6,9	6,8	4,0	4,8	5,7
High school	1,1	34,9	16,3	8,8	5,5	4,4	3,3	3,3	3,6
College	0,3	17,6	6,6	4,3	1,9	1,9	2,4	1,2	1,2
Panel B: Average Wages									
Total	21,2	100,4	87,1	69,2	56,4	52,5	48,5	47,1	44,0
No high school	17,8	82,7	74,2	59,5	47,3	43,7	40,6	39,3	38,2
High school	17,8	92,6	76,8	62,3	49,6	45,4	42,0	41,8	38,8
College	21,1	161,8	128,8	94,6	73,4	66,3	59,8	57,4	58,3
Panel C: Firms									
Total	0,3	6,0	3,2	1,7	1,0	1,0	0,7	0,6	0,8
Industry	0,0	0,2	0,2	0,1	0,1	0,1	0,1	0,1	0,1
Construction	0,0	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0
Commerce	0,1	1,7	0,9	0,5	0,4	0,3	0,2	0,2	0,3
Services	0,1	3,9	2,1	1,0	0,5	0,5	0,4	0,3	0,4
1 employee	0,1	1,4	0,8	0,4	0,2	0,2	0,2	0,1	0,2
2 to 10 employees	0,1	3,3	1,8	0,9	0,5	0,5	0,4	0,3	0,4
11 to 20 employees	0,0	0,6	0,3	0,2	0,1	0,1	0,1	0,1	0,1
more than 20 employees	0,0	0,6	0,3	0,2	0,1	0,1	0,1	0,1	0,1

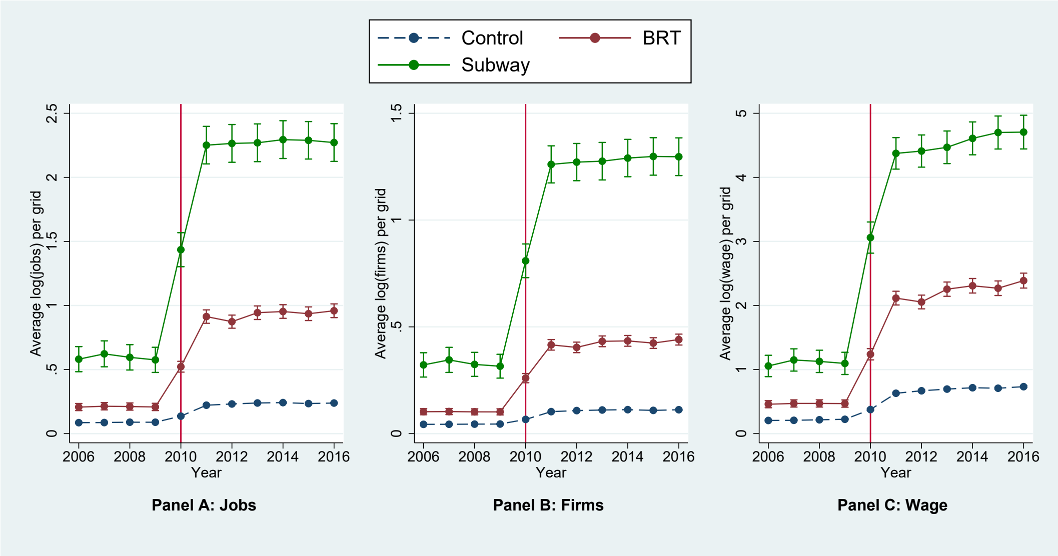
Notes: Data is from 2006 RAIS.

Figure 2.8: Distribution of Jobs and Transport Infrastructure in 2006 and 2016



Notes: Employment data is from RAIS, 2006 and 2016. Transport stations shapefile is from Data Rio, Instituto Pereira Passos, Rio de Janeiro Municipality.

Figure 2.9: Transport station within 250 m of grid



Notes: Data is from RAIS, 2006 to 2016. In the y axis, panel A, I plot average log(number of jobs) per group. In the y axis, panel B, I plot average log(number of firms) per group. In the y axis, panel C, I plot average log(average wage) per group. Control group are composed by grids at least 2 kilometers apart from any BRT, LRT and subway station.

Finally, I analyze the difference between treatment groups. Figure 2.9 presents the evolution of aggregate outcomes for three groups of grids: control, within 250m from a BRT station and within 250m from a subway station. In baseline, subway treated groups had more firms, jobs and higher wages than BRT treated grids and control. At endline (2016), this pattern remains, but the differences between groups are bigger. This difference in trends between BRT and subway treated grids can be interpreted in two, non-excludent, ways. The difference may be due to different impacts of the transportation technologies. In another words, subway stations may have a greater effect because of this specific technology. Or, regardless of the transportation mode, the heterogeneous impacts may reflect the different characteristics in baseline. In fact, considering treatment grids in a 250 meters radius from stations, 90% of BRT treated grids had no jobs in 2006. For subway, this percentage was 76% and for LRT 66%. Even more important, only 2.7% of 250m BRT treated grids are in last quartile of the distribution, against 13% for subway and 26% for LRT. So BRT, subway and LRT can be described, respectively, as interventions in low, medium and high-density environments.

2.3 Empirical Strategy

Evaluating the impacts of interventions in transport has two major challenges. First, as the previous section made clear, transport infrastructure is not randomly assigned. Stations are located in grids with more jobs and firms, which characterize selection bias. In the cross-section, control and treatment groups are not similar and it is not possible to identify the true causal effect. Second, even if treatment was randomly assigned, it is difficult to define an appropriate counterfactual for the absence of transport improvement due to contamination concerns. Grids outside the treated areas can be affected in the case of reorganization of economic activity. In particular, lower commuting costs can lead to the creation of new firms and cause existing firms to reallocate within the city. If firms reallocate from the control to treatment groups, the control grids will be contaminated by the treatment. In this last scenario, grids can be direct or indirectly affected, posing a threat to the definition of a pure control group. I address this identification issues exploring a long panel data set and divide the metropolitan area in a fine grid.

Specifically, in order to recover causal effects, I estimate three specifications using a difference-in-difference methodology. The first specification is detailed below.

$$Y_{irzt} = \beta_0 + \sum_d \beta_1^d T_{irzt}^d + \theta_i + \rho_t + t_r + t_r^2 + \varepsilon_{irzt}, \tag{2-1}$$

where Y_{irzt} denotes the logarithm of the outcome in grid i, district r, zone z, year t. I define as main outcomes the number of firms, number of jobs and average wage. The treatment is represented by the group of variables $\sum_d T_{irzt}^d$, where T_{irzt}^d assumes value 1 if the grid i lies at distance d from a functioning station in year t. The parameter d defines eight categories according to the distance between grid and station: up to 250m, 250 to 500m, 500 to 750m, 750m to 1km, 1.25 to 1.5km, 1.5 to 1.75km and 1.75km to 2km. The objective of the treatment specification is twofold. First, it allows for estimation of the station's influence area. Second, the probability of contamination in the control group is reduced since control grids are 2km apart from functioning stations. According to the literature, the station's influence area ranges from 800 meters to 1.2 kilometers, depending on the transportation technology[3]. Thus I expect zero effects on grids more than 1.25 km apart from stations, which translates into a 750m buffer between treated and control grids.

In case this non-contamination hypotheses fails, coefficients will be biased. Effects are underestimated if positive effects are present at distances

[3]incluir Raio de Influência de Estações de LRT, Metrô e BRT

greater than 2 km. For example, a subway station can increase the demand for transportation in a bus stop 2.5 km away. Then firms close to this bus stop will face higher demand due to the increase in the circulation of people. On the opposite direction, effects are overestimated if the control group is subject to negative effects. This would be case of a firm located in a grid 3 km away from a BRT station that decides to move to a treatment grid in response to the station opening. In particular, the specification cannot differentiate between creation and reorganization of the economic activity.

I control for grid and year fixed effects. Besides, I allow each district to have different linear and quadratic trends. Since districts are the smallest unit of city planning, the inclusion of these variables allows for heterogeneous dynamics led by different policies across districts. The error term ε_{iast} is clustered at the district level.

Specification 2-1 has two identification hypotheses. First, the non-contamination of the control group. Second, treatment and control groups must have the same trend prior to station's inauguration. Nevertheless, Figure 2.4 shows that treatment and control groups have different trends due to Rio's announcement as 2016 Olympics host city. As mentioned in the previous section, transportation plans were public in 2010 and Rio de Janeiro election made them credible. In order to control for differential effects of the announcement by treatment status, I add announcement treatment effects in specification 2-2:

$$Y_{irzt} = \beta_0 + \sum_r \beta_1^d T_{irzt}^d + \sum_r \beta_2^d A_{irzt}^d + \theta_i + \rho_t + t_r + t_r^2 + \varepsilon_{irzt}, \tag{2-2}$$

where announcement is represented by the group of variables $\sum_d A_{irzt}^d$,and A_{irzt}^d assumes value 1 from 2010 to 2016 if the grid i lies at distance r from a station.

As suggested by Figure 2.9, subway station may have stronger effects than BRT stations. To account for heterogeneous effects, specification 2.3 introduces specific treatment variables for each transportation mode:

$$Y_{irzt} = \beta_0 + \sum_d \beta_{11}^d BRT_{irzt}^d + \sum_d \beta_{12}^d SUB_{irzt}^d + \sum_d \beta_{13}^d LRT_{irzt}^d \tag{2-3}$$

$$+ \sum \beta_{21}^d Abrt_{irzt}^d + \sum_d \beta_{22}^d Asub_{irzt}^d + \sum_d \beta_{23}^d Alrt_{irzt}^d + \theta_i + \rho_t + t_r + t_r^2 + \varepsilon_{irzt} \tag{2-4}$$

At the same time, differences may be due to baseline grid characteristics. In particular, if transportation infrastructure is complementary to other urban infrastructure, effects may be non-linear in the initial density of economic activity. For example, firms may decide not to locate in a grid with no Internet access. Since urban infrastructure and the employment density are correlated, the magnitude of the impact may be higher in denser areas. Another possible explanation is coordination failure. Due to agglomeration externalities, firms will not operate in a grid that has no other firms. In equilibrium, if no firms decide to enter, the station's impact is null. I investigate how these two mechanisms interact by estimating specification 2.3 using all grids and five subsamples. I divide grids according to baseline characteristics. First, I create a subsample with grids that had no firms in 2006. Then I split the remaining grids in quartiles. I repeat the same procedure using the number of jobs instead of number of firms. The exercise allows verifying if effects are homogeneous across the grid distribution for each transportation mode.

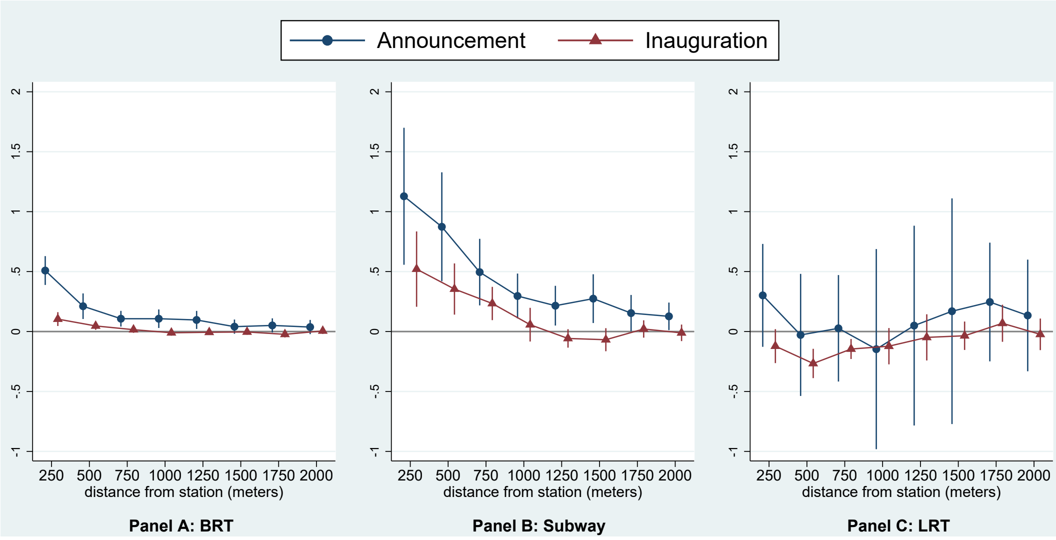
All specifications are estimated using as dependent variables the logarithm of the main outcomes (number of firms, number of jobs and average wage). Since I use the logarithm transformation, coefficients β_1 and β_2 are interpreted as the percentual change caused by the treatment. I also analyze if there are heterogeneous effects across worker's and firm's characteristics. I estimate specification 2.3 using as dependent variables the logarithm of the number of jobs and average wage by educational level: incomplete high school, complete high school and complete college. Concerning firms, I use as outcome the logarithm of the number of firms by sector of activity (construction, service, commerce and industry) and size (0,1, 2 to 10, 11 to 20, and more than 20 employees). The agriculture sector is used to perform robustness exercises since I expect zero or negative effects.

Finally, to investigate how effects accumulate over time, I estimate specification 2.3 with leads and lags from BRT station's inauguration. It is not possible to perform the same exercise for LRT cause I only observe one year of treatment. Concerning subway, there are two types of stations. Stations inaugurated between 2007 and 2014 were planned long ago and construction began in the 1980's, which implies disperse and staggered anticipation effects. For subway stations that opened in 2016, there is only one year of treatment.

2.4 Results

Now I discuss identification comparing estimates from specifications 2-1,2-2 and 2.3. Since results are qualitative the same for the three main outcomes, for simplicity I only present results with number of jobs as dependent variable. Table 2.4 displays results for specifications 2-1 and 2-2. In column

Figure 2.10: Estimated coefficients: BRT, LRT and Subway (specification 3)



Notes: The graphs plot estimated coefficients for specification 3 (Table 14). The dependent variable is the log(number of jobs). Each sub-figure refers to one transportation mode: BRT (a); subway (b); LRT (c). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per zones. Erros are clustered by zones. Dots represent point estimates and bar represents 95% confidence interval.

(1), I show results of specification 2-1 with grid and year fixed effects. In column (2), I add linear and quadratic trends for seven city's zones, while column (3) includes linear and quadratic trends for 33 districts. Significance and magnitude of coefficients increase with the inclusion of controls. Stations affect grids in a 750 meters radius and, as expected, impacts are stronger for grids closer to stations. Column (4) shows estimates from specification 2-2. Inauguration impacts disappear with the inclusion of announcement treatment variables. Announcement effects are much stronger and influence grids up to 2 kilometers from stations.

Results from specification 2.3 are shown in Figure 2.10. Estimates show evidence of important heterogeneities across transportation technologies. In relation to BRT and subway, treated grids are positively affected by announcement and inauguration. For both transportation modes, announcement effects are stronger and have a broader geographical reach. The magnitude of subway impacts is higher and the spatial decay is lower than BRT. In the opposite direction, LRT stations' announcement has zero effect and inauguration has a negative effect up to 750 meters. Since LRT stations are located in highly dense areas and I only observe grids in the year of inauguration, these impacts may represent displacement effects from construction. In the following sections, I carry out a detailed discussion on each transportation mode.

Table 2.4: Regression results: BRT, LRT and Subway (specification 1 and 2)

		(1)	(2)	(3)	(4)
Inauguration	up to 250 m	0.38***	0.39***	0.40***	0.05
		(0.04)	(0.05)	(0.05)	(0.05)
	250m to 500m	0.16***	0.18***	0.19***	-0.03
		(0.04)	(0.03)	(0.03)	(0.04)
	500m to 750m	0.05	0.07***	0.08***	-0.02
		(0.04)	(0.02)	(0.02)	(0.03)
	750m to 1000r	-0.01	0.02	0.04*	-0.03
		(0.04)	(0.03)	(0.02)	(0.02)
	1000m to 125(	-0.02	0.01	0.03	-0.03*
		(0.04)	(0.02)	(0.02)	(0.02)
	1250m to 150(	-0.05	-0.02	0.00	-0.05**
		(0.03)	(0.02)	(0.02)	(0.02)
	1500m to 175(	-0.04	-0.01	0.01	-0.03*
		(0.03)	(0.02)	(0.02)	(0.02)
	1750m to 200(	-0.03	-0.00	0.01	-0.02
		(0.03)	(0.02)	(0.02)	(0.02)
Announcement	up to 250 m				0.67***
					(0.09)
	250m to 500m				0.41***
					(0.09)
	500m to 750m				0.20***
					(0.05)
	750m to 1000r				0.13***
					(0.04)
	1000m to 125(				0.12***
					(0.03)
	1250m to 150(				0.11**
					(0.04)
	1500m to 175(				0.08**
					(0.03)
	1750m to 200(				0.06**
					(0.03)
Observations		1,347,786	1,347,786	1,347,786	1,347,786
R^2		0.04	0.06	0.07	0.08
Number of grids		122,526	122,526	122,526	122,526
FE Year		Yes	Yes	Yes	Yes
FE Grid		Yes	Yes	Yes	Yes

Notes: The table reports estimated coefficients and standard erros in parentheses. The dependent variable is log(number of jobs). Column (1) reports results from specification (1) with grid and year fixed effects. Column (2) reports results from specification (1) with grid and year fixed effects, linear and quadratic trends per zones. Column (3) reports results from specification (1) with grid and year fixed effects, linear and quadratic trends per districts. Column (4) reports results from specification (2) with grid and year fixed effects, linear and quadratic trends per district. In all columns, error are clustered by districts.

In light of relevant heterogeneous effects, specification 2.3 is the preferred one. As discussed in the previous section, estimates are unbiased depending on two identification hypotheses. Although I cannot test the control group non-contamination hypotheses, it is possible to check if treatment and control groups have the same trend pre-announcement. Figures 2.5, 2.6 and 2.7 show compelling evidence of parallel trends for treatment groups specific to each

Table 2.5: F test: Pre-announcement Trend

	BRT	Subway	LRT
up to 250 m	0,18	2,57	2,77
250m to 500m	1,05	0,96	3,15
500m to 750m	0,55	0,99	2,76
750m to 1000m	0,98	1,14	1,15
1000m to 1250m	1,58	1,34	1,95
1250m to 1500m	2,30	2,35	2,57
1500m to 1750m	1,39	2,18	8,81
1750m to 2000m	2,09	2,50	1,91

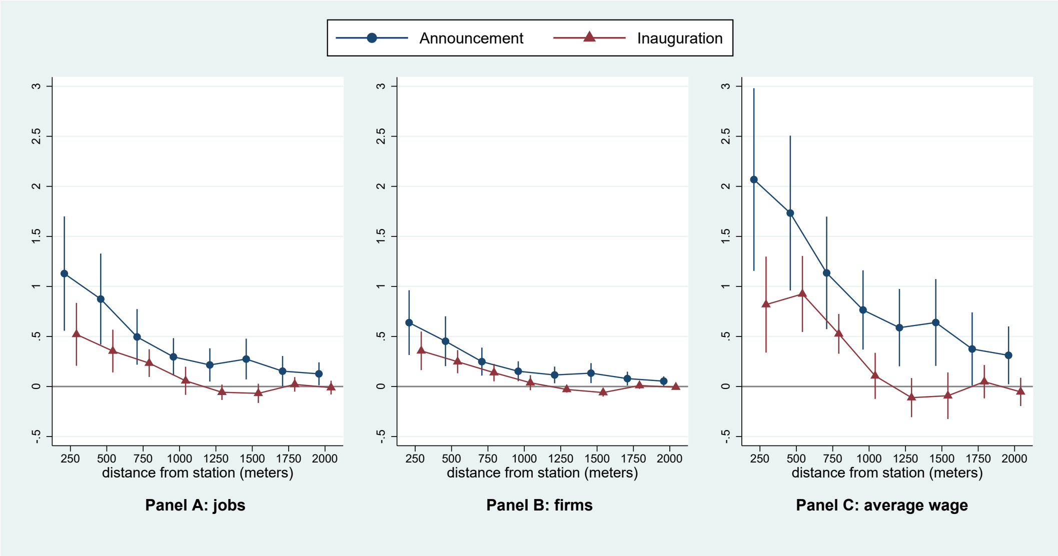
Notes: The table reports F statistics for hypotheses tests. Lines represents each treatment group. I test if the variables that identify the treatment group before annoucement are jointly significant. Columns represent transportation mode. For example, the cell in line "up to 250 meters" and column "BRT" is the result of the test of pre-annoucement trends for grids up to 250 meters from a BRT station.

transportation mode. To formally check the hypotheses, I include placebo treatment variables for years 2009, 2008 and 2007, and test if they are jointly different from zero. Table 2.5 presents F-statistics for each treatment group in specification 2.3. For BRT and subway, pre-trends assumption seems reliable. However, for LRT stations, several groups exhibit different trends pre-announcement, at 10%, 5% and 1% significance levels. Consequently, I will only analyze results for BRT and subway.

Figure 2.11 plots estimated coefficients for subway station's announcement and inauguration for the three main outcomes: number of jobs, firms and average wage. In 250 meters radius, station's announcement increases jobs by 113%, firms by 64% and average wage by more than 200%. The effect of station inauguration is 52%, 36% and 82%, respectively. Announcement effects span up to 2 kilometers from stations, while functioning station have positive impacts up to 750 meters. Although there are no estimates from the literature to compare with, results seem very large.

It is important to highlight that impacts on wages are higher than on number of jobs and firms. Thus, stations not only attracted more firms and jobs, but also the jobs created have higher wages. Two mechanisms can be in place: selection or agglomeration forces. Concerning selection, the composition of the labor force might have changed towards more qualified or productive workers. For example, due to lower commuting costs, firms are now able to hire workers with higher opportunities costs. Selection may also reflect a change in firm's characteristics, such as sector of activity. Besides, even if labor force educational composition remains the same, new workers might be more productive. In particular, higher productivity might be caused by agglomeration externalities. Subway stations attract more workers, which

Figure 2.11: Estimated Coefficients: Subway



Notes: The graphs plot estimated subway coefficients for specification 3 (Table 14). Each sub-figure refers to one dependent variable: log(number of jobs) (a); log(number of firms) (b); log(average wage) (c). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by zones. Dots represent point estimates and bar represents 95% confidence interval.

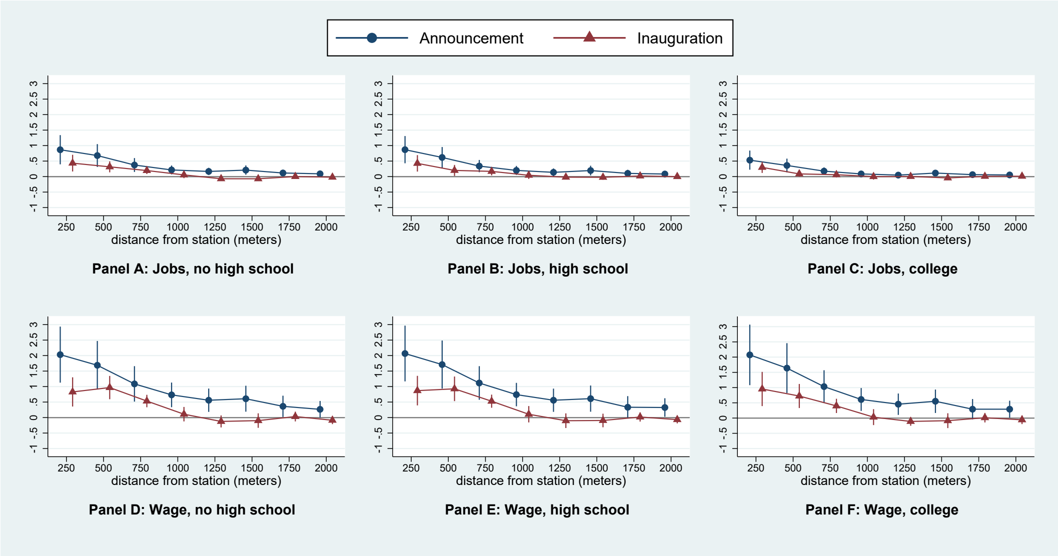
raises the employment density. In turn, higher density increases productivity through sharing, matching and learning[4]

I try to uncover mechanisms by exploring heterogeneous impacts according to firm's and worker's characteristics. Figure 2.12, 2.13 and 2.14 present results. Impacts on wages are not significantly different across worker's educational level: average wages double with station's inauguration in a 250 meters radius. On the other hand, the magnitude of effects on number of jobs for workers up to high school are twice the magnitude of effects for college workers. Since college workers' have on average higher wages, these effects point to lower average wages. Consequently, increase in average wage cannot be explained by a change in the educational composition of the labor force.

Regarding firm's sector of activity, the bulk of the impacts come from commerce and service sector. There are small positive effects for industry and construction sector. Analyzing heterogeneous effects per firm size, impacts are higher for firms up to ten employees. Nevertheless, depending on the exact size of firms, the total effect on firms with more than 11 employees can be more significant. Even more important, since bigger firms are more productive, this increase on the number of larger firms translated into higher productivity, which explains higher wages. More detailed research is needed to confirm this

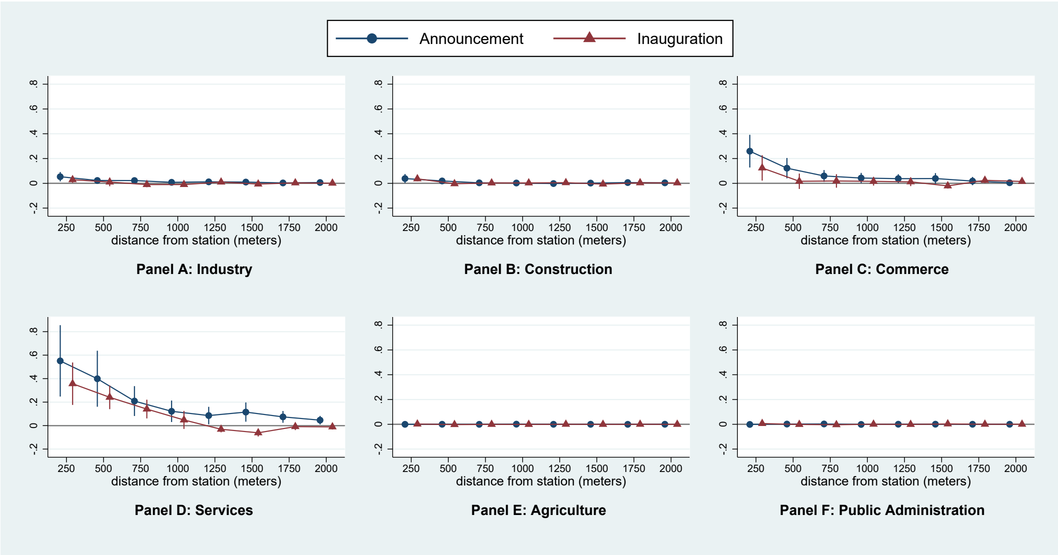
[4]For details in agglomeration economies, see Chapter 1.

Figure 2.12: Estimated Coefficients: Subway, per workers' educational level



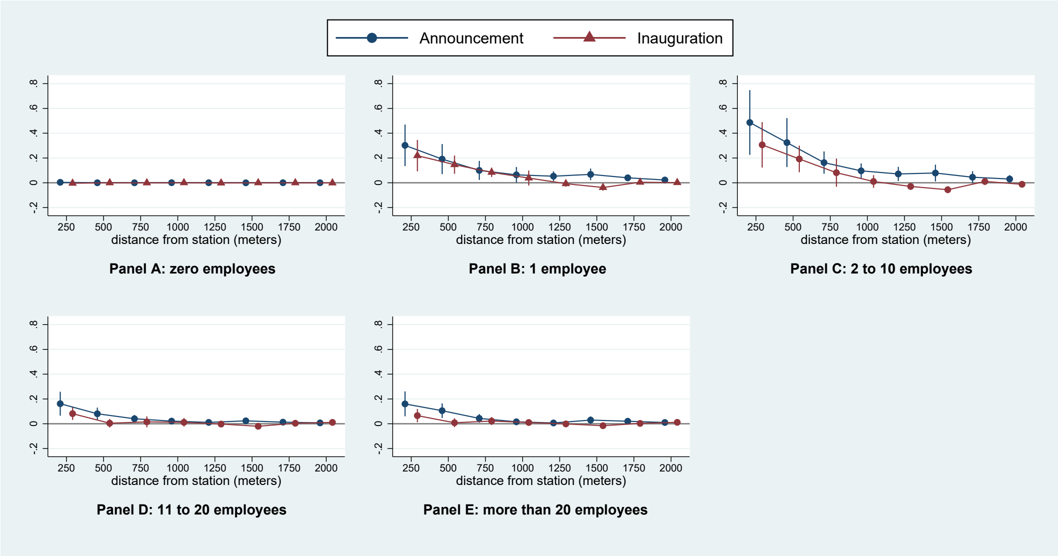
Notes: The graphs plot estimated subway coefficients for specification 3. Each sub-figure refers to one dependent variable: log(number of jobs of workers' with no high school) (a); log(number of jobs of workers' with high school) (b); log(number of jobs of workers' with college) (c); log(average wage of workers' with no high school) (d); log(average wage of workers' with high school) (e); log(average of workers' with college) (f). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval. Coefficients of columns (1), (2) and (3) are presented in Table 12. Coefficients of columns (4), (5) and (6) are presented in Table 13.

Figure 2.13: Estimated Coefficients: Subway, per firms' sector of activity



Notes: The graphs plot estimated subway coefficients for specification 3 (Table 9). Each sub-figure refers to one dependent variable: log(number of firms in the industry sector) (a); log(number of firms in the construction sector) (b); log(number of firms in the commerce sector) (c); log(number of firms in the service sector) (d); log(number of firms in the agriculture sector) (e); log(number of firms in the public administration sector) (f). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Figure 2.14: Estimated Coefficients: Subway, per firms's size



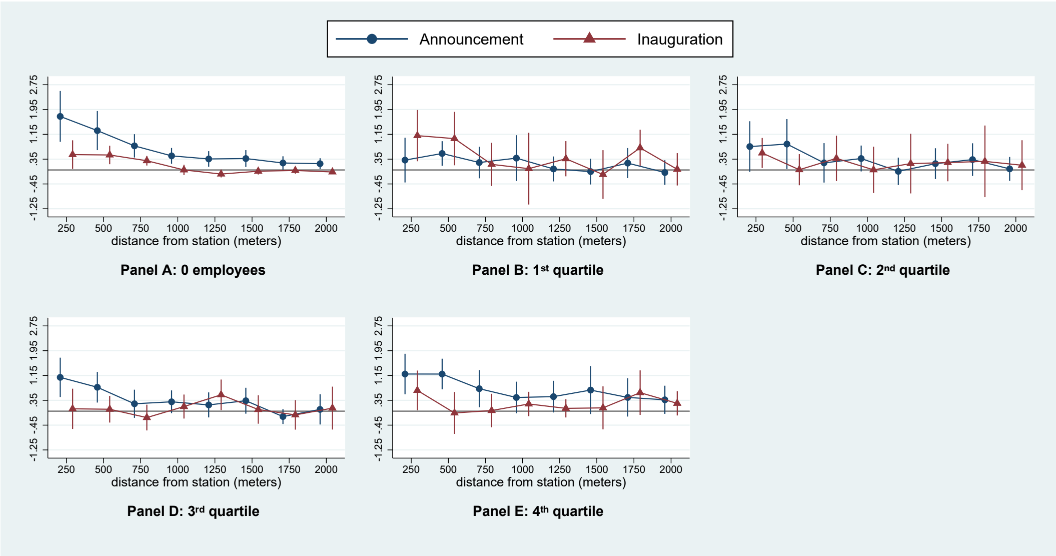
Notes: The graphs plot estimated subway coefficients for specification 3 (Table 14). Each sub-figure refers to one dependent variable: log(number of jobs) (a); log(number of firms) (b); log(average wage) (c). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per zones. Erros are clustered by zones. Dots represent point estimates and bar represents 95% confidence interval.

hypothesis.

Finally, I estimate specification 2.3 in subsamples according to baseline characteristics. Figure 2.15 evinces that announcement and inauguration impacts come from grids with zero and positive employment. Thus it is possible to divide effects in two types: increase in the density of economic activity, and expansion to areas where it was not present. The coefficients on grids with positive employment are higher; pointing that densification impact is stronger. On the other hand, impact on grids with zero employment in baseline has broader geographical reach.

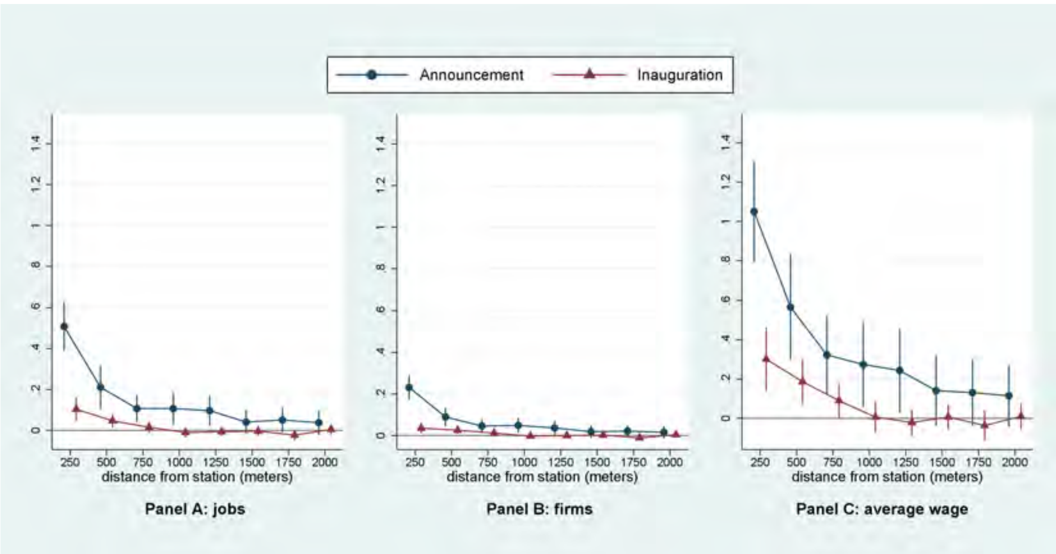
BRT impacts on main outcomes are displayed in Figure 2.16. Some patterns are similar to subway stations: announcement has larger impacts than inauguration; wages are more affected than firms and jobs. Besides, heterogeneous impacts are qualitative the same (Figures 2.17, 2.18 and 2.19). Nevertheless, there are important differences. First, BRT effects are much smaller. In 250 meters radius, BRT station's announcement increases jobs by 51%, firms by 23% and average wage by more than 100%. The effect of station inauguration is 10% on jobs, 4% on firms and 30% on wages. Second, the geographical scope of impacts is more limited. While subway stations's announcement have impact up to 2 kilometer, BRT spans to 1.25 kilometer. Concerning inauguration, subway affect grids up to 750 meters and BRT up to 500m.

Figure 2.15: Estimated Coefficients: Subway, per subsamples



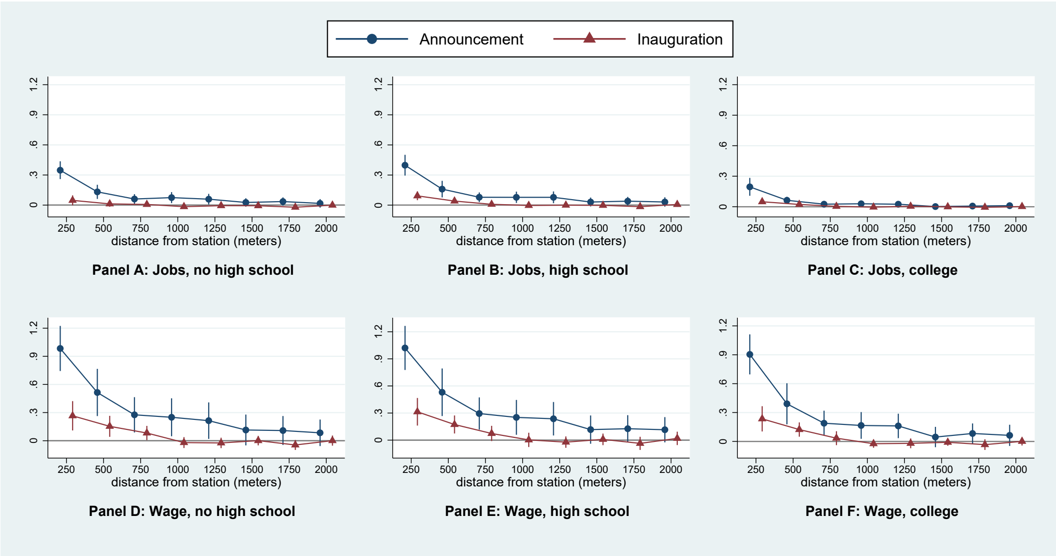
Notes: The graphs plot estimated subway coefficients for specification 3 (Table 11). The dependent variable is log(number of jobs). Each sub-figure refers to one subsample according with the number of jobs in the grid in 2006: 0 jobs (a); first quartile of the distribution of grids with positive employment (b); second quartile of the distribution of grids with positive employment (c); third quartile of the distribution of grids with positive employment (d); forth quartile of the distribution of grids with positive employment (e). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Figure 2.16: Estimated Coefficients: BRT



Notes: The graphs plot estimated BRT coefficients for specification 3 (Figure 14). Each sub-figure refers to one depedent variable: log(number of jobs) (a); log(number of firms) (b); log(average wage) (c). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

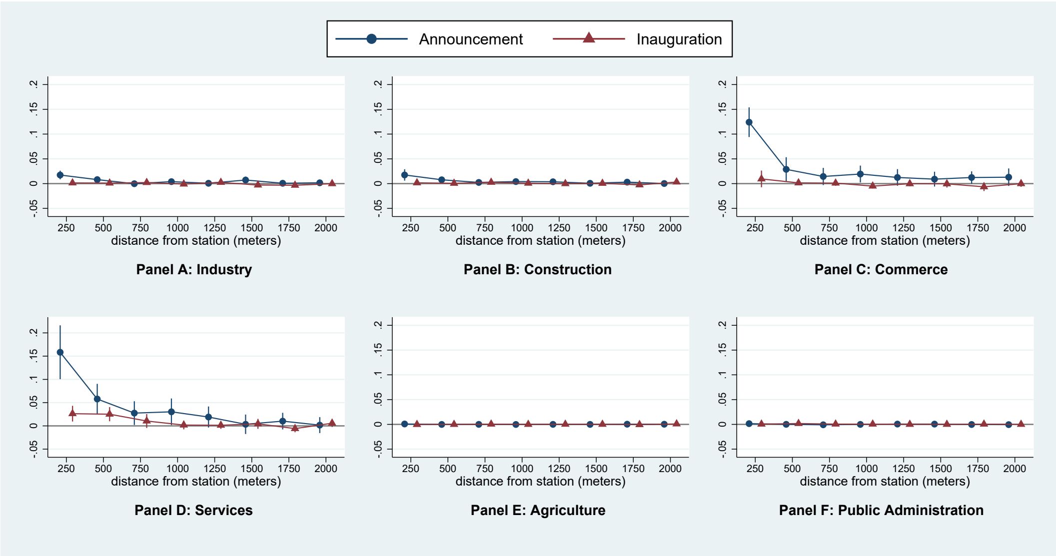
Figure 2.17: Estimated Coefficients: BRT, per workers' educational level



Notes: The graphs plot estimated BRT coefficients for specification 3 (Table 13). Each sub-figure refers to one depedent variable: log(number of jobs of workers' with no high school) (a); log(number of jobs of workers' with high school) (b); log(number of jobs of workers' with college) (c); log(average wage of workers' with no high school) (d); log(average wage of workers' with high school) (e); log(average of workers' with college) (f). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval. Coefficients of columns (1), (2) and (3) are presented in Table 12. Coefficients of columns (4), (5) and (6) are presented in Table 13.

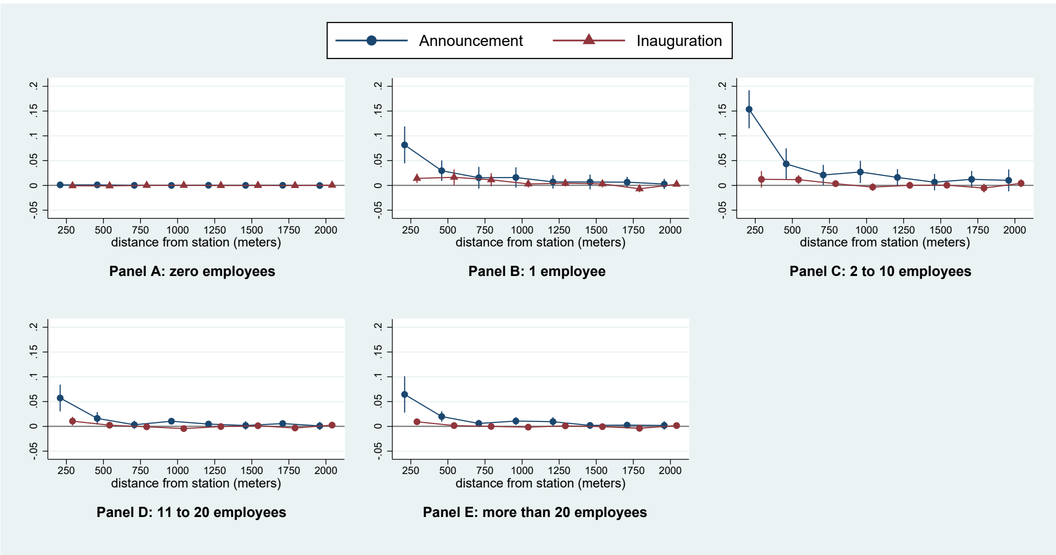
Even more important, Figure 2.20 shows stark differences between interventions: BRT station's opening only have positive effects on grids with zero employment in baseline. And, even though announcement effects have positive impact on grids in the first, second and third quartile, effects span to longer distances in grids with no employment or in the first quartile. So, BRT transport investment led to a city sprawl, developing new areas of economic activity.

Figure 2.18: Estimated Coefficients: BRT, per firms' sector of activity



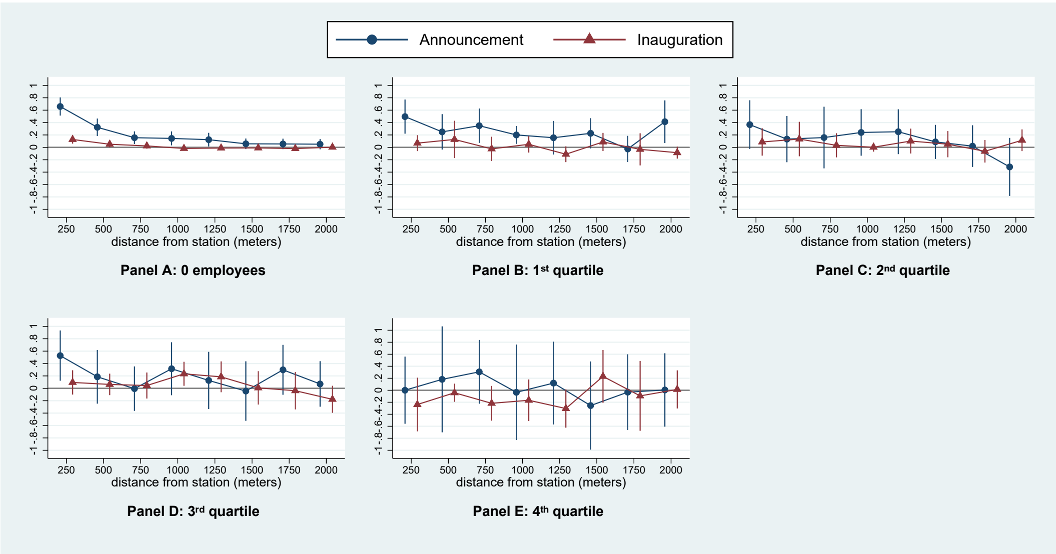
Notes: The graphs plot estimated BRT coefficients for specification 3 (Table 9). Each sub-figure refers to one depedent variable: log(number of firms in the industry sector) (a); log(number of firms in the construction sector) (b); log(number of firms in the commerce sector) (c); log(number of firms in the service sector) (d); log(number of firms in the agriculture sector) (e); log(number of firms in the public administration sector) (f). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Figure 2.19: Estimated Coefficients: BRT, per firms' size



Notes: The graphs plot estimated BRT coefficients for specification 3 (Table 10). Each sub-figure refers to one dependent variable: log(number of firms with 0 employees) (a); log(number of firms with 1 employee) (b); log(number of firms with 2 to 10 employees) (c); log(number of firms with 11 to 20 employees) (d); log(number of firms with more than 20 employees) (e). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Figure 2.20: Estimated Coefficients: BRT, per subsamples



Notes: The graphs plot estimated subway coefficients for specification 3 (Table 11). The dependent variable is log(number of jobs). Each sub-figure refers to one subsample according with the number of jobs in the grid in 2006: 0 jobs (a); first quartile of the distribution of grids with positive employment (b); second quartile of the distribution of grids with positive employment (c); third quartile of the distribution of grids with positive employment (d); forth quartile of the distribution of grids with positive employment (e). The sample includes 1,347,786 grids in the Rio de Janeiro City. Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Chapter 3
Quantifying aggregate and distributional effects: a structural approach

3.1
Introduction

In preparation for the 2014 World Cup and 2016 Olympic Games, Rio de Janeiro underwent a major expansion of public transport infrastructure. The city invested more than 4.5 billion dollars in its public transportation system. The investments included the extension of a subway line, the construction of a light rail system and two Bus Rapid Transit corridors, creating a transportation ring around the city (Figure 2.3).

Evidences from the previous chapters indicate that the new transportation infrastructure had relevant impacts on travel time and on the economic activity in the stations' vicinities. Nevertheless, these reduced-form approaches do not allow me to infer overall effects on inequality and welfare, and its mechanisms. Hence I develop a model of internal city structure that features high- and low-skilled workers, production and residential externalities, and heterogeneous city blocks. The model builds on a recent quantitative urban model (Ahlfeldt et al. (2015)) and extend it to include heterogeneous workers. I estimate the general equilibrium effects of transportation infrastructure expansion on wages, employment, inequality, productivity and welfare.

In the model, the city is defined as a set of heterogeneous blocks. Blocks differ in the exogenous characteristics: floor space, amenities, productivity and commuting times. Amenity relates to characteristics associated with higher utilities from living in that block. For example, proximity to the beach and scenery views. Productivity relates to characteristics that increase firms' productivity, such as proximity to the port area. Commuting times depend on the transportation infrastructure available in the residence and workplace block.

Workers face two decisions. First, whether to move to the city or not. Workers move if expected utility on the city is higher than the reservation utility level. If they do move, workers observe idiosyncratic preference shocks for where to live and work within the city and choose the residence-workplace pair of blocks to maximize utility. Workers prefer pairs with lower rents, higher amenities and wages, and low commuting times.

Each block has a representative firm that uses floor space, high- and low-skilled workers to produce a numeraire good. The market is perfectly competitive and wages adjust to equilibrate supply and demand. Firms, high- and low-skilled workers bid for floor space, so that the proportion of floor space used for productive or residential purposes is endogenously determined.

Besides exogenous characteristics, block also differ due to endogenous agglomeration forces. Total amenities depend on exogenous characteristics and on the density of residents in each block. Higher density increases the utility from living in block i. Symmetrically, total productivity depends on exogenous characteristics and on the density of workers. These endogenous agglomeration forces generate residential and productivity externalities.

The source of inequality in the model is the existence of production agglomeration externalities specific to high-skilled workers. Because of this additional agglomeration force, high-skilled workers yield higher wages since they obtain higher gains from agglomeration. In turn, this agglomeration force can impact segregation through two mechanisms. First, high-skilled jobs will be more concentrated geographically. Thus, depending on the transport infrastructure, residence decision may also be more concentrated around workplace in order to reduce commuting costs. Second, higher agglomeration forces lead to higher wage inequality. Since high- and low-skilled workers bid for floor space, high-skilled workers will concentrate in high amenities residence locations. Higher prices will push low-skilled workers out of these locations, a phenomenon known as gentrification. In the presence of poor transport infrastructure, commuting costs increase rapidly with distance. This exacerbates both mechanisms and give rise to highly segregated cities, where high-skilled workers agglomerate close to the city center and low-skilled workers live in the outskirts of the metropolitan area. This is a common configuration of developing world metropolises, including Rio de Janeiro.

In order to estimate the structural model, I combine information on residence and employment for each skill group inside each city block in 2010. Besides, commuting times between all city blocks (57,122 combinations) are computed using random forest regression and data from restricted access origin-destination survey (2011). Structural parameters are determined according to a 3-step estimation procedure that involves calibration, generalized method of moments and grid search. Estimated parameters indicate large productivity gains from agglomeration for all workers and even larger for high-skilled workers. Besides results suggest that exogenous characteristics, such as proximity to the beach, are much more relevant for total amenities than agglomeration forces, represented by residential density.

The estimated model is then used to perform counterfactual exercises to assess the impacts of the recent transport infrastructure expansion in Rio de Janeiro. In particular, I compare equilibrium outcomes using 2018 Google Maps and counterfactual travel times to infer the overall and distributional effects of transport investments. I consider three counterfactuals scenarios: without BRT, without the subway extension and without both investments.

Results point that transport investments increased the welfare of high- and low-skilled workers. Nevertheless, high-skilled workers experienced a larger increase, raising inequality. The expansion of the transport infrastructure connected new locations with Rio's central business district, which increased the number of residential options with lower commuting costs. This led to a reduction in the concentration of residents and increased the concentration of jobs. Both effects are stronger for high-skilled workers, which raises residential and employment segregation. In particular, a gentrification process took place: higher demand and prices in the newly connected area led to an increase in residential segregation. This process was exacerbated by the fact that, among the newly connected areas, some had high amenities due to proximity to the beach.

The chapter relates to literature strand that uses general equilibrium models to asses agregate and distributional effects of transportation improvements (Ahlfeldt et al. (2015), Redding and Rossi-Hansberg (2017)). In particular, my work is closest related to Tsivanidis (2018), that looks at the aggregate and distributional effects of TransMilenio, Bogota's BRT system. Also based on Ahlfeldt et al. (2015), the model introduces multiple types of workers by incorporating multiple types of firms with different demand for worker groups. The author finds that while the system caused increases in welfare and output larger than its cost, gains accrued slightly more to high-skilled workers. Results suggest an increase in residential segregation by skills. Differently from Tsivanidis (2018), I observe workers' educational level, which enables me to directly introduce heterogeneous workers in the model and estimate how agglomeration forces contribute to wage inequality. Thus my model incorporates the main mechanism that describes how city growth results in higher inequality and spatial segregation. Besides, I contribute to the literature by estimation the effects of different transportation technologies in the same framework.

3.2 Model

The city is composed of heterogeneous blocks, where high and low-skilled workers choose where to live and work. City blocks differ in four dimensions: residential amenities, final good productivity, supply of floor space, and transport infrastructure. The access to transport infrastructure determines travel times between pairs of city locations. Each block has a fixed supply of floor space, which can be used for residence or production.

High- and low-skilled workers choose residence and workplace pair that maximizes their utility, according to residential amenities, rents, wages, and commuting times between locations. Both types of workers observe the same residential amenities, floor prices and commuting times. But, they (potentially) face different commuting costs and wages. For example, high- and low-skilled workers can make different choices of transportation modes.

Homogeneous firms use floor space and high- and low-skilled workers to produce a single final good. Firms and workers post rent offers in each block and prices adjust to clear markets.

Residential amenities and final good productivity are subject to agglomeration externalities. Residential amenities are determined by exogenous characteristics (e.g. proximity to the beach) and the density of residents. If an additional worker decides to live in a block, the residential amenity increases for all residents in that block. The increase in amenities has the same magnitude independently if the additional resident is high or low-skilled. Final good productivity differs among locations, depending on exogenous characteristics and employment density. Different from amenities, high-skilled workers increase productivity by a higher amount than low-skilled workers. Consequently, high-skilled workers yield higher wages due to larger productivity gains from agglomeration.

City structure is then determined by the interaction between agglomeration forces (residential and production externalities) and dispersion forces (commuting costs and inelastic supply of floor space). In particular, the degree of segregation between high and low-skilled workers depends on the difference of magnitude of these forces between worker's type. For example, since production externalities are potentially stronger for high-skilled workers, they are subject to stronger agglomeration forces. On the other hand, high-skilled workers may face lower commuting costs due to different transport mode choices.

3.2.1 Model Setup

A city is defined as a set of discrete locations (blocks), indexed by $i \in I = \{1, .., S\}$. Each block i has a supply of floor space L_i, residential amenities B_i, final goods productivity A_i and a vector of travel times between location i and all other j locations in the city τ_{ij}. There are two groups of works: high- and low-skilled($f = \{H, L\}$).

The city belongs to a wider economy and is populated by endogenous measures of high- and low-skilled workers, respectively H_H and H_L. Workers face no cost of moving within the city and the wider economy. If expected utility in the city is higher that reservation utility level, workers move to the city and observe idiosyncratic utility shocks for each possible pair of residence and employment location. High- and low-skilled workers face different reservation utility levels - $\bar{U}_H$ and $\bar{U}_L$-, and different distribution of utility shocks. Workers pick the residence and workplace pair that maximize utility. Thus, in equilibrium the urban population is determined so that utility from living in the city is equal to the reservation utility for each workers' type.

Residential amenities (B_i) represent how attractive block i is as a residence location. Specifically, amenities depends on an exogenous component (b_i) and an endogenous component (Ω_i). The exogenous component captures fundamental characteristics such as proximity to the beach and scenic views. The endogenous component captures residential externalities, depending on the residential density $\left(\frac{H_{Ri}}{N_i}\right)$ in block i and on the strength of residential externalities (η).

$$B_i = b_i \Omega_i, \qquad \Omega_i = \left(\frac{H_{Ri}}{N_i}\right)^{\eta} \tag{3-1}$$

The commuting cost is modeled as $d_{ijf} = e^{k_f \tau_{ij}} \in [1, \infty]$, where commuting time between block i and j is measured in minutes. The constant k_f is type-specific and regulates the size of commuting costs for high- and low-skilled workers.

Firms produce a single final good using high- and low-skilled workers and floor space. Final goods are traded without cost within the city and the wider economy and chosen as the numeraire ($p = 1$). Firms are attracted to blocks with higher productivity, represented by A_j . Symmetrically to residential amenities, A_j is composed of an exogenous component (a_j) and an endogenous component (Υ_i) - that represent agglomeration externalities. Total productivity depends on total employment density $\left(\frac{H_{MHj}+H_{MLj}}{N_j}\right)$ in block j and on the strength of production externalities:

$$A_j = a_j \Upsilon_j, \quad \Upsilon_j \equiv \left(\frac{H_{MHj} + H_{MLj}}{N_s}\right)^{\lambda} \tag{3-2}$$

To close the model, a construction sector supplies floor space and uses geographical land N and capital K as inputs. Each block has an effective supply of floor space L_i, used for production (θ_i) or residence ($1 - \theta_i$). The final good and construction markets are perfectly competitive and have constant returns to scale. For simplicity, I suppose absentee landlords.

3.2.2 Workers

Worker o, skill group f, living in block i and working in block j has utility equal to:

$$U_{ijfo} = \frac{B_i z_{ijfo}}{d_{ijf}} \left(\frac{c_{ijfo}}{\beta}\right)^{\beta} \left(\frac{l_{ijfo}}{1-\beta}\right)^{1-\beta}, \qquad 0 < \beta < 1 \tag{3-3}$$

Workers derive utility from consumption of the final single good (c_{ijfo}), consumption of residential floor space (l_{ijfo}), residential amenities for living in block i (B_i), disutility from commuting from residence block i to workplace block j, and an idiosyncratic preference shock for the pair residence-workplace (z_{ijfo}). Workers are risk neutral.

The idiosyncratic shock to worker's preference for residence-workplace pair is drawn from a Fréchet distribution, $F(z_{ijfo}) = e^{-z_{ijfo}^{\epsilon_f}}$. The shape parameter ($\epsilon_f > 1$) is also type-specific and controls the dispersion of idiosyncratic utility.

After observing the idiosyncratic preference shock, workers decide where to live and work according to residential amenities, price of floor space, workplace wages, and commuting costs between residence and workplace. Both high- and low-skilled workers face the same amenities B_i and floor prices Q_i, but different shock distribution z_{ijfo}, commuting costs d_{ijf} and wages w_{jf}. The worker solves the problem:

$$Max_{(i,j,c,l)} \quad U_{ijfo} = \frac{B_i z_{ijfo}}{d_{ijf}} \left(\frac{c_{ijfo}}{\beta}\right)^{\beta} \left(\frac{l_{ijfo}}{1-\beta}\right)^{1-\beta} \qquad s.a. w_{jf} = c_{ijfo} + Q_i l_{ijfo} \tag{3-4}$$

Considering optimal choices of final good and floor space consumption, indirect utility of living in block i and working in block j takes the form:

$$u_{ijfo} = \frac{B_i z_{ijfo} w_j Q_i^{\beta-1}}{d_{ijf}} \tag{3-5}$$

Since the idiosyncratic preference shock has a Fréchet distribution, the probability of choosing residence-workplace pair ij is proportional to the utility of that choice against the sum of utilities in all possible residence-workplace pairs in the city. Specifically, the probability that worker type f lives in block i and works in block j is equal to:

$$\pi_{ijf} = \frac{(d_{ijf} Q_i^{1-\beta})^{-\epsilon_f} (B_i w_{jf})^{\epsilon_f}}{\sum_{r=1}^{S} \sum_{s=0}^{S} (d_{rsf} Q_r^{1-\beta})^{-\epsilon_f} (B_r w_{sf})^{\epsilon_f}} \tag{3-6}$$

which means that blocks with higher amenities, lower floor prices and lower commuting times to workplace locations will attract more residents. And blocks with higher wages and lower commuting times to residence locations will attract more workers. Summing across all possible workplace locations defines the probability of living in block i. Symmetrically, summing across residence locations defines the probability of working in block j.

$$\pi_{Rif} = \sum_{s=1}^{S} \pi_{isf} \qquad\qquad \pi_{Mjf} = \sum_{r=1}^{S} \pi_{rjf} \tag{3-7}$$

Consequently, conditional on living in block i, the probability of choosing residence-workplace pair ij depends only on the workplaces characteristics of city blocks: wages and commuting costs.

$$\pi_{ijf|i} = \pi_{ijf} / \pi_{Rif} = \frac{(w_{jf}/d_{ijf})^{\epsilon_f}}{\sum_{s=1}^{S} (w_{sf}/d_{isf})^{\epsilon_f}} \tag{3-8}$$

3.2.3 Production

Homogeneous firms provide a single final tradable good using labor and floor space. The final good market is perfectly competitive and firms have constant returns to scale. Production function follows a Cobb-Douglas form between total labor (H_{Mj}) and floor space used commercially (L_{Mj}). Output in block j is equal to

$$Y_j = A_j H_{Mj}^{\alpha} L_{Mj}^{1-\alpha}, \tag{3-9}$$

where A_j represents total factor productivity in block j.

Total labor (H_{Mj}) is a combination of high- (H_{MHj}) and low-skilled workers (H_{MLj}). It takes the form of a CES function:

$$H_{Mj} = (C_j (H_{MHj})^{1/\delta} + (H_{MLj})^{1/\delta})^{\delta} \tag{3-10}$$

High-skill workers have higher gains from agglomeration, represented by the productivity term C_j . This agglomeration externality depends exclusively on high-skill employment density:

$$C_j = \left(\frac{H_{MHs}}{N_s} \right)^{\lambda_H}, \tag{3-11}$$

where λ_H represents high-skill externality strength.

3.2.4
Land Market

Finally, a perfectly competitive construction sector supplies floor space using land (N_i) and capital (K_i) as inputs. The construction sector production function takes the form of a Cobb-Douglas function:

$$L_i = K_i^{\mu} N_i^{1-\mu} \tag{3-12}$$

The construction sector firms maximize profit taking as given input prices. Since capital price is the same across city blocks, floor space price will only differ between city blocks because of different land prices. Supply of floor space can be simplified to $L_i = \varphi_i N_i^{1-\mu}$, where φ_i captures the capital intensity, or the density of development, in block i.

3.2.5
Equilibrium

Given the model's parameters $\{\alpha, \delta, \phi, \beta, \epsilon_H, \epsilon_L, k_H, k_L, \lambda, \lambda_H, \eta\}$, the reservation levels of utility in the wider economy$\{\bar{U}_H, \bar{U}_L\}$ and the exogenous location-specific characteristics $\{a, b, \varphi, N, \tau\}$ the general equilibrium of the model is referenced by nine vectors $\{\boldsymbol{\pi_{MH}}, \boldsymbol{\pi_{ML}}, \boldsymbol{\pi_{RH}}, \boldsymbol{\pi_{RL}}, Q, \boldsymbol{q}, \boldsymbol{w_L}, \boldsymbol{w_H}, \boldsymbol{\theta}\}$ and total population per worker's skill level $\{H_H, H_L\}$ such that:

1. Firms maximize profits and have zero profits in each employment location within the city (equations 3-13, 3-14 and 3-15);

2. Workers maximize utility between residence and employment location pairs within the city (equation 3-18);

3. Commuting market clearing (equation 3-17);

4. Land market clearing (equations 3-19, 3-22, 3-23 and 3-25)

Firms choose their block of operation and input quantities to maximize their profits given productivity measures A_j and C_j, other firms' and workers' decisions, and final good and input prices. First order conditions imply that equilibrium prices are:

$$w_{Hj} = \alpha A_j C_j H_{Mj}^{(\delta\alpha-1)/\delta} H_{MHj}^{(1-\delta)/\delta} L_{Mj}^{1-\alpha} \tag{3-13}$$

$$w_{Lj} = \alpha A_j H_{Mj}^{(\delta\alpha-1)/\delta} H_{MLj}^{(1-\delta)/\delta} L_{Mj}^{1-\alpha} \tag{3-14}$$

$$q_j = (1-\alpha) A_j H_{Mj}^{\alpha} L_{Mj}^{-\alpha} \tag{3-15}$$

In equilibrium, firms have zero profits in all city blocks, which implies that every increase in productivity will be compensated by an increase in input prices - wages or rents.

Dividing equation 3-13 per equation 3-14, it is possible to define the wage premium as a function of the ratio of high- to low-skilled workers and the size of high-skilled agglomeration externalities in block j:

$$w_{Hj}/w_{Lj} = C_j(H_{MHj}/H_{MLj})^{(1-\delta)/\delta} \tag{3-16}$$

Since $\delta > 1$, if an additional high-skilled worker decides to commute to block j, she will impact the wage premium in two opposite directions. First, the workers' high- to low-skilled ratio increases, reducing the wage premium. Second, the high-skilled employment density increases, which leads to a higher wage premium. For equilibrium existence, the overall effect has to be negative. This condition imposes an upper bound to high-skill agglomeration strength: $\lambda_H < \frac{\delta-1}{\delta}$. More importantly, equation 3-16 shows that the equilibrium wage premium depends on the magnitude of these two parameters. If the elasticity of substitution is relatively low and high-skill agglomeration externality is high, wage premiums and, consequently, equilibrium wage inequality will be higher.

Concerning workers, equilibrium implies two conditions. First, the number of workers in workplace location j is equal to the sum of commuters to j across all residence locations. The commuting market clearing condition is defined as:

$$H_{Mfj} = \sum_{i=1}^{S} \pi_{ijf|i} H_{Rfi} \tag{3-17}$$

Second, the expected utility of moving to the city is equal to reservation utility level for each worker type. Workers move to the city if the utility of living in the city is higher than living in the wider economy. As more workers enter the city, congestion forces take place, which increases floor prices and, consequently, diminishes utility until workers are indifferent between the city and the wider economy. City size is then determined endogenously by the following equation:

$$E[u_f] = \gamma_f \left[\sum_{r=1}^{S} \sum_{s=0}^{S} (d_{rsf} Q_r^{1-\beta})^{-\epsilon_f} (B_r w_{sf})^{\epsilon_f} \right]^{1/\epsilon_f} = \bar{U_f} \tag{3-18}$$

where $\gamma_f = \Gamma(\frac{\epsilon_f - 1}{\epsilon_f})$ and $\Gamma(.)$ is the gamma function.

Firms and high- and low-skilled workers bid for floor space and the prices adjust to clear markets. Then, the observed prices in data are the maximum between equilibrium prices in residential and production markets:$\mathbb{Q}_i = max\{q_i, Q_i\}$ As in Ahlfeldt et al. (2015), the model allows a wedge between prices of floor space used for residence and production due to land use regulations. For simplicity, I suppress notation. No arbitrage condition between markets implies that:

$$\mathbb{Q}_i = q_i \; if \; q_i < Q_i \qquad \theta_i = 1 \tag{3-19}$$

$$\mathbb{Q}_i = q_i \; if \; q_i = Q_i \qquad \theta_i \in [0,1] \tag{3-20}$$

$$\mathbb{Q}_i = Q_i \; if \; q_i < Q_i \qquad \theta_i = 0 \tag{3-21}$$

where θ_i is the share of floor space used for production. In equilibrium, the market clearing conditions impose that floor space supply and demand are equal in residential and production markets. From equation 3-15, the firms' first order condition, the production market clearing condition is:

$$\left(\frac{(1-\alpha)A_j}{q_j}\right)^{1/\alpha} H_{Mj} = \theta_j L_j \tag{3-22}$$

Since workers spend a constant share $(1-\beta)$ of their wages in floor space consumption, residence market clearing condition equals:

$$\frac{(1-\beta)}{Q_i}\left[E[w_{Hs}|i]H_{RHi} + E[w_{Ls}|i]H_{RLi}\right] = (1-\theta_i)L_i \tag{3-23}$$

where expected wage per residence location is defined as:

$$E[w_{fs}|i] = \sum_{s=1}^{S} \pi_{isf|i} w_{fs} \tag{3-24}$$

With market clearing and profit maximization, total demand equals total supply:

$$(1-\theta_i)L_i + \theta_i L_i = \varphi_i N_i^{1-\mu} \tag{3-25}$$

3.3 Data

To estimate structural parameters I construct an extensive database for the Rio de Janeiro Metropolitan Area. The database combines information on workplace and residence employment for each skill group inside each city block, commuting times between all city blocks and floor prices. Additionally, I compile data moments on commuting flows, dispersion of wages per residence and work location. Below, I describe the dataset.

3.3.1 Residence

City blocks correspond to the 2010 census statistical areas. The census divided the Rio de Janeiro MSA into 338 statistical areas. Since this is the finest available disaggregation for residence employment per skill group, I chose the statistical area as my unit of observation.

Residence employment, commuting flows, rents and wages per residence location come from Census 2010. I count the number of low and high-skilled workers, between 18 and 70 years old that live in each block. Workers with education level up to high school are considered low-skilled. Workers with a college degree are categorized as high-skilled.

Concerning commuter behavior, Censo asks workers about their one-way commuting time span: up to 30 minutes, between 30 minutes and 1 hour, between 1 and 2 hours and more than 2 hours. I compile the proportion of workers that commute up to an hour in the Rio de Janeiro metropolitan area for each skill level.

Floor prices are computed using rent prices and housing characteristics: average rent price, average rent price per room and average rent price per bedroom. Finally, I compute the dispersion of average wages per skill level and residence location.

3.3.2 Work

Workplace employment comes from RAIS 2010, a restricted access administrative data set that contains firm-level information on address, number of workers, workers' skill level and wages. Firms' addresses were geocoded and matched to the 2010 census block structure. One drawback of using RAIS data is that it only comprises the universe of formal workers in Brazil. Consequently, the population of workers in the Census outnumbers the population of workers in RAIS. To adjust the data, I multiply workplace employment to match the Census' total population. Tsivanidis (2018) uses the same adjustment procedure and shows that the distribution of informal employment is equivalent to the distribution of formal employment for Bogotá, Colombia. Unfortunately, since there is no information available on the within city distribution of informal workers for Rio de Janeiro, I cannot test this hypotheses. If the distribution of formal and informal employment is different, this will introduce a measurement error and can bias estimates.

3.3.3 Commuting time

The quantitative analyses of the model requires information on travel times between all city blocks, which represents a 338x338 travel time matrix. To compute this matrix, I use the 2011 travel time information from restricted access origin-destination survey for Rio de Janeiro Metropolitan Area. Nevertheless, since this household survey does not have information on travel times between all blocks (57,122 combinations), I need to predict commuting times out of sample. To accomplish this, I use the origin-destination dataset to estimate a travel time production function using random forest regression. Then I predict commuting times between all city blocks.

Finally, to perform counterfactual analyses, I conduct a two step procedure. First, I collect the 2018 travel times from Google Maps API for the same sample as the origin-destination survey. I construct a panel data set and regress the travel time difference between 2011 and 2018 in a set of dummy variables that characterize the transport infrastructure built in the period. This allows me to estimate counterfactual travel times in the absence of the transport infrastructure investments. In particular, I estimate counterfactual times with no investments, without only the BRT stations and without only the new subway stations. Second, I repeat the procedure conducted for the 2011 origin-destination sample: estimate a travel time production function using random forest regression and predict counterfactual travel time matrices.

3.4 Estimation

To estimate structural parameters of the model, I follow a three-step estimation strategy. First, I calibrate production functions and workers' utility parameters $\{\alpha, \delta, \phi, \beta\}$ according to estimates from the literature. Second, I use moments from observed data to estimate the structural parameters $\{\epsilon_H, \epsilon_L, k_H, k_L, \lambda_H\}$ using GMM. Third, I simulate the model to perform a grid search over parameter space to calibrate $\{\lambda, \eta\}$ that are consistent with the data being an equilibrium of the model.

Next, I describe each estimation step in detail. Finally, I discuss model identification and present results.

3.4.1 Calibration

I set the share of consumer expenditure on residential floor space $(1-\beta)$ equal to 0.25, consistent with Davis and Davis and Ortalo-Magné (2011) and data on consumer expenditure available for the Rio de Janeiro Metropolitan Area (POF 2008-2009). Concerning the production function, I set the share of floor space in firms cost $(1-\alpha)$ equal to 0.2, consistent with Valentinyi and Herrendorf (2008). In line with estimates from Card (2009) and estimates from Pecora and Menezes-Filho (2014) for Brazilian economy, the elasticity of substitution between high- and low-skilled workers (δ) is equal to 1.3. The share of land in construction costs $(1-\phi)$ is set to 0.25, consistent with Combes et al. (2012) and Epple et al. (2010) estimates.

3.4.2 Generalized Method of Moments

I use five moments for observed data to estimate parameters $\{\epsilon_H, \epsilon_L, k_H, k_L, \lambda_H\}$: share of workers with commuting time up to 60 minutes per skill group; dispersion of residence wages per skill group and dispersion of high-skill workplace wages. Next I discuss each moment condition.

First, I estimate the semi-elasticity of workers decision to commuting time ($v_f = \epsilon_f k_f$), using the share of workers with commuting time up to 60 minutes per skill group from 2010 Census. I rewrite the commuting market clearing condition as a function of transformed wages ($\omega_{jf} = w_{jf}^{\epsilon_f}$) and commuting costs ($e^{v_f \tau_{ij}} = d_{ij}^{\epsilon_f}$), as following:

$$H_{Mjf} = \sum_{i=1}^{S} \frac{(\omega_{jf}/e^{v_f \tau_{ij}})}{\sum_{s=0}^{S}(\omega_{sf}/e^{v_f \tau_{sj}})} H_{Rif}, \qquad f = \{H, L\} \tag{3-26}$$

Since the vectors$\{H_{Mf}, H_{Rf}, \tau\}$ are observed in data, for each value of v_f, I can pin down the equivalent transformed wages[1]. With the transformed wages (ω_{jf}) in hand, I use the moment condition below to estimate v_f:

$$E[\Psi . H_{Mjf} - \sum_{i=1}^{\aleph_j} \frac{(\omega_{jf}/e^{\vartheta_f \tau_{ij}})}{\sum_{s=0}^{S}(\omega_{sf}/e^{\vartheta_f \tau_{sj}})} H_{Rif}] = 0, \qquad f = \{H, L\} \tag{3-27}$$

where Ψ is the share of workers that commute up to 60 minutes in data, and $\aleph_j$ are the locations with commuting time up to 60 minutes from block i. So the estimation algorithm consists of a fixed point estimation with a embedded minimization problem.

Second, I use the dispersion of residence wages per skill group to estimate $\{\epsilon_H, \epsilon_L\}$. Taken as given transformed wages (ω_{jf}) and transformed commuting costs ($e^{v_f \tau_{ij}}$), average wages per residence location can be defined as a function of ϵ_f, the Fréchet shape parameter that controls the dispersion of idiosyncratic utility.

[1]Transformed wages are identified up to a normalization. I set the transformed wages geometric mean equal to one.

$$E[w_{sf/i}] = f(\epsilon_f|v_f, \omega_{jf,}\tau_{ij}) = \sum_{s=1}^{S} \frac{(\omega_{sf}/e^{\vartheta_f \tau_{is}})}{\sum_{j=0}^{N}(\omega_{jf}/e^{\vartheta_f \tau_{ij}})} \omega_s^{1/\epsilon_f}, \qquad f = \{H, L\} \tag{3-28}$$

Then I estimate ϵ_f to minimize the distance between the variance of log residential wages in the model and observed in data. The moment condition is:

$$E[ln(E[w_{sf/i}])^2 - \sigma^2_{ln\,E[w_{j/i}]}] = 0, \qquad f = \{H, L\} \tag{3-29}$$

Third, I set λ_H to minimize the distance between the variance of log high-skilled workplace wages in the data and the model. From equation 3-16, conditional on low-skill workplace wages (ω_{jL}), ratio of high- to low-skilled workers $\left(\frac{H_{MHj}}{H_{MLj}}\right)$, density of high-skilled workers $\left(\frac{H_{MHs}}{N_s}\right)$ and parameter δ, high-skill workplace wage is a function of high-skill agglomeration externality parameter λ_H:

$$w_{Hj} = f(\lambda_H|w_{Lj,}H_{MHj}, H_{MLj,}N_j, \delta) = w_{Lj}\left(\frac{H_{MHs}}{N_s}\right)^{\lambda_H}\left(\frac{H_{MHj}}{H_{MLj}}\right)^{\frac{1-\delta}{\delta}} \tag{3-30}$$

Moment condition is defined as:

$$E[ln(w_{Hj})^2 - \sigma^2_{ln\,w_{Hj}}] = 0 \tag{3-31}$$

I use these five moment functions in a two-step GMM procedure to estimate the parameter vector $\Phi = \{v_H, v_L, \epsilon_H, \epsilon_L, \lambda_H\}$, where $k_f = v_f/\epsilon_f$.

3.4.3 Grid Search

In order to calibrate parameters $\{\lambda, \eta\}$, I perform a grid search over parameter space. I simulate the model with 400 different parameters combinations and pick the pair that best matches the observed distribution of high- to low-skilled residents ratio. This grid search requires a 3-step procedure. First, with calibrated and estimated parameters values $\{\alpha, \delta, \phi, \beta, \epsilon_H, \epsilon_L, k_H, k_L, \lambda_H\}$ and observed variables $\{H_{MH}, H_{RH}, H_{ML}, H_{RL}, \mathbb{Q}, \tau, N\}$, I use model equilibrium equations to obtain equilibrium vectors$\{A, B, \varphi, w_L, w_H, \theta\}$ and reservation utility levels$\{\bar{U}_H, \bar{U}_L\}$. Second, I simulate the model for different combinations of $\{\lambda, \eta\}$ using the parameters $\{\alpha, \delta, \phi, \beta, \epsilon_H, \epsilon_L, k_H, k_L, \lambda_H\}$ and exogenous characteristics $\{\varphi, \tau, N, \bar{U}_H, \bar{U}_L\}$. Due to the possibility of multiple equilibrium, I select the equilibrium closest to the endogenous variables previous obtained. So the vectors$\{H_{MH}, H_{RH}, H_{ML}, H_{RL}, \mathbb{Q}, A, B, w_L, w_H, \theta\}$

serve as an initial guess for the simulation. Third, I pick the pair $\{\lambda, \eta\}$ with the minimal distance between simulated and observed moments. Due to the shape of the distribution of high- to low-skilled residents ratio, I chose the median, skewness and kurtoses as moment conditions.

3.4.4 Identification

Now I discuss the identification of estimated parameters $\{v_H, v_L, \epsilon_H, \epsilon_L, \lambda_H\}$. Moment conditions 3-27 show that a higher value of v means that workers choices are more responsive to commuting times. Consequently, the probability of working in block j declines more rapidly with commuting time. Then the share of workers with commuting time up to 60 minutes would be smaller. Concerning the dispersion of worker's preference, equation 3-28 indicate that a higher value of the Frechet shape parameter would make worker's choices of residence-workplace more similar. Conditional on residence location, variance of average wage would be smaller. Finally, the identification of the high-skill agglomeration externality parameter relies on equation 3-30. If the value of λ_H was larger, high-skill productivity gains from agglomeration would also be larger. So high-skilled workers would be more concentrated geographically and variance of workplace wages would increase. Monte Carlo simulations results, reported in the appendix (Figure 2), give evidence that moment conditions identify structural parameters. I create an hypothetical the city using know parameters and show that the estimation procedure recovers true parameters values.

Concerning the parameters $\{\lambda, \eta\}$, I argue that each parameter will impact the ratio of high to low-skilled residents distribution in different margins, which permits identification of both parameter's values. Lower values of λ mean that the general agglomeration externalities are relatively weak relative to high-skill agglomeration externalities. Thus high-skilled workers would be relatively much more concentrated in a few workplace locations, which would impact residence choices. In order to reduce commuting costs, high-skilled workers would also have an incentive to agglomerate in residence location close to workplace. The skewness and kurtosis of the distribution of high to low-skill ratio per residence location would be larger. Additionally, due to this difference in production agglomeration externalities, a marginal larger value of η will attract more high- than low-skilled workers to move to the city. The overall population ratio of high- to low-skill workers would increase, which would increase the median of the distribution. At the same time, since

Table 3.1: GMM Results

Parameter	Estimate
v_H	0.12***
	(0.004)
v_L	0.15***
	(0.04)
ϵ_H	2.84***
	(0.31)
ϵ_L	3.8***
	(1.11)
λ_H	0.11***
	(0.03)

Notes: Estimates are from two-step efficient GMM procedure. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

the parameter is the same for the two groups of workers, it does not affect right-tail of the distribution.

3.4.5 Results

Table 3.1 reports the GMM estimates. It is important to draw attention to three points. First, results show that low-skilled workers are more responsive to commuting times since $v_L > v_H$. Although low-skilled workers have lower commuting costs ($k_H > k_L$), their preference are less disperse which make them more capable to adjust residence-workplace choices to reduce commuting costs. Second, estimates indicate large productivity gains from agglomeration. High-skill productivity externality is equal to 0.11 (table 3.1) and grid search picked general productivity externality equal to 0.18. Most of the estimates from the literature lie within 0.03-0.08, although Greenstone et al. (2010) report an estimate of 0.12 and Kline and Moretti (2013) of 0.2. Nevertheless, all this estimates use data from developed countries. Closer to my estimates, Tsivanidis (2018) estimate productivity externality for Bogotá equal to 0.23. Concerning high-skill productivity externality, I cannot compare estimates since I believe this is first paper to estimate intra-city parameter for developing countries. Third, from grid search procedure, amenities agglomeration externality estimate is 0.02, a significant lower value if compared with estimates for Berlin: 0.11 (Ahlfeldt et al. (2015)). This means that exogenous characteristics are much more relevant for amenities in Rio de Janeiro Metropolitan Area than agglomeration forces.

Now I discuss model fit of GMM and grid search estimation procedure

Table 3.2: GMM Model Fit

	Model	Data
Panel A: Targeted moments		
% commuters up to 60 minutes		
High-skill	68.17	68.17
Low-skill	64.82	64.82
Dispersion of residence wages		
High-skill	0.096	0.096
Low-skill	0.059	0.059
Dispersion of workplace wages		
High-skill	0.30	0.30
Panel B: Non-targeted moments		
Workplace wages		
High-skill		
1th quintile	0.62	0.58
2th quintile	0.80	0.77
3th quintile	1.00	0.97
4th quintile	1.24	1.25
5th quintile	1.64	1.62
Low-skill		
1th quintile	0.68	0.83
2th quintile	0.83	0.90
3th quintile	0.99	0.96
4th quintile	1.20	1.04
5th quintile	1.48	1.19

Notes: Estimates are from two-step efficient GMM procedure.

(Tables 3.2 and 3.3). Table 3.2, panel B, shows that model‘s prediction of high-skill workplace wages are more accurate than for low-skill. Two comments are in order. First, data from workplace wages come from RAIS, which covers only workers with a formal contract. Since informality is higher among low-skilled workers, low-skill workplace wages are more prone to measurement error than high-skill wage. Second, low-skilled workers include all educational levels up to incomplete college degree. Thus, this category can include highly heterogeneous workers, which harms model fit. For future work, I intend to estimate the model for three worker's skill groups.

3.5 Counterfactual

To asses the overall and distributional effects of urban mobility investments, I perform counterfactual exercises using previously estimated counterfactual commuting times. Counterfactual equilibrium - shutting down

Table 3.3: Grid Search Model Fit

	Model	Data
Panel A: Targeted moments		
Ratio high/low skill residents		
Median	0.11	0.11
Kurtosis	2.46	2.49
Skewness	10.08	8.83
Panel B: Non-targeted moments		
Ratio high/low skill residents		
1th quartile	0.07	0.04
2th quartile	0.10	0.08
3th quartile	0.12	0.14
4th quartile	0.13	0.32

BRT and subway extension, only BRT and only subway extension stations - are compared with estimated equilibrium with 2018 travel time matrix.

All tables and figures report the percentage change in outcomes. Table 3.4 presents overall results, while tables 3.5, 3.6 and 3.7 explore heterogeneous effects by workers' skill level. Since total population in Rio de Janeiro remained fairly stable between 2010-2018, I assume a close city hypotheses in simulations, which means that total high- and low-skilled population are constant and utility level will adjust accordingly.

The first column in table 3.4 shows the effects of shutting down all BRT and new subway stations: GDP, rents and inequality reduce by a significant amount. Although the welfare of both worker's group diminish, high-skilled workers experience a larger reduction, which ends up reducing inequality.

To uncover mechanisms, I calculate the herfindhal and dissimilarity index for the distribution of jobs and residents in the city[2]. In the absence of investments, concentration of jobs would diminish, while concentration of residents would increase. In turn, segregation between worker's type decreases for jobs and enhances for residents. Figure 3.1 presents the percentage change in the number of residents and jobs in the absence of the BRT and the subway extension. Results suggest that transport investments led to an agglomeration of economic activity and a sprawl of residents.

Second and third column of table 3.4 present estimates of shutting down only BRT or only the subway extension. Results show evidence of large heterogeneous effects. BRT leads to a decrease in concentration and

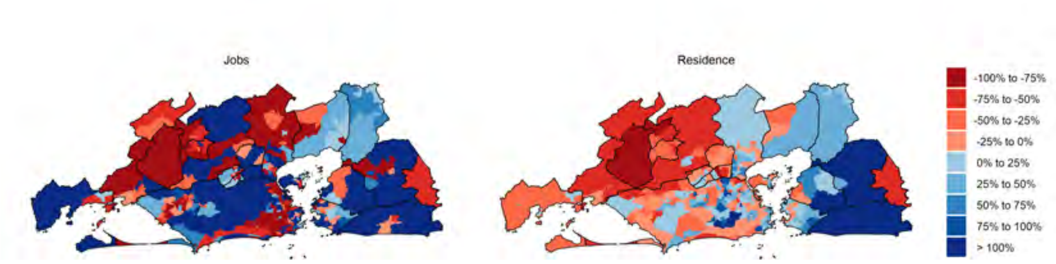
[2]The herfindhal index is a measure of concentration, where higher values point to higher concentration. The dissimilarity index measures the degree of segregation: the percentage of workers that would have to move so that high- and low-skilled workers are equally distributed throughout the city.

Table 3.4: Contrafactuals

	No BRT	No subway extension	No BRT No subway extension
GDP	4.4	-6.8	-11.1
Rents	7.7	-4.0	-8.6
Welfare Low	-33.4	1.4	-41.9
Welfare High	-38.2	-0.7	-48.6
Inequality	-7.2	-2.1	-11.6
Jobs			
Herfindhal Index	2.3	-10.1	-27.9
Dissimilarity Index	14.6	-4.4	-2.5
Residents			
Herfindhal Index	3.4	7.5	2.7
Dissimilarity Index	17.5	-5.4	1.3

Notes: The simulation uses as initial point the estimated equilibrium in 2010.

Figure 3.1: Contrafactuals, no BRT and subway extension: number of residents and jobs



Notes: The simulation uses as initial point the estimated equilibrium in 2010.

segregation of jobs, which ends up reducing GDP. On the other hand, the subway extension induced higher concentration of jobs and segregation, evincing larger agglomeration effects for high-skilled workers. Concerning residents location, both investments led to city sprawl. Nevertheless, while BRT reduces segregation, subway enhances it. This suggests that subway investments induced a process of gentrification. The fact that, without the subway, low-skilled welfare would be higher, and rents lower, reinforces the argument.

Table 3.5 describes results per skill type and transport investment. Considering results in the absence of all investments (first column), there are two main points to highlight. First, for both skill groups, BRT and subway

Table 3.5: Contrafactuals: concentration and segregation

	No BRT	No subway extension	No BRT No subway extension
Panel A: Jobs			
Herfindhal Index			
High-skill	28.9	-11.9	-27.0
Low-skill	-0.8	-10.3	-27.9
Panel B: Residents			
Herfindhal Index			
High-skill	25.6	6.9	15.9
Low-skill	2.3	6.8	0.8

Notes: The simulation uses as initial point the estimated equilibrium in 2010.

translate into a large and comparable employment agglomeration. Second, in stark contrast, residents sprawl and the bulk of the effects come from high-skilled citizens. Thus estimates imply that lower commuting time allowed high-skilled citizens to agglomerate in workplace and live in blocks further away. Analyzing the effects of transport investments separately, results indicate that the subway extension (third column) provokes similar impacts for high- and low-skilled workers, while BRT has a greater impact for high-skilled. Besides, it is import to highlight that investments exhibit strong complementarity since total effects are not the sum of partial effects.

In table 3.6, I investigate the impacts on the mean and dispersion of wages. In accordance with the increase in inequality, the employment agglomeration, induced by transport investments, reduce the mean and increase the dispersion of the wage premium. From the perspective of wages per residence location, wages across the city are on average higher and more similar. Once again, this is consistent with results previous described. Since investments lead to residents sprawl, mostly high-skill, the inequality in average wage per residence location diminishes. Finally, table 3.7 shows that high-skilled workers experienced a larger reduction in commuting times. Additionally, most of the effects is due to BRT or the combination of BRT and new subway stations.

All in all, results indicate that transport investments increased the welfare of high- and low-skilled workers. Nevertheless, high-skilled workers experienced a larger increase, raising inequality. The expansion of the transport infrastructure connected new locations with Rio's central business district,

Table 3.6: Contrafactuals: wage's mean and dispersion

	No BRT	No subway extension	No BRT No subway extension
Panel A: Wages per workplace			
Wage premium (high/low)			
Mean	5.5	3.9	9.0
Dispersion	-5.6	-3.9	-10.4
Panel B: Average wage per residence			
High-Skill			
Mean	-2.9	-3.8	-15.4
Dispersion	21.0	-3.6	15.4
Low-skill			
Mean	-2.3	-2.3	-11.4
Dispersion	17.3	5.3	18.3

Notes: The simulation uses as initial point the estimated equilibrium in 2010.

Table 3.7: Contrafactuals: average commuting time per residence location

	No BRT	No subway extension	No BRT No subway extension
Low-skill	14.2	-1.3	21.4
High-Skill	17.9	-0.9	27.3

Notes: The simulation uses as initial point the estimated equilibrium in 2010.

which increased the number of residential options with lower commuting costs. This led to a reduction in the concentration of residents and increased the concentration of jobs. Both effects are stronger for high-skilled workers, which raises residential and employment segregation. In particular, a gentrification process took place: higher demand and prices in the newly connected area led to an increase in residential segregation. This process was exacerbated by the fact that, among the newly connected areas, some had high amenities due to their proximity to the beach.

Final Remarks

This dissertation goal is to assess the impacts of a major transportation infrastructure expansion in Rio de Janeiro (Brazil). I construct a novel data set and employ machine learning, reduced-forms, and general equilibrium approaches to infer effects in commuting times, economic activity, welfare, inequality and segregation.

I have three main results. First, the new infrastructure successfully reduced commuting times. Besides, transportation investments show important complementarities, which point to the importance of connecting the entire network. Second, the agglomeration of jobs increased while the agglomeration of residents diminished. Both effects are larger for high-skilled workers. Third, BRT and subway stations had a positive impact in the level of economic activity in the vicinity of stations. The phenomenon is residence-led: sprawl and dispersion od residents induced more firms and jobs in these areas. The welfare of both types of workers increased, but high-skilled workers benefited the most, which resulted in higher inequality and segregation.

These results pose an important question: Is the development of cities inevitably associated with higher inequality and segregation? This is particular important for developing countries cities. If public authorities aim to increase welfare and reduce inequality, results show that general equilibrium effects and agglomeration externalities are relevant for transportation investments. In this regard, this dissertation develops new tools that can guide policymakers. Empirical exercises in all three chapters can be replicated for other cities in Brazil and, upon data availability, for other developing countries' metropolises.

Bibliography

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, 83(6):2127–2189.

Atack, J., Haines, M. R., and Margo, R. A. (2008). Railroads and the Rise of the Factory: Evidence for the United States, 1850-70. Working Paper 14410, National Bureau of Economic Research.

Baum-Snow, N. and Kahn, M. E. (2000). The effects of new public projects to expand urban rail transit. *Journal of Public Economics*, 77(2):241 – 263.

Behrens, K. and Robert-Nicoud, F. (2015). Chapter 4 - Agglomeration Theory with Heterogeneous Agents. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of regional and urban economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 171–245. Elsevier.

Billings, S. B. (2011). Estimating the value of a new transit option. *Regional Science and Urban Economics*, 41(6):525 – 536.

Black, D. A., Kolesnikova, N., and Taylor, L. J. (2014). Why do so few women work in New York (and so many in Minneapolis)? Labor supply of married women across US cities. *Journal of Urban Economics*, 79:59–71.

Bocarejo, J. P., Portilla, I., and Meléndez, D. (2016). Social fragmentation as a consequence of implementing a Bus Rapid Transit system in the city of Bogotá. *Urban Studies*, 53(8):1617–1634.

Bocarejo, J. P., Portilla, I. J., Velásquez, J. M., Cruz, M. N., Peña, A., and Oviedo, D. R. (2014). An innovative transit system and its impact on low income users: the case of the Metrocable in Medellín. *Journal of Transport Geography*, 39:49 – 61.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Combes, P.-P., Duranton, G., and Gobillon, L. (2012). The costs of agglomeration: Land prices in French cities.

Davis, M. A. and Ortalo-Magné, F. (2011). Household expenditures, wages, rents. *Review of Economic Dynamics*, 14(2):248–261.

Dorna G, R. (2017). Evaluación de impacto del proyecto de mejora integral del Ferrocarril Gral. Roca, Ramal Plaza Constituición. unpublished.

Duranton, G. and Puga, D. (2004). Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. In J. Vernon Henderson and Jacques-François Thisse, editor, *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2063–2117. Elsevier.

Epple, D., Gordon, B., and Sieg, H. (2010). A new approach to estimating the production function for housing. *American Economic Review*, 100(3):905–24.

Gershenson, S. (2013). The causal effect of commute time on labor supply: Evidence from a natural experiment involving substitute teachers. *Transportation Research Part A: Policy and Practice*, 54:127–140.

Gibbons, S. and Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, 57(1):148 – 169.

Gonzalez-Navarro, M. and Turner, M. A. (2018). Subways and urban growth: Evidence from earth. *Journal of Urban Economics*, 108:85 – 106.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598.

Gutiérrez-I-Puigarnau, E. and van Ommeren, J. N. (2010). Labour supply and commuting. *Journal of Urban Economics*, 68(1):82–89.

Kline, P. and Moretti, E. (2013). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *The Quarterly Journal of Economics*, 129(1):275–331.

Martinez, D., Mitnik, O., Oviedo, D., Scholl, L., and Yañez-Pagans, P. (2018). BRT impacts on housing values: The case of Lima, Peru. Inter-American Development Bank.

Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Pecora, A. R. and Menezes-Filho, N. (2014). O papel da oferta e da demanda por qualificação na evolução do diferencial de salários por nível educacional no Brasil. *Estudos Econômicos (São Paulo)*, 44(2):205–240.

Perdomo, J. A. (2011). A methodological proposal to estimate changes of residential property value: Case study developed in Bogota. *Applied Economics Letters*, 18(16):1577–1581.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9:21–58.

Rodríguez, D. A. and Mojica, C. H. (2009). Capitalization of BRT network expansions effects into prices of non-expansion areas. *Transportation Research Part A: Policy and Practice*, 43(5):560 – 571.

Scholl, L., Innao, M., Martinez, D., Mitnik, O., Oviedo, D., Pedraza, L., and Yañez-Pagans, P. (2018a). Employtment impacts of BRT systems: The case of Lima, Peru. unpublished.

Scholl, L., Mitnik, O. A., Oviedo, D., and Yañez-Pagans, P. (2018b). A Rapid Road to Employment? The Impacts of a Bus Rapid Transit System in Lima.

Silva, S. and Torres, H. (2013). Traffic and public transport planning for olympic games: Challenges for the city of rio de janeiro. unpublished.

Tsivanidis, N. (2018). The aggregate and distributional effects of urban transit infrastructure: Evidence from bogotá's transmilenio. Technical report, Working Paper, University of Chicago Booth School of Business.

Tyndall, J. (2017). Waiting for the R train: Public transportation and employment. *Urban studies*, 54(2):520–537.

Valentinyi, A. and Herrendorf, B. (2008). Measuring factor income shares at the sectoral level. *Review of Economic Dynamics*, 11(4):820–835.

Appendix

Table 8: Origin-destination survey: sample size per origin and destination municipality

	Destination						
Origin	Berford Roxo	Duque de Caxias	Guapimirim	Itaboraí	Itaguaí	Japeri	Magé
Berford Roxo	331	43	0	0	0	0	1
Duque de Caxias	44	670	5	0	0	2	18
Guapimirim	0	5	44	1	1	0	13
Itaboraí	0	0	0	165	0	0	2
Itaguaí	0	0	1	0	77	0	0
Japeri	0	1	0	0	0	63	0
Magé	1	17	13	2	0	0	92
Mangaratiba	0	0	0	0	6	0	1
Maricá	0	0	0	0	0	0	0
Mesquita	3	2	0	0	0	0	0
Nilópolis	6	2	0	0	0	1	0
Niterói	2	8	0	13	0	0	1
Nova Iguaçu	32	2	4	0	1	12	0
Paracambi	0	0	0	0	1	6	0
Queimados	0	0	0	0	0	12	0
Rio de Janeiro	80	187	7	5	4	34	17
São Gonçalo	1	1	1	18	0	0	1
São João de Meriti	18	32	0	0	0	0	0
Seropódica	0	1	0	0	4	0	0
Tanguá	1	0	0	6	0	0	1
Total	519	971	75	210	94	130	147

	Destination						
Origin	Mangaratiba	Maricá	Mesquita	Nilópolis	Niterói	Nova Iguaçu	Paracambi
Berford Roxo	0	0	4	9	1	35	0
Duque de Caxias	0	0	2	1	8	2	0
Guapimirim	0	0	0	0	1	0	0
Itaboraí	0	0	0	0	14	0	1
Itaguaí	6	0	0	0	0	1	1
Japeri	0	0	0	1	0	13	5
Magé	0	0	0	0	2	0	0
Mangaratiba	42	0	0	0	0	0	0
Maricá	0	123	0	0	17	0	0
Mesquita	0	0	82	14	0	25	0
Nilópolis	0	0	13	120	1	4	0
Niterói	0	15	1	1	739	3	0
Nova Iguaçu	0	0	21	6	2	511	3
Paracambi	0	0	0	0	0	5	90
Queimados	0	0	1	1	0	11	0
Rio de Janeiro	2	9	28	47	113	131	6
São Gonçalo	0	13	0	0	154	1	0
São João de Meriti	0	0	2	6	2	11	0
Seropódica	0	0	0	0	0	9	4
Tanguá	0	0	0	0	0	0	0
Total	50	160	154	206	1054	762	110

	Destination						
Origin	Queimados	Rio de Janeiro	São Gonçalo	São João de Meriti	Seropódica	Tanguá	Total
Berford Roxo	0	92	0	19	0	1	536
Duque de Caxias	0	203	1	32	0	0	988
Guapimirim	0	4	1	0	0	0	70
Itaboraí	0	6	17	1	0	5	211
Itaguaí	0	5	0	0	5	1	97
Japeri	12	36	0	0	0	0	131
Magé	0	14	1	0	0	1	143
Mangaratiba	0	0	0	1	0	0	50
Maricá	0	10	13	0	0	0	163
Mesquita	1	29	0	2	0	0	158
Nilópolis	0	53	0	7	0	0	207
Niterói	0	120	152	1	0	0	1056
Nova Iguaçu	11	157	1	8	10	0	781
Paracambi	0	10	0	0	4	0	116
Queimados	93	28	0	4	0	0	150
Rio de Janeiro	30	6499	83	83	7	0	7372
São Gonçalo	0	82	627	3	0	7	909
São João de Meriti	4	79	4	352	1	0	511
Seropódica	0	9	0	2	42	0	71
Tanguá	0	2	5	0	0	60	75
Total	151	7438	905	515	69	75	13795

Notes: Data is from 2011 Origin-Destination Survey, Rio de Janeiro State.

Table 9: Regression results: Firms, per sector of activity

			(1)	(2)	(3)	(4)	(5)	(6)
			industry	construction	commerce	services	public administration	agriculture
Announcement	BRT	up to 250 m	0.02***	0.02***	0.12***	0.16***	0.00*	0.00
		250m to 500m	0.01**	0.01***	0.03**	0.06***	0.00	-0.00
		500m to 750m	-0.00	0.00*	0.01*	0.03**	-0.00	0.00
		750m to 1000m	0.00*	0.00***	0.02**	0.03**	-0.00	-0.00
	SUBWAY	up to 250 m	0.05***	0.04**	0.26***	0.55***	-0.00	-0.00
		250m to 500m	0.02**	0.02***	0.12***	0.40***	0.00	0.00
		500m to 750m	0.02***	0.00	0.06**	0.21***	0.00*	-0.00
		750m to 1000m	0.01	0.00	0.04**	0.12**	-0.00	0.00**
Open	BRT	up to 250 m	0.00	0.00	0.01	0.03***	0.00	-0.00
		250m to 500m	0.00	0.00	0.00	0.03***	0.00***	-0.00
		500m to 750m	0.00*	0.00	0.00	0.01	0.00	0.00**
		750m to 1000m	-0.00	0.00	-0.00**	0.00	0.00	-0.00
	SUBWAY	up to 250 m	0.03**	0.04***	0.12**	0.36***	0.01*	0.00
		250m to 500m	0.01	-0.00	0.02	0.24***	-0.00	-0.00
		500m to 750m	-0.01	0.00	0.02	0.14***	-0.00*	-0.00
		750m to 1000m	-0.01**	0.00	0.02	0.05	0.00	-0.00
Observations			1,347,786	1,347,786	1,347,786	1,347,786	1,347,786	1,347,786
R^2			0.01	0.00	0.03	0.07	0.00	0.00
Number of grids			122,526	122,526	122,526	122,526	122,526	122,526
FE Year			Yes	Yes	Yes	Yes	Yes	Yes
FE Grid			Yes	Yes	Yes	Yes	Yes	Yes
Trend			RA	RA	RA	RA	RA	RA
Trend2			RA	RA	RA	RA	RA	RA
Cluster			RA	RA	RA	RA	RA	RA

Notes: The table reports estimated coefficients and standard erros in parentheses for specification (3). Each column refers to one dependent variable: log(number of jobs of workers with no high school) (1); log(number of jobs of workers with high school) (2); log(number of jobs of workers with college) (3). Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Table 10: Regression results: Firms, per size

			(1)	(2)	(3)	(4)	(5)
			0 employees	1 employee	2 to 10 employees	11 to 20 employees	> 20 employees
Announcement	BRT	up to 250 m	0.00*	0.08***	0.15***	0.06***	0.06***
		250m to 500m	0.00**	0.03***	0.04***	0.02**	0.02***
		500m to 750m	0.00	0.02	0.02**	0.00	0.01
		750m to 1000m	-0.00	0.02	0.03**	0.01***	0.01***
	SUBWAY	up to 250 m	0.00**	0.30***	0.49***	0.16***	0.16***
		250m to 500m	0.00	0.19***	0.32***	0.08***	0.11***
		500m to 750m	0.00	0.10**	0.16***	0.04***	0.04**
		750m to 1000m	0.00***	0.06**	0.10***	0.02**	0.02
Open	BRT	up to 250 m	-0.00	0.01***	0.01	0.01**	0.01***
		250m to 500m	-0.00***	0.02**	0.01**	0.00	0.00
		500m to 750m	0.00	0.01*	0.00	-0.00	-0.00
		750m to 1000m	0.00	0.00	-0.00	-0.00***	-0.00
	SUBWAY	up to 250 m	-0.00***	0.22***	0.31***	0.08***	0.07**
		250m to 500m	-0.00	0.15***	0.19***	0.00	0.01
		500m to 750m	-0.00**	0.08***	0.08	0.02	0.02
		750m to 1000m	0.00	0.04	0.01	0.01	0.01
Observations			1,347,786	1,347,786	1,347,786	1,347,786	1,347,786
R^2			0.00	0.03	0.05	0.02	0.02
Number of grids			122,526	122,526	122,526	122,526	122,526
FE Year			Yes	Yes	Yes	Yes	Yes
FE Grid			Yes	Yes	Yes	Yes	Yes
Trend			RA	RA	RA	RA	RA
Trend2			RA	RA	RA	RA	RA
Cluster			RA	RA	RA	RA	RA

Robust standard errors in parentheses

Notes: The table reports estimated coefficients and standard erros in parentheses for specification (3). Each column refers to one dependent variable: log(number of firms with 0 employees) (1); log(number of firms with 1 employee) (2); log(number of firms with 2 to 10 employees) (3); log(number of firms with 11 to 20 employees) (4); log(number of firms with more than 20 employees) (5). Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Table 11: Regression results: Jobs, per subsamples

			(1)	(2)	(3)	(4)
			0 employees	1st quartile	2nd quartile	3rd quartile
Announcement	BRT	up to 250 m	**0.66*** **	**0.50*** **	0.37*	**0.53** **
		250m to 500m	**0.32*** **	0.25*	0.13	0.19
		500m to 750m	**0.16*** **	**0.35** **	0.16	-0.01
		750m to 1000m	**0.14** **	**0.20*** **	0.24	0.32
	SUBWAY	up to 250 m	**1.73*** **	0.32	0.76*	**1.09*** **
		250m to 500m	**1.27*** **	**0.53** **	**0.84** **	**0.77*** **
		500m to 750m	**0.78*** **	0.24	0.23	0.24
		750m to 1000m	**0.46*** **	0.38	0.37*	0.30
Open	BRT	up to 250 m	**0.13*** **	0.07	0.09	0.10
		250m to 500m	**0.05** **	0.13	0.13	0.06
		500m to 750m	0.02	-0.02	0.03	0.05
		750m to 1000m	-0.02	0.05	0.00	**0.23** **
	SUBWAY	up to 250 m	**0.50** **	**1.11** **	**0.55** **	0.08
		250m to 500m	**0.48*** **	**1.01** **	0.01	0.06
		500m to 750m	**0.30*** **	0.18	0.38	-0.21
		750m to 1000m	0.00	0.04	0.01	0.15
Observations			1,271,325	19,129	19,129	19,140
R^2			0.18	0.06	0.23	0.42
Number of grids			115,575	1,739	1,739	1,740
FE Year			Yes	Yes	Yes	Yes
FE Grid			Yes	Yes	Yes	Yes
Trend			RA	RA	RA	RA
Trend2			RA	RA	RA	RA
Cluster			RA	RA	RA	RA

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The table reports estimated coefficients and standard erros in parentheses for specification (3). Each column refers to one subsample according with the number of jobs in the grid in 2006: 0 jobs (1); first quartile of the distribution of grids with positive employment (2); second quartile of the distribution of grids with positive employment (3); third quartile of the distribution of grids with positive employment (4); forth quartile of the distribution of grids with positive employment (4). Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Table 12: Regression results: Jobs, per educational level

			(1)	(2)	(3)
			no high school	high school	college
Announcement	BRT	up to 250 m	**0.35*** **	**0.40*** **	**0.20*** **
		250m to 500m	**0.13*** **	**0.16*** **	**0.06*** **
		500m to 750m	**0.06** **	**0.08*** **	**0.03** **
		750m to 1000m	**0.07*** **	**0.08*** **	**0.03*** **
	SUBWAY	up to 250 m	**0.87*** **	**0.87*** **	**0.53*** **
		250m to 500m	**0.68*** **	**0.62*** **	**0.36*** **
		500m to 750m	**0.37*** **	**0.34*** **	**0.18*** **
		750m to 1000m	**0.21*** **	**0.20*** **	**0.09** **
Open	BRT	up to 250 m	**0.05* **	**0.09*** **	**0.05*** **
		250m to 500m	0.01	0.04***	**0.02*** **
		500m to 750m	0.01	0.01	0.00
		750m to 1000m	-0.02*	-0.00	-0.00
	SUBWAY	up to 250 m	**0.43*** **	**0.43*** **	**0.30*** **
		250m to 500m	**0.31*** **	**0.20** **	0.09*
		500m to 750m	**0.20*** **	**0.17** **	0.06
		750m to 1000m	0.05	0.04	0.00
Observations			1,347,786	1,347,786	1,347,786
R^2			0.06	0.07	0.04
Number of grids			122,526	122,526	122,526
FE Year			Yes	Yes	Yes
FE Grid			Yes	Yes	Yes
Trend			RA	RA	RA
Trend2			RA	RA	RA
Cluster			RA	RA	RA

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The table reports estimated coefficients and standard erros in parentheses for specification (3). Each column refers to one dependent variable: log(number of jobs of workers with no high school) (1); log(number of jobs of workers with high school) (2); log(number of jobs of workers with college) (3). Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Table 13: Regression results: Average wage, per educational level

			(1)	(2)	(3)
			no high school	high school	college
Announcement	BRT	up to 250 m	**0.98*** **	**1.02*** **	**0.90*** **
		250m to 500m	**0.51*** **	**0.53*** **	**0.39*** **
		500m to 750m	**0.28*** **	**0.29*** **	**0.19*** **
		750m to 1000m	**0.25** **	**0.25** **	**0.17** **
	SUBWAY	up to 250 m	**2.03*** **	**2.07*** **	**2.07*** **
		250m to 500m	**1.69*** **	**1.71*** **	**1.64*** **
		500m to 750m	**1.09*** **	**1.12*** **	**1.03*** **
		750m to 1000m	**0.73*** **	**0.74*** **	**0.61*** **
Open	BRT	up to 250 m	**0.27*** **	**0.31*** **	**0.23*** **
		250m to 500m	**0.15*** **	**0.17*** **	**0.12*** **
		500m to 750m	**0.08** **	0.07*	0.03
		750m to 1000m	-0.02	0.00	-0.02
	SUBWAY	up to 250 m	**0.83*** **	**0.87*** **	**0.95*** **
		250m to 500m	**0.97*** **	**0.93*** **	**0.72*** **
		500m to 750m	**0.53*** **	**0.53*** **	**0.40*** **
		750m to 1000m	0.11	0.11	0.03
Observations			1,327,357	1,322,157	1,263,385
R^2			0.11	0.12	0.10
Number of grids			122,525	122,510	122,352
FE Year			Yes	Yes	Yes
FE Grid			Yes	Yes	Yes
Trend			RA	RA	RA
Trend2			RA	RA	RA
Cluster			RA	RA	RA

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The table reports estimated coefficients and standard erros in parentheses for specification (3). Each column refers to one dependent variable: log(average wage of workers with no high school) (1); log(average wage of workers with high school) (2); log(average wage of workers with college) (3). Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Table 14: Regression results: main outcomes

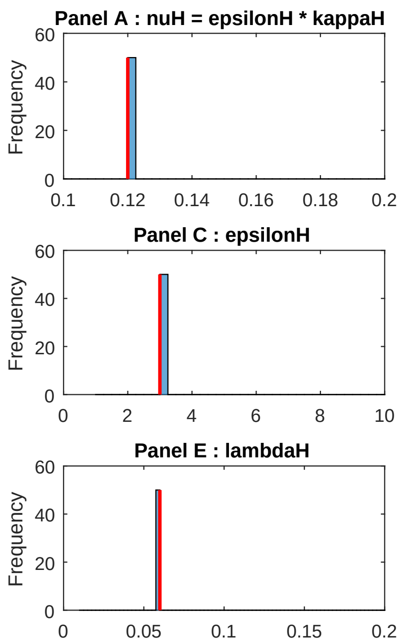
	log(jobs)		log(firms)		log(average_wage)	
	Announcement	Openning	Announcement	Openning	Announcement	Openning
Panel A: BRT						
up to 250 m	0.51***	0.10***	0.23***	0.04***	1.05***	0.30***
	(0.06)	(0.03)	(0.03)	(0.01)	(0.13)	(0.08)
250m to 500m	0.21***	0.05***	0.09***	0.03***	0.57***	0.19***
	(0.05)	(0.02)	(0.02)	(0.01)	(0.13)	(0.06)
500m to 750m	0.11***	0.02	0.05***	0.01	0.32***	0.09**
	(0.03)	(0.01)	(0.02)	(0.01)	(0.10)	(0.04)
750m to 1000m	0.11***	-0.01	0.05**	-0.00	0.27**	0.01
	(0.04)	(0.01)	(0.02)	(0.01)	(0.11)	(0.04)
1000m to 1250m	0.10**	-0.01	0.04**	-0.00	0.24**	-0.02
	(0.04)	(0.01)	(0.02)	(0.00)	(0.11)	(0.03)
1250m to 1500m	0.04	-0.00	0.02	0.00	0.14	0.01
	(0.03)	(0.01)	(0.01)	(0.01)	(0.09)	(0.03)
1500m to 1750m	0.05*	-0.02*	0.02	-0.01*	0.13	-0.04
	(0.03)	(0.01)	(0.01)	(0.01)	(0.08)	(0.04)
1750m to 2000m	0.04	0.00	0.02	0.00	0.11	0.01
	(0.03)	(0.01)	(0.01)	(0.01)	(0.08)	(0.03)
Panel B: Subway						
up to 250 m	1.13***	0.52***	0.64***	0.36***	2.07***	0.82***
	(0.28)	(0.15)	(0.16)	(0.09)	(0.45)	(0.24)
250m to 500m	0.87***	0.35***	0.45***	0.25***	1.73***	0.92***
	(0.22)	(0.10)	(0.12)	(0.06)	(0.38)	(0.19)
500m to 750m	0.50***	0.23***	0.25***	0.14***	1.14***	0.53***
	(0.14)	(0.07)	(0.07)	(0.04)	(0.28)	(0.10)
750m to 1000m	0.30***	0.06	0.15***	0.04	0.76***	0.11
	(0.09)	(0.07)	(0.05)	(0.04)	(0.19)	(0.11)
1000m to 1250m	0.22**	-0.06	0.12***	-0.03*	0.59***	-0.11
	(0.08)	(0.04)	(0.04)	(0.02)	(0.19)	(0.10)
1250m to 1500m	0.27***	-0.07	0.13**	-0.06***	0.64***	-0.09
	(0.10)	(0.05)	(0.05)	(0.02)	(0.21)	(0.11)
1500m to 1750m	0.15**	0.02	0.08**	0.01	0.37**	0.05
	(0.07)	(0.04)	(0.03)	(0.02)	(0.18)	(0.08)
1750m to 2000m	0.13**	-0.01	0.05**	-0.01	0.31**	-0.05
	(0.06)	(0.03)	(0.02)	(0.01)	(0.14)	(0.07)

Panel C: LRT						
up to 250 m	0.30	-0.12*	0.12	-0.06	0.05	0.05
	(0.21)	(0.07)	(0.12)	(0.04)	(0.30)	(0.13)
250m to 500m	-0.03	-0.27***	-0.08	-0.11***	-0.11	-0.28
	(0.25)	(0.06)	(0.12)	(0.03)	(0.52)	(0.17)
500m to 750m	0.03	-0.15***	-0.05	-0.06***	-0.13	-0.11
	(0.22)	(0.04)	(0.10)	(0.02)	(0.33)	(0.15)
750m to 1000m	-0.15	-0.12	-0.11	-0.01	-0.24	-0.11
	(0.41)	(0.07)	(0.20)	(0.03)	(0.68)	(0.16)
1000m to 1250m	0.05	-0.05	-0.04	-0.03	0.19	-0.03
	(0.41)	(0.09)	(0.20)	(0.03)	(0.73)	(0.14)
1250m to 1500m	0.17	-0.04	-0.03	0.07***	0.22	0.21
	(0.46)	(0.06)	(0.22)	(0.02)	(0.63)	(0.18)
1500m to 1750m	0.25	0.07	0.08	0.02	0.62	0.19
	(0.24)	(0.08)	(0.10)	(0.04)	(0.43)	(0.20)
1750m to 2000m	0.13	-0.02	0.01	-0.01	0.28	0.11
	(0.23)	(0.06)	(0.09)	(0.03)	(0.34)	(0.11)
Observations	1,347,786		1,347,786		1,347,038	
R-squared	0.08		0.08		0.12	
Number of id_grid	122,526		122,526		122,526	
FE Year	Yes		Yes		Yes	
FE Grid	Yes		Yes		Yes	
Trend	RA		RA		RA	
Trend2	RA		RA		RA	
Cluster	RA		RA		RA	
Pre-announcement '	No		No		No	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The table reports estimated coefficients and standard erros in parentheses for specification (3). Each column refers to one dependent variable: log(number of jobs) (1); log(number of firmsl) (2); log(average wage) (3). Besides treatment variables, covariates include year and grid fixed effects, linear and quadratic trends per district. Erros are clustered by districts. Dots represent point estimates and bar represents 95% confidence interval.

Figure 2: Monte Carlo Results for GMM Estimation Procedure



Notes: Results based on 100 simulations of hypothetical cities with the same size as Rio de Janeiro Metropolitan Area. I draw random productivity vectors and choose structural parameters to simulate the equilibrium based on the theorical model. With the same set of information that is observed in real data, I estimate structural parameters following the GMM procedure.