UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Rubens Lopes Pereira da Silva

Three Essays on the Relevance of Transport Accessibility Infrastructure:
Labor Market Size, Covid-19 Contagion, and School Absences

Recife
2024

Rubens Lopes Pereira da Silva

Three Essays on the Relevance of Transport Accessibility Infrastructure:
Labor Market Size, Covid-19 Contagion, and School Absences

Tese submetida ao programa de Pós-graduação em Economia da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de **Doutor em Economia**.

Área de Concentração: Teoria Econômica

Orientador: Prof. Dr. Raul da Mota Silveira Neto

Recife

2024

Catalogação na Fonte
Bibliotecária Maria Betânia de Santana da Silva, CRB4-1747

S586t Silva, Rubens Lopes Pereira da
Three Essays on the Relevance of Transport Accessibility Infrastructure: Labor Market Size, Covid-19 Contagion, and School Absences / Rubens Lopes Pereira da Silva. - 2024.
90 folhas: il., 30 cm.

Orientador: Prof. Dr. Raul da Mota Silveira Neto.
.
Tese (Doutorado em Economia) – Universidade Federal de Pernambuco, CCSA, 2024.
Inclui referências.

1. Rodovias. 2. Migração pendular. 3. COVID-19 (Doença). 4. Frequência escolar. I. Silveira Neto, Raul da Mota (Orientador). II. Título.

330.91732 CDD (22. ed.) UFPE (CSA 2024 – 050)

RUBENS LOPES PEREIRA DA SILVA

THREE ESSAYS ON THE RELEVANCE OF TRANSPORT ACCESSIBILITY INFRASTRUCTURE: Labor Market Size, Covid-19 Contagion, and School Absences

Tese apresentada ao Programa de Pós-Graduação em 25 de agosto de 2023 da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Economia.

Aprovado em: 25/08/2023.

BANCA EXAMINADORA

____
Profº. Dr. Raul da Mota Silveira Neto (Orientador)
Universidade Federal de Pernambuco

____
Profº. Dr. André Matos Magalhães (Examinador Interno)
Universidade Federal de Pernambuco

____
Profº. Dr. Diego Firmino Costa da Silva (Examinador Interno)
Universidade Federal de Pernambuco

____
Profª. Dr. Julia Rocha Araujo (Examinadora Externa)
Universidade Federal do Rio Grande do Norte

____
Profª. Dr. Isabel Pessoa de Arruda Raposo (Examinadora Externa)
Fundação Joaquim Nabuco

AGRADECIMENTOS

Primeiramente, sou grato a Deus, a razão de tudo o que sou e tenho. Agradeço a Ele por me conceder a vida, saúde e sabedoria para concluir este trabalho. Agradeço também ao meu orientador, Prof. Dr. Raul Silveira Neto, cuja orientação e sabedoria foram inestimáveis. Sem a sua assistência e paciência, este trabalho não teria sido possível. À minha amada esposa e à minha família, por sua compreensão, amor e apoio incansável durante os períodos mais difíceis desta jornada. Seu encorajamento constante foi a força que me impulsionou a seguir em frente. Agradeço também aos colegas pela amizade e à universidade pelo ambiente propício. Agradeço especialmente aos amigos Dr. Ricardo Fonseca e Dr. Bruno Ogava, pelas discussões e sugestões que contribuíram para o desenvolvimento deste trabalho, além da amizade e companheirismo. Obrigado a todos que contribuíram, direta ou indiretamente, para a realização desta pesquisa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

ABSTRACT

This thesis investigate the relationship between transport accessibility and several key societal phenomena in Brazil. The three chapters offer unique insights, employing rigorous analytical tools, geographical data manipulation, and rich datasets to discover previously unexplored causal relationships. The first paper delves into the influence of highways on long-distance commuting, underscoring the role of transport infrastructure in providing access to broader labor markets. The second paper illuminates the impact of highways on COVID-19 mortality rates, highlighting the essential role of transportation networks in disease spread and control. Lastly, the third paper presents an exploration into the role of commuting distances on school absenteeism, a critical factor in shaping future outcomes, especially for low-income residents in peripheral areas. In essence, the thesis delineates the nuanced and substantial effects that transportation infrastructure has on the labor market, public health, and education.

Chapter 1. Highways and Travel to Work: Evidence from Brazil

A recent report from the National Confederation of Transport (CNT, 2017) stated that investments in Brazil's transportation infrastructure have stagnated at deficient levels since the 1970s. This is alarming, as transportation solutions are essential for promoting growth, especially for a continental country like Brazil. Among studies that consider the determinants of city growth for Brazil, few treat the impact of transportation infrastructure specifically. Using Census data and regression analysis with instrumental variable, we investigate how proximity to a highway causally affects urban employment growth and travel to work from 2000 to 2010. We adopted a distinct identification strategy based on the construction of Brasília, the capital city of Brazil, which influenced the configuration of the national highway system. We found a positive causal relationship between proximity to highways and daily commuting to a different city, which implies that transportation infrastructure can offer workers access to a broader labor market.

Chapter 2. The Impact of Highways on COVID-19 Mortality: A Case Study for Brazil

This study explores the causal relationship between the extent of highways and the COVID-19 mortality rate in Brazil, shedding light on the role of transportation infrastructure in the spread of infectious diseases. We implemented an instrumental variables regression to investigate this relationship, utilizing maps of the Brazilian railway network from the 1960s as an instrumental variable, given the shared geographical challenges of constructing railways and highways. We discovered a positive effect, where cities with twice the average length of highways experienced a 14% higher COVID-19 mortality rate. This research emphasizes the need for considering broader societal factors like transportation networks in the design and implementation of disease control measures.

Chapter 3. Commuting and School Absenteeism: Evidence from Brazil

School absenteeism hinders the development of important non-cognitive skills and behaviors in students. Thus, understanding the causes of this problem is essential for effectively addressing the issue. Few studies have explored the impact of commuting on absenteeism in school, particularly for developing countries. Using a rich survey dataset, this study employs causal inference analysis to explore the impact of travel distance on student absenteeism on Recife, a densely populated city with significant transportation challenges particularly for low-income residents in peripheral areas. The instrumental variable strategy leverages the city government's enrollment policy, which plays a role in influencing school choice. The findings reveal a significant association between commute distance and absenteeism, indicating that a doubling of the distance leads to an 18.8% increase in absences. This research contributes to the literature by highlighting the role of school attendance in shaping future outcomes.

Keywords: Highways; Travel to work; COVID-19; School absenteeism.

RESUMO

Esta tese investiga a relação entre a acessibilidade dos transportes e diversos fenômenos sociais importantes no Brasil. Os três capítulos oferecem perspectivas únicas, empregando ferramentas analíticas rigorosas, manipulação de dados geográficos e conjuntos de dados ricos para descobrir relações causais anteriormente inexploradas. O primeiro artigo explora a influência das rodovias no deslocamento de longa distância, sublinhando o papel da infraestrutura de transporte em fornecer acesso a mercados de trabalho mais amplos. O segundo artigo analiza o impacto das rodovias nas taxas de mortalidade por COVID-19, destacando o papel essencial das redes de transporte na disseminação e controle de doenças. Por último, o terceiro artigo apresenta uma exploração sobre o papel das distâncias de deslocamento no absenteísmo escolar, um fator crítico em termos de ganhos futuros, especialmente para residentes de baixa renda em áreas periféricas. Em essência, a tese delineia os efeitos substanciais que a infraestrutura de transporte tem no mercado de trabalho, saúde pública e educação.

Capítulo 1. Rodovias e Migração Pendular: Evidências do Brasil

Um relatório recente da Confederação Nacional do Transporte (CNT, 2017) afirmou que os investimentos na infraestrutura de transporte do Brasil estagnaram em níveis deficientes desde a década de 1970. Isso é alarmante, pois as soluções de transporte são essenciais para promover o crescimento, especialmente para um país continental como o Brasil. Entre os estudos que consideram os determinantes do crescimento urbano para o Brasil, poucos tratam especificamente do impacto da infraestrutura de transporte. Utilizando dados do Censo e análise de regressão com variável instrumental, investigamos como a proximidade a uma rodovia afeta causalmente o crescimento do emprego urbano e a migração pendular de 2000 a 2010. Adotamos uma estratégia de identificação baseada na construção de Brasília, a capital do Brasil, que influenciou a configuração do sistema rodoviário nacional. Descobrimos uma relação causal positiva entre a proximidade às rodovias e o deslocamento diário para uma cidade diferente, o que implica que a infraestrutura de transporte pode oferecer aos trabalhadores acesso a um mercado de trabalho mais amplo.

Capítulo 2. O Impacto das Rodovias na Mortalidade por COVID-19: Um Estudo de Caso para o Brasil

Este estudo explora a relação causal entre a extensão das rodovias e a taxa de mortalidade por COVID-19 no Brasil, lançando luz sobre o papel da infraestrutura de transporte na disseminação de doenças infecciosas. Implementamos uma regressão de variáveis instrumentais para investigar essa relação, utilizando mapas da rede ferroviária brasileira dos anos 1960 como uma variável instrumental, dado os desafios geográficos compartilhados na construção de ferrovias e rodovias. Descobrimos um efeito positivo, onde cidades com o dobro do comprimento médio de rodovias experimentaram uma taxa de mortalidade por COVID-19 14% maior. Esta pesquisa enfatiza a necessidade de considerar fatores sociais mais amplos, como redes de transporte, no design e implementação de medidas de controle de doenças.

Capítulo 3. Commuting e Absenteísmo Escolar: Evidências do Brasil

O absenteísmo escolar prejudica o desenvolvimento de habilidades e comportamentos não cognitivos importantes nos alunos. Assim, entender as causas desse problema é essencial para abordá-lo eficazmente. Poucos estudos exploraram o impacto do commuting no absenteísmo, particularmente para países em desenvolvimento. Utilizando um rico conjunto de dados de pesquisa, este estudo emprega análise de inferência causal para explorar o impacto da distância de viagem no absenteísmo escolar em Recife, uma cidade densamente povoada com desafios significativos de transporte, especialmente para residentes de baixa renda em áreas periféricas. A estratégia de variável instrumental aproveita a política de matrícula do governo da cidade, que desempenha um papel na influência da escolha da escola. Os resultados revelam uma associação significativa entre a distância do deslocamento e a ausência, indicando que o dobrar da distância leva a um aumento de 18,8% nas ausências. Esta pesquisa contribui para a literatura ao destacar o papel da frequência escolar na formação de futuros resultados.

Palavras-chave: Rodovias; Migração pendular; COVID-19; Absenteísmo escolar.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 HIGHWAYS AND TRAVEL TO WORK: EVIDENCE FROM BRAZIL

1.1 INTRODUCTION

Daily commuting still is a central part of the day-to-day of many workers, even with the proliferation of teleworking opportunities (SANDOW, 2011). For many who live in rural areas and in the fringe of large centers, the possibility to commute long distances is a valuable alternative to out-migration (GREEN et al., 2019; HAM, 2002). As better transportation can offer workers access to a larger labor market (BJARNASON, 2014), it can be argued that the lack of investment in transportation infrastructure is an important obstacle to a better labor market matching. In fact, the literature has affirmed that better transportation network improves employment mainly in urban areas (ZENOU, 2000; COOMBES; CASADO-DÍAZ, 2005; BASTIAANSSEN; JOHNSON; LUCAS, 2022).

Much research has been done to show the interaction between long-distance commuting and workers' access to job opportunities for the developed world (SANDOW, 2008; SILVA; JOHNSON; WADE, 2011; BOUSSAUW; NEUTENS; WITLOX, 2012). Studies have found that workers are willing to face longer commutes to have better access to job opportunities, higher gains and cheaper rent. There are strong reasons, however, to believe that these relationships are even more pronounced in developing countries. In South Africa, for instance, cities' shape has been largely affected by policies of segregation from before and during the Apartheid era. Work opportunities are often located far from worker's homes, particularly for black South Africans. This created high commuting costs, both in terms of expenses and travel time. According to the 2003 National Travel Survey, black South Africans spent almost double the average commute time of an american worker in 2002. A decade later, the 2013 South African National Travel Survey reported that the average commuting time had increased by an additional 14 minutes for both black and white South Africans (KERR, 2017). In China, the past few decades were marked by fast urbanization and expansion of motorisation. Due to the rapid economic growth and large migration movements to China's metropolitan areas, the transport development gap has been widening between developed and underdeveloped areas. Investiment in transportation solutions in eastern China has been historically much higher than in the central and western regions. In addition, the income gap of Chinese residents resulted in a class-based stratification of housing choices in China, as high-income workers are more likely to accept longer commutes in order to fulfill their residential preferences (LI et al., 2019). In Latin America, after a period of import-substitution policies and intense population growth from the 1930s to the 1970s, the new economic liberalization and export-oriented developments in the 1980s improved access to new technology in the areas of telecommunication and transport (ROWE; BELL, 2018). With an inflow of new international investiments, large-scale and mainly natural resource-based enterprises started to appear distant from the main metropolitan centers at the time (United Nations, 2013). These developements provoked transformations that reshaped both the labor market and the infrastructure landscape of latin countries ever since.

Specifically for Brazil, long-distance commuting is an important subject for four reasons. First, Brazil is a large developing country known for the high density of its urban areas (INGRAM; CARROLL, 1981; MATA et al., 2007; FERNÁNDEZ-MALDONADO et al., 2014). The Portuguese colonization produced a country with isolated cities across a great territory, which later hindered the development of transportation infrastructures (NATAL, 1991). When compared to other countries (US, India, and China), Brazil has the largest percentage of inhabitants living in very small cities (100k to 250k people) (CHAUVIN et al., 2017). The insufficient supply of transport solutions, together with opportunities dispersed over a large territory, led to the development of a labor market significantly affected by spatial mismatch. Workers often need to travel long distances to reach their job. Poor workers are the most harmed by this configuration, as they face much longer commutes than richer workers (HADDAD; BARUFI, 2017). Second, in the 2000 decade the country saw a lot of growth of its agricultural and mining industries, mainly in the Central-West region, North region and in the state of Minas Gerais. These developments took place far from major cities, forcing workers to travel further afield. Third, the Brazilian transportation infrastructure is notably inadequate for its size and economy. Investments in transport have stagnated at exceptionally low levels since the 1970s (CNT, 2017b). While the highway network had an expansion of 96.7% of its extension in the 1960s, it only grew 11% from 2000 to 2010. Historically, the federal government have failed to execute around 30% of the authorized road infrastructure annual budget, causing an investment deficit that is reflected in the accelerated process of wear and tear on the federal road network (CNT, 2017a). Fourth, long-distance commuting is a growing phenomemon in Brazil, despite the developement of information technology. Interestingly, the growth of people who travel to work or study in the 2000s was higher in percentage than the growth of the whole group of people who work or study. In fact, the number of workers who commute to a different city for working increased by 209% from 2000 to 2010 (from 7.4 million to 15.5 million) (MOURA; DELGADO; COSTA, 2013), while the total number of workers grew 135% in the same period.

This research investigates if the availability of transportation infrastructure has a causal impact on inter-urban access to employment. Specifically, we use instrumental variables regression to evaluate if being near a highway had a causal effect on the growth of workers who traveled daily between cities for working from 2000 to 2010 in Brazil. We adopt a distinct identification strategy, leveraging an inconsequential units approach, similar to Morten e Oliveira (2018). This methodology capitalizes on a historical event — the construction of Brazil's national highway system during the 1950s — which was designed to connect all state capitals to the newly built capital, Brasília. We constructed a minimum spanning tree network, mirroring the radial layout of these highways. We use this artificial network as an instrumental variable in our model. To reinforce the validity of our instrument, we also include several important control variables. Our dependent variable is outbound commuting and we also investigate heterogeneous results by sex and education levels.

Among studies that consider the determinants of long-distance commuting (SANDOW, 2008; SILVA; JOHNSON; WADE, 2011; BOUSSAUW; NEUTENS; WITLOX, 2012; CASSEL et al., 2013; BERGANTINO; MADIO, 2016; CHEN; VOIGT; FU, 2021), we find several gaps that we aim to fill with the present study. First, despite the existence of a large literature that investigates the impact of transportation on employment and wealth (HOLL, 2004b; HOLL, 2004a; BANERJEE; DUFLO; QIAN, 2012; DURANTON; TURNER, 2012; REDDING; TURNER, 2014; FABER, 2014; GIBBONS et al., 2019), no previous work treats the specific impact of transportation infrastructure on travel to work[1] patterns using causal inference. Second, there is little evidence for the role of transportation infrastructure in lessening spatial mismatch in developing countries. Third, the recent studies that examine the spatial mismatch hypothesis for Brazil only consider intra-urban access to the labor market, and not inter-urban access (HADDAD; BARUFI, 2017; DUARTE, 2020). Our study also contributes to expanding a growing literature that seeks to understand the historical process of metropolization and urban expansion in Brazil paying close attention to workers commuting patterns (MOURA; BRANCO; FIRKOWSKI, 2005; OJIMA; JR; PEREIRA, 2010; MOURA; DELGADO; COSTA, 2013).

We found that, on average, Brazillian cities that had access to a highway saw their outbound commuting grow 49.6% over ten years more than their counterparts. When we analyze specific groups, we obtain statistical significance for men but not for women, indicating that women

[1] In this study, we use the term *travel to work* to refer to the daily commute of workers who work in a city different than the one where they live. In the literature, this movement is also known as *long-distance commuting* (ÖHMAN; LINDGREN, 2003), *pendulum migration* (ZASLAVSKY; GOULART, 2017), or *daily migration* (SILVA; FREITAS, 2020).

have more restrictions to accept distant jobs than men. Access to highways had no significant impact on workers with no education attainment, while it was most significant for workers with at least primary education. These results provide strong evidence that Brazilian workers – mainly skilled men – take advantage of highways to find better job opportunities in adjacent employment centers.

By studying the history of Brazil, we note that the combination of poor transportation infrastructure and low-income population produced, over the years, a pattern of cities with very high density and remarkably high commuting costs (INGRAM; CARROLL, 1981; FERNÁNDEZ-MALDONADO et al., 2014; NETO; DUARTE; PÁEZ, 2015). Better access to transportation infrastructure, thus, can be a solution to counterbalance the hardships imposed by such a spatially concentrated job market. Moreover, when transport infrastructures are underdeveloped, as in the Brazillian case (CNT, 2017b), even small expansions should promote big gains in terms of job allocation and employment. Given the scarcity of public resources, investing in transport is also a matter of better resource allocation for promoting equality and growth.

This paper is organized as follows. In the second section, we present a review of the literature; in the third section we give the historical context of transport in Brazil; in the fourth section we present our empirical strategy and data; in the fifth section we present the results; and the last section brings our conclusions.

1.2 LITERATURE REVIEW

Since the development of the intra-urban model by Alonso (1964), Mills (1967), and Muth (1969), many authors expanded the standard theoretical framework to take commuting into account. Wheaton (2004) proposed a model that ties commuting and congestion to agglomeration forces. When employment is more centralized, the model predicts long commute distances and high congestion levels. A city with a dispersed job market, on the other hand, should lead to lower marginal commuting costs and shorter travel distances. Vandyck e Proost (2012), via a theoretical model, examine how transport investment policies can encourage or discourage commuting flow between regions. Bergantino e Madio (2016) proposed a simple theoretical model showing how wages impact commuting costs and time. They argue that psychological barriers may lead workers to commute longer distances instead of migrating.

Empirically, many studies examined commuting patterns of workers, exploring how far they are willing to travel and for what reasons. Sandow (2008) used logistic regression with data for northern Sweden to study commuting behavior in low-density areas. He found a positive relationship of longer commuting with living in low-density regions, being male, and being younger. For Australia, Silva, Johnson e Wade (2011) analysed commuting patterns and found that the difference in income and education levels between origin and destination of commuters were determinant for the number of workers who traveled to work. Boussauw, Neutens e Witlox (2012) assessed the relationship between spatial aspects with commuting distance for Flanders, Belgium, using spatial regression models. They found that high residential density, high degree of spatial diversity, and elevated level of job accessibility were all related to short commute distances. Bergantino e Madio (2016) studied inter-regional mobility in the UK. They used multinomial logit to examine if monetary incentives affect the likelihood of moving to other regions and found a positive relationship that is larger for males than for females workers. Chen, Voigt e Fu (2021) explored inter-city commuting data for Germany between 1994 and 2018 using Machine Learning algorithms. They found that income is a determinant factor for travel to work: cities with a lower average income have more outbound commute towards higher income cities. Housing prices is also related with travel to work, as expensive apartments and higher rent prices are related to more inbound commuting from workers who live elsewhere. They also found that long-distance commute is more intense in the secondary and tertiary sectors, while farmers from the primary sector tend not to commute. Summarizing, authors argue that those who are willing to face longer commutes do it because they expect higher gains. They have access to more job opportunities (for those who live in low-density areas), receive a better salary (for the more skilled), and pay less rent (as they avoid moving to an high-rent metropolis).

For Brazil, specifically, Moura, Branco e Firkowski (2005) produced a rich review of the concept of pendulum migration and its applications in the literature, along with a survey of earlier empirical studies in the subject. Brito e Souza (2005) studied the travel to work movement in the Metropolitan Region of Belo Horizonte, in the southeast of Brazil. Using data from 2002, they found an interesting fact: almost 70% of those who travel to work to the capital city of Belo Horizonte used to live in the capital. This suggests that the peripherization of the region was caused by pressures from the real estate market, a phenomenon also observed in other parts of Brazil in the 1970s (OJIMA; JR; PEREIRA, 2010). Moura, Delgado e Costa (2013) examined the travel to work movement in Brazil by proposing a typology to differentiate cities in terms of direction flow and size. In terms of flow, municipalities can be in one of three conditions: *evader*, when the predominant commute movement is outbound (the case of dormitory towns, for example), *receptor*, when the prevailing commute movement is inbound, or *bidirectional*, when both inbound and outbound daily commute is approximately equivalent. In terms of flow magnitude, they characterize cities as large, medium, or small. Next, they use Census data from 2010 to classify Brazilian municipalities according to these criteria. They found that aproximately 51% of cities fall in the "small-evader" category[2]. Silva e Freitas (2020) examined the characteristics of those who travel to work in the southeast of Brazil using Logit models applied to Census data of 2010. They found that men, urban dwellers, and high-income workers are more prone to live and work in different cities. Meanwhile, being white and having more children are factors that reduce the chance of daily migration.

An important source of heterogeneity in commuting distances is gender. Empirical evidence suggests that women have more restrictions to accept longer-hours work due to household and childcare responsibilities, which is known as the "household responsibility hypothesis" (TURNER; NIEMEIER, 1997). The studies reviewed above in this section largely support this hypothesis for developed regions (MCQUAID; GREIG; ADAMS, 2001; SHEARMUR, 2006; SANDOW, 2008; CASSEL et al., 2013; BERGANTINO; MADIO, 2016), but there is also important evidence for developing countries. Neto, Duarte e Páez (2015) examined data for the Metropolitan Region of São Paulo using ordinal probit models. They inferred that being a married female implied an increase of about 7% on the probability of having shorter commuting times. Additionally, they found that single or divorced women have a shorter commute than their male counterparts. Olivieri e Fageda (2021) also found support for the household responsibility hypothesis in Montevideo, Uruguay. Through multilevel regression models, they obtained evidence that, on average, women make fewer trips and travel shorter distances. In the presence of children, the female breadwinner reduces their travel distance while the male breadwinner increases it. Also, women make more use of transit, thus residential areas with less access to public modes will negatively affect women's mobility.

The long-distance commuting also appears associated with education level. Silva, Johnson e Wade (2011) found that higher skills are a dominant factor prompting long-distance commuting. Cassel et al. (2013) found that more skilled people in Dalarna (Sweden) were more willing to travel for a longer duration. Similarly, Silva e Freitas (2020) found that lower levels of education is associated with less daily migration in Brazillian cities. These findings suggest that more skilled people benefit more from a larger job market, both in developed and developing countries. This is an important result for Brazil. As the country ranks among the nations with the highest levels of inequality (BEGHIN, 2008), it is expected that infrastructure will affect

[2] This number is based on the subset of cities where the daily flow is at least a thousand people.

people differently according to their qualification. As an example, access to transportation can help a computer scientist from a small town to find a much better-paying job in a bigger city. If this worker lives in a big city, he can also travel daily to another big city without having to incur migration costs. However, for a low-skilled worker, like a plumber, the prospect of gains is much smaller. The costs of commuting probably will not compensate the benefits. The worker needs to have a minimum qualification to take advantage of the long-distance commuting.

1.3 TRAVEL TO WORK AND TRANSPORT INFRASTRUCTURE IN BRAZIL

1.3.1 Evolution of land transportation in Brazil

Natal (1991) wrote that there is often a relationship between the occupation of a country and its transportation infrastructure. Although historical facts are not completely responsible for the current conditions of the country, understanding its initial circumstances can shed some light on the state of a country's transport system. In this regard, Brazil's occupation process is an interesting case.

For early colonizers, a common occupation strategy was to first settle on the coast and then move to the interior regions of the country. That was the case, for instance, in the colonization process of North America. After arriving in Brazil in 1500, the Portuguese logic, however, was to rapidly penetrate the interior of the country to take profit from its abundant natural resources (NATAL, 1991). Throughout the colonial period, Brazilian spatial occupation was based on the exploration of primary exportation products (basically: sugar cane, gold, and coffee). This approach generated Brazilian first important cities, but also produced "economic islands": highly specialized cities isolated and distant from each other (HORTENCIO; MILANI, 2019). In the first centuries, there were no economic connections between Brazilian cities or regions as the transport infrastructure was poorly developed (NATAL, 1991).

With the Industrial Revolution came a demand for faster transportation in western countries, which in the 19th century was solved by the construction of railways. The dispersed occupation of Brazil, however, imposed a challenge, as it was difficult and expensive to build an integrated system throughout its great territory (NATAL, 1991). Despite the large system of navigable waterways concentrated in Brazil's north and west, they could not substitute railway development as most production and commerce was located in regions that required land transportation (SUMMERHILL, 2005). Natal (1991) writes that the answer was to build a radial railway system connecting the main cities in the interior to ports on the coast. The railroads in this system were rarely interconnected, which was not a problem initially as the Brazilian economy was export-oriented. However, as the internal markets started to develop in the later 19th century, the limited transportation structure turned out to be insufficient due to the lack of interregional connections.

It was in this context that highways started to appear, often near important railways (NAZARETH, 1978). Paved roads were a cheaper way to connect different regions of the country and supply transportation from the train stations to cities nearby. Despite that, highway construction developed slowly until the mid-20th century, when it became the most important mode of transportation in Brazil (NATAL, 1991). The inflection point occurred when the Brazilian capital was transferred to a new city built from the ground up in the center of the country: Brasília. As a result, a new highway system was constructed to connect the new capital with all the other capitals of the country (MORTEN; OLIVEIRA, 2018).

According to Nazareth (1978), because of increased investments in road systems and lack of maintenance of railways and navigation infrastructure, the 1960s ended with 95% of passengers and 73% of goods being transported by roads. The 1960s started a period of major changes in the population geography of Brazil. The urban centers of the country were going through an intense process of industrialization, and many Brazilians migrated from rural regions to urban cities looking for better working conditions. In the second half of the 20th century, the urban population of Brazil increased 7.3 times – from 19 million to 138 million. This movement was strikingly fast compared to capitalist countries (European countries, US, and Japan) (BRITO; SOUZA, 2005). According to Heidrich (2004), this rapid expansion of the urban landscape was possible, among other reasons, by the declining costs of transportation.

According to Ojima, Jr e Pereira (2010), it was around that time that many dormitory towns emerged in the country. Also known as *satellite cities*, these towns are characterized by large residential areas in which many of their dwellers work in another city, usually in the same metropolitan region. Although this phenomenon is also observable in other countries, ie. US and England, the Brazilian case is special. While American and British satellite cities originated from middle and high-income white workers looking to live far away from polluted industrial urban centers, the Brazilian dormitory town was a result of great real-estate speculation that repelled poorer workers to the fringes of employment centers (OJIMA; JR; PEREIRA, 2010).

Consequently, the urban regions of Brazil expanded rapidly, much faster than the country's transport infrastructures could handle. While in the 1920s the US had $7,500 of GDP per capita (in 2012 dollars) and 50% of urbanization, Brazil only reached that same GDP per capita level in 2011, but with 80% of urbanization. Nowadays Brazil is more urbanized than the US, but much poorer (CHAUVIN et al., 2017).

The 1970s brought a world oil crisis that severely damaged the Brazilian economy. As transportation infrastructure had been mainly financed with public money, the Brazilian transportation system started to deteriorate (NATAL, 1991). Since then, investments in transportation infrastructure have stagnated at extremely low levels (CNT, 2017b). The federal government have failed to execute around 30% of the authorized road infrastructure annual budget, causing an investment deficit that is reflected in the accelerated process of wear and tear on the federal road network (CNT, 2017a). Table 1 shows the extension of national highways from 1960 to 2010 and its growth per decade.

Table 1 – Paved highways extension growth from 1960 to 2010 (Km)

Period	Extension	Relative to past decade
1960	8,675	-
1970	24,146	178.34%
1980	47,487	96.67%
1990	50,372	6.08%
2000	56,097	11.37%
2010	62,351	11.15%

Elaborated by the authors.

Data from the Ministry of Transport, Ports and Civil Aviation.

The combination of quick industrialization, poor transportation infrastructure and low-income population produced, over the years, a pattern of cities with very high density and remarkably high commuting costs (INGRAM; CARROLL, 1981; MATA et al., 2007; FERNÁNDEZ-MALDONADO et al., 2014).

1.3.2 Travel to work in Brazil

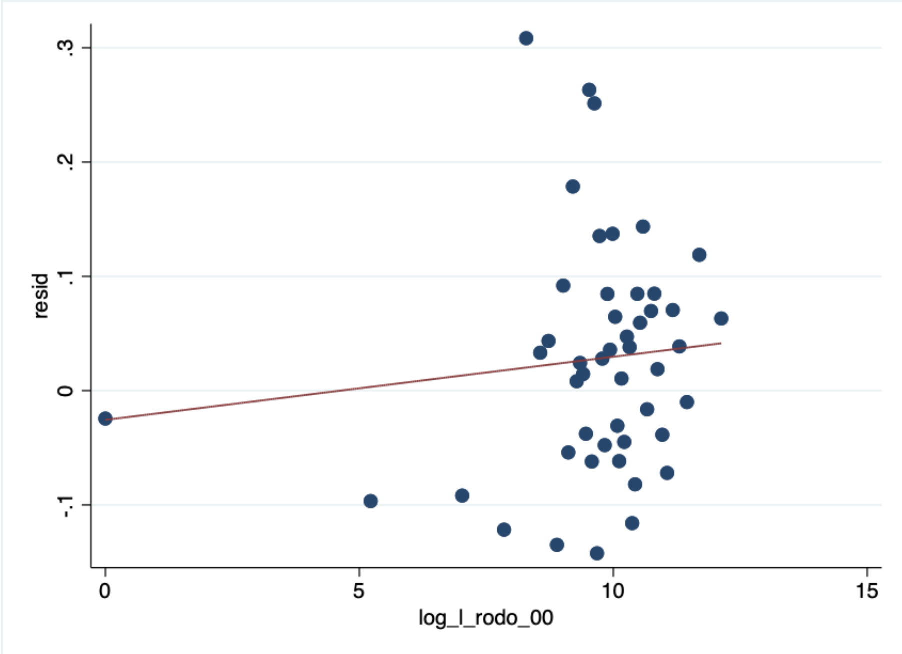
According to Moura, Delgado e Costa (2013), of the almost 110 million people who worked or studied in 2000, 6,7% performed this activity outside their municipality of residence. This percentage rose to 10,6% in 2010, and the number of people more than doubled: it went from 7,403,456 to 15,472,863 people. Of the 5.565 municipalities in Brazil in 2010, 56,7% had a daily inbound and outbound flow of at least a thousand people. The authors affirmed that the growth of people who traveled to work or study in the 2000s was even higher in percentage than the growth of the whole group of people who worked or studied.

The evolution of travel to work numbers is consistent with key policies that took place at the same period in Brazil. First, a government program provided some public services to remote areas of the country, such as public universities. This had a two-fold effect: many new jobs were created because of those universities, which imediately improved the lives of the city residents (CARAZZA, 2016); and students that lived in the city and in neighbouring towns now had an incentive to commute to the university in search of better education opportunities, which improved the country's education level especially for rural towns. The latter effect had a lingering impact in travel to work numbers even years later, as those better-skilled workers gained access to more work opportunities in more competitive centers outside their hometown after graduating (NETO; AZZONI, 2011).

Second, the 2000 decade saw an expansion of industrial activities outside of the traditional urban centers of the country. The North and Northeast regions of Brazil received both public and subsidized foreign investiments for the construction of refineries, automotive assembly plants, and port and shipping industry (GUMIERO, 2015; HISSA-TEIXEIRA, 2018; SILVA; BAGATTOLLI, 2020). As those investiments were mainly located in metropolitan areas, new commuting paths were emerged that stimulated more travel to work.

In spite of the influence of national policies, there is also evidence that the network of highways was a strong conductor of the travel to work advance in this period. Moura, Delgado e Costa (2013) stated that, in the 2000s, city connections grew from the typical core-periphery pattern to more complex arrangements in the last decades. This is true for the states that form the Brazilian South Region (Porto Alegre, Paraná, and Santa Catarina) that are interwoven by two important highways (the BR 116 and BR 101). This phenomenon is also observable in the Northeastern coast, where the metropolitan region of Recife and the urban agglomerations of Maceió, João Pessoa, and Natal are all linked by the national road network and have expressive numbers of daily migration. The authors noted that, in some cases, travel to work flows materialized following highway paths. To illustrate, we regressed our dependent variable, Travel to Work growth, against all of our control variables using OLS. We then plotted the residuals of that regression against the log of highway extension in 2000. Figure 1 shows a positive relationship between the two variables. From this evidence, we believe that the avaiability of a national highway system was crucial to the remarkable expansion of travel to work in the 2000s.

Figure 1 – OLS regression plot of Travel to Work growth regression residuals against log of length of highways in municipalities in 2000



1.4 EMPIRICAL STRATEGY AND DATA

1.4.1 Specification

In this subsection, we propose an empirical model that depicts how highways affect travel to work.

Let i be the index of cities and j the index of states. The cross-section model is:

$$y_i = \alpha + \beta \cdot \text{Highway access}_i + \Gamma \cdot \text{Controls}_i + \lambda_{s(i)} + \varepsilon_i \tag{1.1}$$

in which y_i is log difference of the total of people who traveled outbound to work in the 2000-2010 period. As our unit of study is the municipality, *outbound commuting* describes workers who travel away from their city of residence daily to reach their workplace. To build our explanatory variable, we borrow from Faber (2014)'s approach. Highway access$_i$ is a dummy variable that takes the value of one if any part of the municipality was within 10 km distance of a federal highway in 2000 (and zero otherwise). $Controls_i$ is a vector of eight variables. Three variables are related to geography and space: log distance to the nearest capital, log distance to the coast, and log distance to Brasília. They control for paths that already had intense levels of transportation before the analysis. Three other variables measure historical characteristics of the city: proximity indices to the nearest city related to Brazilian historical sugar, gold and coffee business cycles. Cities related to these periods needed a lot of transportation so they are historically better connected. We also include log of GDP per capita in 2000 and log of population in 2000 to account for previous levels of development. We also include log of GDP per capita in 2000 and log of population in 2000 to account for previous levels of development.

Every variable was measured at the municipality level, which is the lowest administrative level in Brazil. While this choice provided us with detailed data, it also imposed a problem: from 2000 to 2010, the number of municipalities went from 5,507 to 5,565. As many municipalities had their borders modified, using the original municipality configuration as the unit of study could lead to bias. Instead, we use MCAs (Minimum Comparable Areas), an aggregation of municipalities developed by IPEA (REIS et al., 2008). This aggregation strategy allows us to compare the same spatial units throughout time. Similar to Faber (2014), we exclude the MCAs close to the capital cities from the sample. In our preferred specification, we use the range of 50 km from the nearest capital. The final dataset is comprised of 5,162 MCAs.[3]

The model is estimated with state fixed effects $\lambda_{s(i)}$. The country has 26 states and one federal district. As we are working only with federal highways, the state fixed effects can help to partial out the effect caused by state highways; we also cluster errors at the microrregion level.

1.4.2 Endogeneity

The transportation literature has shown that naive OLS estimations can suffer from endogeneity problems (BAUM-SNOW; FERREIRA, 2015). While new infrastructure can promote job accessibility, it is also true that when decision makers expect a city to grow in employment, they may promote transport investments preemptively. This reverse causality problem could introduce bias to our estimates. In addition, our estimations are prone to suffer from omitted variable bias as there can be unobservable characteristics affecting both provision of infrastructure and access to employment centers.

To deal with these issues, several identification strategies have been proposed in the literature. Baum-Snow (2007) used 1947 plans of the Interstate Highway System as an instrument for the network that was built later, being the first work to use construction plans to instrument for realized infrastructure. Duranton e Turner (2012) use the same plans, along with railway

[3] In this study, we use municipality and MCA as interchangeable terms.

networks from the 19th century and expedition routes in past centuries as instruments for current highways' location in the USA. Garcia-López, Hémet e Viladecans-Marsal (2017) use Roman roads and 1870 railways as instruments for the Regional Express Rail (RER) infrastructure in Paris metropolitan area.

Besides planned routes and ancient roads, other works have used the inconsequential units approach (REDDING; TURNER, 2014). In this form of identification, straight lines connecting important cities serve as an instrument for highways or railways that were built between them. Smaller cities that conveniently lie in these paths are said to have received a road "accidentally", as if in a random assignment. Banerjee, Duflo e Qian (2012) used this method as an instrument for China's railway system. More recently, Minimum Spanning Tree (MST) algorithms have been used for generating these artificial networks. These algorithms connect a set of nodes with straight lines minimizing the total length of the segments, thus emulating the job of road planners: connect many cities with a single network minimizing the cost (FABER, 2014). Faber (2014) used this method while studying the impact of China's highway system, and Morten e Oliveira (2018) produced a Spanning Tree Network to predict the Brazilian highway network.

Following previous literature, our identification strategy will rely on a inconsequential units approach borrowed from Morten e Oliveira (2018). The authors built a minimum spanning tree network tailored to mimic the Brazilian government highway plans in 1950s. With the construction of the capital Brasília in a central region of the country, the government simultaneously started building a national highway system connecting the capital to all state capitals. Along with longitudinal, transversal and diagonal roads, they built eight radial highways starting in Brasilia and expanding to the rest of the country. Based on this historical data, the authors divided Brazil into eight segments and built a minimum spanning tree network that connected Brasília to all state capitals within each segment. Our MST network is shown in Figure 2.

Our first stage regression is:

$$\text{Highway access}_i = \sigma + \mu \cdot \text{Access MSTN}_i + \Gamma \cdot \text{Controls}_i + \lambda_{s(i)} + \eta_i \tag{1.2}$$

Specifically, our instrument Access MSTN_i is a dummy variable that takes the value of one if any part of the municipality was within 10 km distance of the artificial network.

The relevancy condition states that the instrument should be a good predictor of the explanatory variable. As discussed above, historical evidence suggests that the national highways system, from the 1950s, was mainly developed to connect capital cities of Brazil to Brasília as

Figure 2 – The Minimum Spanning Tree Network

straightforwardly as possible (MORTEN; OLIVEIRA, 2018). Thus, a minimal spanning tree that connects capital cities to Brasília with straight lines and using the smallest possible total extension should be a good proxy for the engineers' and social planners' strategy at the time. We are confident that our instrument satisfies the relevancy condition, as shown by the first stage F well above the threshold value of 10 (see Table 3). Also, correlation analysis shows that the number of kilometers of highways in each municipality becomes more correlated with the artificial network along the decades.[4] We also believe that our instrument satisfies the exogeneity condition. Due to the inconsequential nature of the generated lines, it is reasonable to affirm that the instrument cannot affect the travel to work levels via any other way except by predicting highway access.

However, as much as the inconsequential units approach is already consolidated in the literature, we take an additional step: we include control variables to ensure the exogeneity of our instrument. The log distance from the nearest capital accounts for the fact that cities near the capital are more likely to receive roads.[5] The log distance from Brasília, the country's capital city, is included because the network was intentionally generated to connect all capital cities to Brasília. The log distance to the coast controls "for access to traditional methods of transportation that existed before the lines of interest were constructed" (BANERJEE; DUFLO; QIAN, 2012).

[4] We ran correlation analysis for each decade from 1970 to 2010.
[5] As noted in Section 1.4.1, we also exclude the capital cities from our dataset.

Finally, we include three variables that measure proximity to important cities in the economic history of Brazil: key cities to the sugar, gold and coffee business cycles. Cities related to these periods needed a lot of transportation so they are historically better connected. These variables are indices: they take the value 1 if the municipality was directly affected by the respective economic cycle, and a value between 0 and 1 for nearby cities within 200 km based on distance according to a functional form.[6] We also include log of GDP per capita in 2000 and log of population in 2000 to account for previous levels of development.

1.4.3 Data

Our dataset was obtained from 3 different sources. The Brazilian Demographic Census is performed by the Brazilian Institute of Geography and Statistics (IBGE) every ten years. We use Census data from 2000 and 2010 to obtain our dependent variable at the municipality level for each decade.

The Ministry of Transport, Ports and Civil Aviation provides digital maps and pictures of the highways of Brazil for each decade from 1960 to 2010, and the IBGE provides digital maps for the Brazilian municipalities. From these data, we were able to compute the explanatory variable using GIS techniques. We also used GIS and Python libraries to produce the instrument, which is based on Morten e Oliveira (2018).

We also use a set of geographical and historical controls. The distance to the nearest capital and to the coast was computed by the authors. As historical controls, we used three indexes that measure the proximity of each municipality to the central locations of three critical moments of Brazilian economic history: the sugar cycle, the gold rush, and the coffee boom. These indices were produced by Naritomi, Soares e Assunção (2012).

In Table 2 we show some descriptive statistics for the variables used in the regression analysis. We notice that men's travel to work growth is higher then women, on average. This effect can be seen in the categories: "with children", "without children", and "single"; being "married" the only exception. Among women's categories of outbound commute, women without children shows the largest growth of the decade, while single women had the smallest growth. The same can be said about men. Also, the data suggest that married workers increased their

[6] The function is $I_i = \begin{cases} \left(\frac{200-d_i}{200}\right)^2 & \text{if } d_i \leq 200 \text{ km} \\ 0 & \text{otherwise} \end{cases}$. For more details, see Naritomi, Soares e Assunção (2012).

level of travel to work more than single workers, and people without children engaged more in long-distance commuting than workers with children.

Table 2 – Descriptive Statistics

	count	mean	std	min	median	max
Outbound growth	5480	0.985	0.773	-3.376	0.874	5.718
Outbound men growth	5480	1.078	0.844	-2.624	0.933	6.078
Outbound men with children growth	5480	1.075	0.967	-3.424	0.874	5.839
Outbound men no children growth	5480	1.632	1.142	-2.707	1.448	5.635
Outbound single men growth	5480	0.993	1.015	-3.217	0.840	5.876
Outbound married men growth	5480	1.205	0.933	-3.424	1.013	5.650
Outbound women growth	5480	0.813	0.980	-3.890	0.718	5.239
Outbound women with children growth	5480	1.060	1.242	-4.174	0.908	5.384
Outbound women no children growth	5480	1.370	1.291	-3.552	1.346	5.125
Outbound single women growth	5480	0.581	1.097	-4.005	0.510	5.078
Outbound married women growth	5480	1.264	1.178	-3.185	1.105	5.558
Log population 2000	5480	9.334	1.069	6.678	9.247	13.886
Log GDP per capita 2000	5480	10.439	1.343	7.338	10.210	16.424
Log nearest capital dist	5480	12.121	0.723	8.768	12.246	13.742
Log coast dist	5480	12.375	1.119	7.369	12.553	14.853
Log Brasília dist	5480	13.781	0.513	10.549	13.848	14.916
Sugar proximity index	5479	0.074	0.200	0.000	0.000	1.000
Coffee proximity index	5479	0.092	0.255	0.000	0.000	1.000
Gold proximity index	5479	0.123	0.281	0.000	0.000	1.000

1.5 RESULTS

1.5.1 Main results

Our main results are reported in the Panel A of Table 3. Our first conclusion is that highways had a statistically significant and causal effect on the workers commuting patterns. According to our estimates, when the average Brazilian municipality had access to a highway in 2000, many residents were encouraged to look for job opportunities in neighboring employment centers – a growth of 49.6% over ten years. As Table 2 reports an average travel to work growth of 98,5% along the 2000 decade, our results suggest that highways were responsible for aproximately half of this growth.

Panel B shows that the first stage regression. The statistics show that the instrument is significant and its first stage F is above 10, even when several control variables are taken into account.

To further test the strength of our instrument, we run the Kleibergen-Paap rk Wald statistic that measures weak instruments and we reject the null-hypothesis that the model is underidentified. We also report the p-value from the chi-squared Anderson-Rubin test, a test that evaluates the relevancy of an instrument in just-identified models (MOREIRA, 2009). We are able to reject the null-hypothesis of the A-R test at the 5% level in all of our models where the explanatory variable is significant. Given that we have a strong first stage, we can move on to explore the impact of highways on our variables of interest.

Our findings suggest that being connected to a highway favored the evading movement of workers towards those new employment centers. Better transportation infrastructure offer workers faster access to a larger radius, so when a city is connected to a highway, more workers can use that infrastructure to look for job opportunities in a larger labor market outside their city of residence (BJARNASON, 2014). As we exclude municipalities close to capital cities, we are able to capture the evasion-inducing effect of highways on the average Brazilian municipality.

This interesting result is aligned with past descriptive data from the literature. According to the typology proposed by Moura, Delgado e Costa (2013), more than 56% of Brazilian cities are of the evader type, while 31% are of the bidirectional type, and 12% are receptors. The authors explain that, between 2000 and 2010, job opportunities became more spread out in the country's territory. As some cities developed thriving employment centers, these places became attractive to workers from surrounding municipalities that had limited job offers. These workers started looking for jobs in these flourishing locations. As each new employment hub attracted workers from several nearby municipalities, many more cities became "evader" type.

Table 3 – Main model results

Panel A		
	OLS	Second Stage
	(1)	(2)
	Outbound growth	Outbound growth
Highway access	0.00565	0.496**
	(0.0288)	(0.240)
Controls	Yes	Yes
State FE	Yes	Yes
Observations	5162	5162
R^2	0.128	0.0621

Panel B	
	First stage
	(1)
	Highway access
MSTN access	0.165***
	(0.0250)
Controls	Yes
State FE	Yes
Observations	5162
R^2	0.142
First stage F	43.41
Kleibergen-Paap p-value	1.11e-08
Anderson-Rubin p-value	0.0399

Clustered standard errors in parentheses. The dependent variables are the Highway acess for the first stage and the log difference of the total of people who traveled to work in the 2000-2010 period in the second stage. Highway access is a dummy that takes the value of one if any part of the municipality was within 10 km distance of a federal highway in 2000. Control variables are log of population in 2000, log of GPD in 2000, log distance to the nearest capital city, log distance to the coast, log distance to Brasília, and indices indicating proximity to cities related to important economic cycles in Brazil (sugar, coffee and gold). * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

1.5.2 Heterogenous results

1.5.2.1 Gender

Empirical studies have long asserted that the pattern of commuting of men and women differ, and women generally have more restrictions to accept longer-hours work than men (MCQUAID; GREIG; ADAMS, 2001; SHEARMUR, 2006; SANDOW, 2008; CASSEL et al., 2013; BERGANTINO; MADIO, 2016). Past studies had already pointed out that women commute less than men in Latin America (NETO; DUARTE; PÁEZ, 2015; OLIVIERI; FAGEDA, 2021). This has been associated with different explanations. First, because of the aforementioned Household Responsibility Hypothesis (see Section 1.2), travel to work may impose a higher cost to women than for men. As women generally have more responsabilities at home, a better access to a highway will not benefit her if she is not able to take advantage of this improved access. The opportunity cost – leaving the house work behind and working far away from her children – may be too high for the average brazilian women. Second, there is evidence that women are more likely to work close to home. This is largely because female-dominated occupations are more evenly distributed in space when compared to male-dominated jobs (SANDOW; WESTIN, 2010). Third, the percentage of women who work in part-time jobs is larger than for men (IBGE, 2010a). If they work less they also receive less on average, which implies that for many women the potential earnings do not compensate the costs of travelling to work. Fourth, according to Neto e Moura (2019), in Brazil, long commute time is causally associated with chance of being a victim of violence in urban centers, especially for women. It is possible that this effect is even more pronunced in the context of travelling between cities, as in the highway the driver is more alone and more vulnerable than inside an urban area.

We explore this effect by running regression analysis for men and women separately. The new set of estimates is presented in the following Table 4. In our estimates, we found a causal impact of highway access on travel to work for men, but not for women. This result is consistent with the idea that reducing commuting costs may not be enough for promoting travel to work, as the worker can have other reasons to not leave their city of residence. In the case of women, this can be explained by the Household Responsibility Hypothesis, by the fact that female-dominated jobs are more evenly distributed in space, by the fact that women tend to work in part-time jobs, and because of risks associated with travelling alone. It is interesting to notice that there was no significant result for women regardless of marital status or the presence of children.

This result suggests that the Household Responsibility Hypothesis is not the only factor behind the lack of women's travel to work. Regarding men, the results are similar among columns. For men without children (column 3), the results suggest that highways were responsible for aproximately half of the average travel to work growth in the 2000 decade (compare to Table 2). The only result without significance was for single men (column 4). This subgroup of the male workers are likely younger, less skilled, and less experienced, thus they should find fewer opportunities outside of their hometown than their counterparts. Moreover, this group often is financially dependent on their parents, which may explain their lack of travel to work.

Although past studies have also found that women commute less than men (SANDOW, 2008; NETO; DUARTE; PÁEZ, 2015; BERGANTINO; MADIO, 2016; OLIVIERI; FAGEDA, 2021), our analysis showed a stronger result: for Brazil, between 2000 and 2010, better access to highway had no impact on female workers.

Table 4 – Results for Men and Women

Men	(1)	(2)	(3)	(4)	(5)
	All	With child	No child	Single	Married
Highway access	0.595**	0.688**	0.893**	0.331	0.754**
	(0.269)	(0.320)	(0.399)	(0.306)	(0.324)
Controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Observations	5162	5162	5162	5162	5162
First stage F	43.41	43.41	43.41	43.41	43.41
Kleibergen-Paap p-value	1.11e-08	1.11e-08	1.11e-08	1.11e-08	1.11e-08
Anderson-Rubin p-value	0.0245	0.0229	0.0151	0.287	0.0129
Women					
	(1)	(2)	(3)	(4)	(5)
	All	With child	No child	Single	Married
Highway access	0.249	-0.116	0.353	0.224	0.124
	(0.291)	(0.388)	(0.382)	(0.320)	(0.373)
Controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Observations	5162	5162	5162	5162	5162
First stage F	43.41	43.41	43.41	43.41	43.41
Kleibergen-Paap p-value	1.11e-08	1.11e-08	1.11e-08	1.11e-08	1.11e-08
Anderson-Rubin p-value	0.398	0.767	0.362	0.489	0.740

Clustered standard errors in parentheses.

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

For more details, see Table 3.

1.5.2.2 Education

In Table 5, we estimate our regression model by education levels. We expected to see some variation in the coefficients, as the level of education may dictate the worker's ability to arbitrate between modes of transport. Non-skilled workers, for instance, usually occupy more precarious jobs and are more dependent on public transport. The transport alternatives that most benefit from highways are less accessible to the non-skilled, as these solutions may weigh more on their budget. High-skilled workers, on the other hand, are less bonded by migration costs and should be more willing to move closer to their workplace. As they earn higher salaries, they should de able to afford higher costs of living than their counterparts.

In Table 5, “Non-skilled” refers to those who did not finish the Primary education level. The “Primary” columns comprise those who completed at least the first nine years of the mandatory educational path (also known as Elementary School). “Secondary” columns refer to those who finished at least High School, and “Tertiary” includes those who have an university degree or beyond.

The results confirmed our expectations. The estimates show no significance for the “Non-skilled” category (column 1), significance for workers with primary and secondary education (columns 2 and 3), and no significance for workers with tertiary education. When compared to our main results, the effect of highway access for the primary and secondary education groups are larger than for the general population (compare to Table 3).

These findings emphasize that people with middle-level education seem to be more willing to travel greater distances and for a longer duration than either non-skilled or very skilled workers. Possible reasons could explain these effects. Non-skilled workers are generally poorer than their counterparts, and, according to Duarte (2020), low-income workers gravitate towards informal work near home due to the high costs of commuting. Those low-skilled workers find little incentive travelling long distances, as they cannot benefit from a larger labor market that offer better opportunities for skilled workers. Having access to a highway, thus, should not drastically change those workers travelling patterns. We can also find reasons to explain the absence of significance for workers with an university degree. Travel to work is usually a means to avoid migration costs; it allows the worker to benefit from a larger labor market without paying more in rent and costs of living. Chen, Voigt e Fu (2021), for instance, found that cities with higher rent have more inbound commuting of workers from nearby cities, showing that most workers avoid moving to an expensive city and prefer to live elsewhere, pay cheaper rent, and travel to work. High-skilled workers, on the other hand, earn higher salaries, and should be less restrained by costs of living. They have higher opportunity costs, thus they are more willing to move closer to work even if this means higher costs. There, these skilled workers should be near better professional opportunities and can benefit from the amenities of a larger city.

Although a positive relationship between education and commuting was already detected by Cassel et al. (2013) in Sweden and by Silva e Freitas (2020) in Brazil, our study is the first to explore how each educational level respond causally to the provision of highways.

Table 5 – Results for education level

	(1) Non-qualified	(2) Primary	(3) Secondary	(4) Tertiary
Highway access	0.460	0.843**	0.710**	0.537
	(0.289)	(0.387)	(0.334)	(0.384)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5162	5162	5162	5162
First stage F	43.41	43.41	43.41	43.41
Kleibergen-Paap p-value	1.11e-08	1.11e-08	1.11e-08	1.11e-08
Anderson-Rubin p-value	0.115	0.0267	0.0217	0.147

Clustered standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

For more details, see Table 3.

1.5.3 Robustness analysis

Our instrumental variable choice assume that the national roads network aimed to connect capital cities, so other cities that are located in their path are there accidentaly. However, to strengthen our exclusion restriction, we followed Faber (2014) and excluded from the sample all cities within 50 km from the state capital cities for our main estimations. As this threshold may appear ad-hoc, in this section we use different sample of cities according to other criteria; the capital cities, however, are always excluded from the sample.

Our data show that most of the highway network expansion occurred in the past (see Table 1) as capital cities were prioritized. The roads that were built in the last decade were mainly further from capital cities, which may suggest that the relationship between roads and travel to work growth is stronger far from the capitals. We thus investigate if the inclusion of more cities near the capitals weaken the relationship between roads and travel to work growth. We first vary the threshold from 50 km to 30 km, as it is possible that a threshold of 50km was too conservative. Thus, in the columns 1 and 2 of panel “All” of Table 6, we run regressions letting more cities in the sample. Compared to the main results (see Section 1.5.1), the coefficients are statistically the same and maintain significance at least at the 10% level. If the estimations were significantly different, this could indicate that, despite our identification strategy, the highways of our sample were still being influenced by the economy of the municipalities. Put differently, our explanatory variable would be endogenous. These results, however, validate our estimation strategy, and also demonstrate that our estimations are robust to the variation of subsamples.

In columns 3 and 4 of panel "All" of Table 6, we only remove cities that are part of metropolitan regions. This is an important robustness test, as being part of a metropolitan region implies participation in policies that benefit the whole area (metropolitan transport solutions being a meaningful example). The estimated coefficients are similar to the main estimation (Table 3) in significance, signals, and size. This is evidence that our results are not simply a consequence of policies that unite municipalities of a same metropolitan region.

As maritime transport has a close relationship to land transport for cargo, the presence of ports can drastically change the transportation dynamics and infrastructure of a city. Cities with ports tend to have a historical link with others based on the transport infrastructure necessary for the flow of production. Also, populous cities with highways are generally denser cities, as infrastructure is only justified if it will be used by many people. So we also exclude from the sample cities with more than 300 thousand residents and that had access to a highway to control for unobservable factors associated with agglomeration gains and transportation infrastructure at the state level.

For the next results, we return to our original 50 km threshold and impose additional restrictions. In columns 1 and 2 of panel "All" Table 6 we ran our analysis excluding every municipality that had a port. The coefficients remain significant at the 10% level, but become smaller. In columns 3 and 4 of panel "All" Table 6, we exclude cities with populous cities that had highway access. The estimated effects are of the same magnitude and sign of our main results, and also maintain a significance of 5%. These results show that our results are robust to important sample variations.

We also run our analysis for men and women separetely in the panels "Men" and "Women" of Table 6. In accordance to our results in Subsection 1.5.2.1, all results are significant for men but not for women.

Table 6 – Varying the subset of cities by gender

All				
	(1)	(2)	(3)	(4)
	30km from capitals	Metropolitan regions	Removing cities with ports	Removing populous cities
Highway access	0.442*	0.447**	0.501**	0.495**
	(0.226)	(0.219)	(0.239)	(0.240)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5357	4240	5127	5144
First stage F	45.04	49.63	43.35	43.81
Kleibergen-Paap p-value	5.86e-09	6.13e-09	1.16e-08	9.96e-09
Anderson-Rubin p-value	0.0509	0.0458	0.0366	0.0399
Men				
	(1)	(2)	(3)	(4)
	30km from capitals	Metropolitan regions	Removing cities with ports	Removing populous cities
Highway access	0.528**	0.471*	0.604**	0.592**
	(0.252)	(0.245)	(0.268)	(0.269)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5357	4240	5127	5144
First stage F	45.04	49.63	43.35	43.81
Kleibergen-Paap p-value	5.86e-09	6.13e-09	1.16e-08	9.96e-09
Anderson-Rubin p-value	0.0342	0.0589	0.0218	0.0251
Women				
	(1)	(2)	(3)	(4)
	30km from capitals	Metropolitan regions	Removing cities with ports	Removing populous cities
Highway access	0.221	0.300	0.246	0.252
	(0.281)	(0.286)	(0.290)	(0.290)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5357	4240	5127	5144
First stage F	45.04	49.63	43.35	43.81
Kleibergen-Paap p-value	5.86e-09	6.13e-09	1.16e-08	9.96e-09
Anderson-Rubin p-value	0.435	0.300	0.401	0.390

Clustered standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

For more details, see Table 3.

For completeness, in Table 7 we run the robustness analysis for every education level and compare it to the main results in Table 5. The first panel shows results for the non-qualified group. Only the third column showed a significant result, albeit at the 10% level and smaller in magnitude than the main result. The insignificance of three out of four variations of the sample indicates that highways did not change the commuting patterns of this group significantly, which is aligned with the results obtained in Table 5. The second and third panel show results for the primary and secondary education groups, respectively. For both groups, the results are all positive, significant and similar in magnitude to their main results. Their standard deviations show that the coefficients are statistically the same with the ones on Table 5. The fourth panel show results for the tertiary education group. As all estimations provided insignificant results, we have strong evidence to affirm that the effect of highway access on the commuting patterns of workers with an university degree is non-existent. These results show that our results are robust. Even when the sample is modified in important ways, the results remain significant and statistically equivalent to the main results.

Table 7 – Varying the subset of cities by education level

	(1)	(2)	(3)	(4)
Non-qualified				
	30km from capitals	Metropolitan regions	Removing cities with ports	Removing populous cities
Highway access	0.384	0.366	0.475*	0.461
	(0.269)	(0.267)	(0.286)	(0.288)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5357	4240	5127	5144
First stage F	45.04	49.63	43.35	43.81
Kleibergen-Paap p-value	5.86e-09	6.13e-09	1.16e-08	9.96e-09
Anderson-Rubin p-value	0.161	0.179	0.0996	0.113
Primary				
	(1)	(2)	(3)	(4)
	30km from capitals	Metropolitan regions	Removing cities with ports	Removing populous cities
Highway access	0.722**	0.617*	0.840**	0.837**
	(0.363)	(0.373)	(0.386)	(0.386)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5357	4240	5127	5144
First stage F	45.04	49.63	43.35	43.81
Kleibergen-Paap p-value	5.86e-09	6.13e-09	1.16e-08	9.96e-09
Anderson-Rubin p-value	0.0447	0.102	0.0267	0.0273
Secondary				
	(1)	(2)	(3)	(4)
	30km from capitals	Metropolitan regions	Removing cities with ports	Removing populous cities
Highway access	0.727**	0.818***	0.706**	0.708**
	(0.321)	(0.309)	(0.333)	(0.333)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5357	4240	5127	5144
First stage F	45.04	49.63	43.35	43.81
Kleibergen-Paap p-value	5.86e-09	6.13e-09	1.16e-08	9.96e-09
Anderson-Rubin p-value	0.0133	0.00379	0.0219	0.0217
Tertiary				
	(1)	(2)	(3)	(4)
	30km from capitals	Metropolitan regions	Removing cities with ports	Removing populous cities
Highway access	0.517	0.402	0.546	0.534
	(0.385)	(0.357)	(0.384)	(0.381)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5357	4240	5127	5144
First stage F	45.04	49.63	43.35	43.81
Kleibergen-Paap p-value	5.86e-09	6.13e-09	1.16e-08	9.96e-09
Anderson-Rubin p-value	0.160	0.249	0.139	0.146

Clustered standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

For more details, see Table 3.

1.6 CONCLUSIONS

This research aimed to assess the impact of highway access on commuting patterns, with a particular focus on travel to work growth. Our approach leverages an instrumental variable strategy in a unique and compelling way to address potential endogeneity concerns. The main model estimated reveals a statistically significant and economically meaningful relationship between the presence of a highway and the travel to work growth.

Our findings indicate that the construction of highways significantly impacts the commuting patterns, leading to an increase in travel to work growth. The effect appears to be more substantial for men than women, a result that adds a new dimension to the existing literature on transportation and labor market outcomes. Furthermore, we found that the level of education plays a crucial role in this relationship. The construction of highways had a more substantial impact on commuting patterns of individuals with primary and secondary education, but no significant effect was found for those with tertiary education.

The results were found to be robust to a variety of sample modifications and specification checks. In particular, our analysis controlled for the impact of proximity to capital cities, participation in metropolitan areas, presence of ports, population size, and the dynamics related to agglomeration gains and transportation infrastructure at the state level.

The study also addresses the role of transportation infrastructure in mitigating spatial mismatch in developing countries, emphasizing the importance of better resource allocation for promoting equality and economic growth. The results suggest that investing in transport infrastructure can be a viable solution to counterbalance the challenges posed by a spatially concentrated job market and alleviate the hardships faced by workers with limited job opportunities. Furthermore, this research adds to the growing literature that seeks to understand the historical process of metropolization and urban expansion in Brazil, with a specific focus on workers' commuting patterns.

Our research also opens up avenues for future research. For instance, exploring the reasons behind the observed gender differences could provide further insights into the role of transport infrastructure in shaping labor markets. Similarly, understanding empirically why highways do not seem to impact the commuting patterns of tertiary-educated individuals could be an interesting area of exploration.

In conclusion, our study provides strong evidence of the influence of transport infrastructure on commuting patterns. By applying an innovative methodology and rigorous robustness checks, we deliver reliable insights into this critical area of study.

2 THE IMPACT OF HIGHWAYS ON COVID-19 MORTALITY: A CASE STUDY FOR BRAZIL

2.1 INTRODUCTION

The COVID-19 outbreak has emphazised the need for effective strategies to control the spread of infectious diseases. Understanding how a pathogen spreads within a region is of paramount importance in combating its transmission and mitigating its impact on public health. There is already evidence that increased mobility resulted in heightened exposure to COVID-19 (BAUM-SNOW; GLAESER; ROSENTHAL, 2022; GLAESER; GORBACK; REDDING, 2022; MANGRUM; NIEKAMP, 2022), however, there is little research on the role of transportation infrastructures in influencing the proliferation of the disease. It is crucial to investigate how significant is their impact, and what implications does it have for countries like Brazil, characterized by ongoing development and an extensive[1] but poorly maintained road network (CNT, 2017b).

This article aims to explore the relationship between highways and COVID-19 mortality rates in Brazil, shedding light on the potential implications for disease control strategies in a developing nation. Although Brazil is a large contry with spread-out cities, the majority of its population is concentrated in urban areas. The combination of poor transportation infrastructure and low-income population produced, over the years, a pattern of cities with very high density and remarkably high commuting costs (INGRAM; CARROLL, 1981; FERNÁNDEZ-MALDONADO et al., 2014; NETO; DUARTE; PÁEZ, 2015). By examining the interplay between transportation networks and disease transmission, we aim to determine the extent to which highways contributed to the spread of the virus in this setting.

Specifically, we employ an instrumental variables regression approach to examine the causal effect of the length of highways in 2017 on the COVID-related mortality rate in Brazil in 2020. To construct our instrumental variable, we utilize maps of the Brazilian railway network from the 1960s. Historical data indicate that past railway development in Brazil primarily aimed to support an export-oriented economy, with limited consideration for future service accessibility (NATAL, 1991). Over time, railways became less relevant. During the 1950s, the construction of the present-day Brazilian highway system took place. The project aimed to connect the country's capitals as part of a national integration policy (MORTEN; OLIVEIRA, 2018; CHEIN; PINTO, 2016). Due to the shared geographical challenges involved in constructing railways and

[1] The federal highways network of the country comprises 65.6 thousand km.

highways, the presence of old railway systems can be utilized as a predictive indicator for the current state of road infrastructure (DURANTON; TURNER, 2012; GARCIA-LÓPEZ; HÉMET; VILADECANS-MARSAL, 2017).

Although the disease initially transmitted through airports (CANDIDO et al., 2020), a substantial portion of its propagation between cities occurred via roads (CARMO et al., 2020; NICOLELIS et al., 2021; DESMET; WACZIARG, 2022; LI; MA, 2022). However, to our knowledge, the works that examine the impact of roads on COVID-19 transmission do not employ causal methods. By investigating this relationship using causal inference, we aim to fill this gap and provide insights into how highways can causally impact the spread of contagious diseases, particularly for an important developing country.

Our study uncovered that cities possessing twice the average length of highways noted a 14% higher mortality rate in COVID-19-related deaths. This finding underscores the pivotal role of highways in the spread of contagious diseases. By acknowledging this result, public health officials can formulate better methods for limiting disease spread and improving public health outcomes in future crises.

This paper is organized as follows. In the second section, we present a review of the literature; in the third section we present our empirical strategy and data; in the fourth section we present the results; and the last section brings our conclusions.

2.2 LITERATURE REVIEW

Desmet e Wacziarg (2022) propose two conceptualizations of the impact of spatial disparities on disease transmission. The first postulates that the observed geographic variation in disease spread and severity merely reflects time-delayed similarities across all regions, implying that infection, mortality, and hospitalization rates will eventually converge nationally. Consequently, policy interventions need not take local specificities into account. Alternatively, the second view asserts that intrinsic geographic differences – such as population density, transportation systems, and general health status – significantly influence the spread of the disease in different regions. Thus, this perspective underscores the need for policies that are fine-tuned to local conditions.

In this section, we will show that the evidence tends to support the latter approach. Specifically, we will highlight the importance of transportation networks in explaining geographic disparities in infectious disease transmission. Population density does not singularly determine which locations will transform into spreader nodes; as cities are integrated within intricate transportation networks, the connectivity of these networks significantly influences the geographical dispersion of infected individuals (COELHO et al., 2020; NICOLELIS et al., 2021; MITZE; KOSFELD, 2022).

2.2.1 Transportation Infrastructure and Diseases Transmission

Correlations have been detected between higher occurrence of infectious diseases and areas adjacent to significant transportation routes. Past literature have found that highways can be conduits for disease transmission, exemplified by instances such as the rural transmission of syphilis in North Carolina, and the diffusion of H1N1 throughout China (COOK et al., 1999; FANG et al., 2012; SOUCH; COSSMAN; HAYWARD, 2021). In this literature, two studies deserve further inspection.

First, the study conducted by Oster (2012). The author uncovered a strong and positive correlation between export growth and the incidence of new HIV infections in Africa, employing regression analysis to derive these insights. Importantly, the strength of this relationship was more robust in contexts with denser road networks. The author also showed that areas near roads exhibited higher HIV prevalence overall, as confirmed by utilizing HIV rates and GIS data on African road networks.

Second, the work of Adda (2016) that sought to examine the relationship between economic activity, transportation infrastructure, and the spread of viruses. The author used high frequency data from France spanning a quarter of a century. The study employed a quasi-experimental design to evaluate the effects of policies aimed at limiting interpersonal contact, such as the closure of schools and public transportation networks. A notable outcome of the study was the discovery that expansions of transportation networks, including railway lines, were correlated with increased virus spread, suggesting that such infrastructural developments pose significant health costs.

The results of these works underscore the importance of road networks and people movement in driving an epidemic.

2.2.2 Transportation Infrastructure and COVID-19 Transmission

During and after the pandemic, several studies have investigated the role of transportation infrastructure specifically in the spread of COVID-19.

Coelho et al. (2020), for instance, employed an exponential growth model to analyze the spread of COVID-19 on time series data for 65 countries. In their findings, variables such as Gross National Income, annual population growth, healthcare investment, mean temperature, and precipitation demonstrated no significant effect on the exponential phase of COVID-19's spread. Nevertheless, two crucial factors influencing COVID-19's growth rate were identified: a country's population size and its centrality in the global transportation network. Larger populations, particularly in urban areas, were linked with faster virus spread, which could be ascribed in part to the challenges associated with managing virus outbreaks in areas with dense populations. Also, the position of a country in the global transportation network was a significant factor influencing the disease's exponential growth.

Harris (2021) investigated the role of the subway system in the spread of SARS-CoV-2 in New York City in early 2020. The study demonstrated that the rapid citywide transmission was facilitated by subways and that the subsequent reduction in subway travel led to decreased virus transmission. Additionally, areas with a lesser decline in subway use eventually became virus hotspots. This research employed spatial regression models, phylogenetic analyses of viral isolates, and smartphone tracking data to estimate subway visits per zip code tabulation area (ZCTA).

In their 2022 study, Desmet e Wacziarg investigated the determinants of spatial variation in COVID-19 severity across the United States. Through county-level regression analysis, they identified the role of various dimensions of density, including local transit, as a key predictor of disease severity. They found that initial disease spread was correlated with proximity to major international airports, indicating global connectivity's role in disease diffusion. However, this influence decreased over time as the virus penetrated more isolated areas. The study further revealed the variable impact of public transportation. While it initially had a strong positive effect on death rates, this correlation lessened over time, potentially due to changes in behavior, such as increased caution or the shift to remote work. The authors also noted difficulties related to working with COVID-19 data. Both cases and deaths data impose challenges, but the problem is more acute for the former. Reported cases depend on testing and, during the initial stages of the pandemic, testing was neither uniform nor prevalent.

Wan e Wan (2022) conducted an empirical study to ascertain the effect of intercity high-speed railway (HSR) connections on the transmission of COVID-19 during the outbreak in Wuhan, China, when government interventions were absent. Using the instrumental variable method and adopting a least cost routing approach akin to Faber (2014), they found that cities connected to Wuhan via the HSR system experienced a notable increase in COVID-19 cases. This increase was quantified as an addition of 0.029 cases per 10,000 population, contributing to 45% of total infections. They argued that this elevated intercity transmission was primarily due to the facilitated human mobility between cities via HSR connections.

Using a spatial general equilibrium model, Li e Ma (2022) argued that the spatial transmission of COVID-19 in China was significantly influenced by migration flows and transportation infrastructure, particularly in prefectures closely connected to Hubei, the epicenter of the outbreak. The study highlighted that 28% of the infections outside Hubei were attributed to the recent developments in transportation networks (roads, railways, high-speed railways, and waterways) and liberalization of migration policies. Specifically, without the expansions in the transportation networks from 2005-2015, infections could have been reduced by 15.31%, while maintaining the migration policy as it was in 2005 could have resulted in a 17.82% reduction in infections. Notably, if both factors were reverted to their 2005 configurations, a substantial reduction of 28.21% in the spread of COVID-19 could have been achieved, particularly benefiting coastal areas and regions closer to Hubei.

2.2.3 Roads and COVID-19 Transmission

Even with the swift expansion of aviation, maritime, and train transportation in various global regions, highways remain the predominant mode of human transit on sub-national and regional levels (STRANO et al., 2018). Thus, it is also important to consider the specific role of roads in the spread of COVID-19.

Souch, Cossman e Hayward (2021) explored the influence of transportation and mobility on the spread of COVID-19 in the United States by analyzing spatiotemporal patterns using county-level data. They identified a more rapid arrival of the virus in counties intersected by interstate highways, with the time difference most pronounced in rural areas, highlighting the role of road travel in the transmission of the disease. The researchers linked this pattern to human mobility, suggesting that interstate travel restrictions might have strengthened mitigation efforts during the early stages of the pandemic and decreased transmission through network contact.

Studies have also investigated the role of roads in the spread of COVID-19 in Brazil. According to Nicolelis et al. (2021), SARS-CoV-2 initially infiltrated Brazil through its major international airports in March 2020. The virus then rapidly disseminated to adjacent large metropolitan areas, primarily state capitals. Once the virus became entrenched and community transmission started to rise sharply in these cities, the virus began its national spread. This was further facilitated by the lack of major roadblocks during the initial phase of the epidemic. Consequently, a handful of these state capitals played a critical role in transmitting SARS-CoV-2 across the country via Brazil's extensive highway network.

Carmo et al. (2020) investigated the role of a major highway, the BR-232, in the spread of COVID-19 in the Brazilian state of Pernambuco. They performed an ecological study in which they gathered data during twelve epidemiological weeks, starting from the first confirmed case on March 12, 2020. The found that initial cases were confined to Recife, the capital, potentially facilitated by commercial flights, as the city has the busiest international airport in the region. However, a marked shift in the pattern of spread was noticed a few weeks later, with a steady increase of cases in municipalities located along the BR 232 highway. This highway, which spans 563 kilometers from the coast to the hinterland, was identified as a crucial factor in the interiorization of the virus, indicating that ground transportation routes may have significantly contributed to the virus's spread to rural areas.

The study conducted by Chauvin (2021) offers an analysis of the COVID-19 pandemic's varying impacts across over 2,500 Brazilian cities, drawing comparative insight from the United States when feasible. The research utilized OLS regressions with state fixed effects to control for non-observed state-specific factors. Population density was found to amplify the local impact of the disease in both the US and Brazil. Unlike in the US, however, higher income levels in Brazilian cities, correlating with increased mobility, experienced a greater toll from the pandemic. Commuting played a significant role in Brazil, as cities with longer average commute times witnessed fewer cases. However, this connection with commute times weakened and, in most cases, disappeared after the first wave's peak.

Nicolelis et al. (2021) employed an epidemiological compartmental model to elucidate the factors influencing the spread of COVID-19 in Brazil. They discovered that the geographical spread of the disease was significantly influenced by transportation networks, particularly roads, as well as the distribution of intensive care units (ICUs) across the country. The researchers found that São Paulo, termed a "super-spreading city", contributed to more than 85% of the initial case spread throughout Brazil, with an additional 16 cities accounting for nearly all the cases reported in the first three months of the pandemic. Notably, the spread was found to correlate with 26 federal highways, contributing to approximately 30% of the case spread. Moreover, the uneven distribution of ICUs, primarily concentrated in state capitals, created a “boomerang effect”, as patients from rural areas were transported to these capitals, thereby intensifying the geographical distribution of COVID-19 deaths.

To our knowledge, no study used identification strategies to evaluate the causal impact of highways on COVID-19 mortality rates. We not only innovate by employing causal analysis to examine the relationship of road networks with COVID-19 deaths, but we also accomplish this in a large and important developing country, known for its spread-out cities and its extensive but poorly maintained road network.

2.3 EMPIRICAL STRATEGY AND DATA

2.3.1 Specification

In this subsection, we propose an empirical model that describes how highways affect COVID-19 mortality rates. Let i be the index of cities and j the index of states. The model is:

$$\text{Log(Mortality rate)}_i = \alpha + \beta \cdot \text{Log(Highway length)}_i + \Gamma \cdot \text{Controls}_i + \lambda_{j(i)} + \varepsilon_i \qquad (2.1)$$

In the econometric model, the variable Mortality rate$_i$ represents the COVID-19-related death count per 100,000 inhabitants in the year 2020. We worked with mortality data instead of confirmed cases because, as argued by Desmet e Wacziarg (2022), cases data should be less reliable than COVID-19 related deaths data. During the most acute phase of the pandemic, tests were not as prevalent and were not conducted uniformly across the country. The variable Highway length$_i$ denotes the length of highways, measured in kilometers, within each municipality i in 2017.

The vector Controls$_i$ comprises a set of seven variables used as control factors in the analysis. Three variables are related to geography and space: log distance to the nearest capital, log distance to the coast, and log distance to Brasília. They control for paths that already had intense levels of transportation before the analysis. Additionaly, we include 2010 level of formality of the workforce of cities. Formal workers may have different working conditions, such as better workplace safety measures or the ability to work remotely, reducing their risk of contracting the virus, which may influence mortality rates. Also, the degree of formality can also reflect access to healthcare insurance plans. We also include the degree of workers who work in service industries, so we can capture the potential impact of transmission dynamics within this sector. Service Jobs often involve close interpersonal contact, especially in industries such as hospitality, retail, and personal services. Also, the services sector typically requires workers to commute and interact with a diverse range of individuals, which may lead to higher levels of exposure to the virus. Finally, we include income per capita and population density of 2010 to account for previous levels of development.

The model is estimated with state fixed effects $\lambda_{j(i)}$. The country has 26 states and one federal district. As our data include federal highways only, the state fixed effects can help to partial out the effect caused by state roads. We also cluster errors at the state level.

To avoid outliers, we exclude the capital cities from our sample. For dealing with possible endogeneity problems, we use an instrumental variable approach which is described in the next section.

2.3.2 Endogeneity

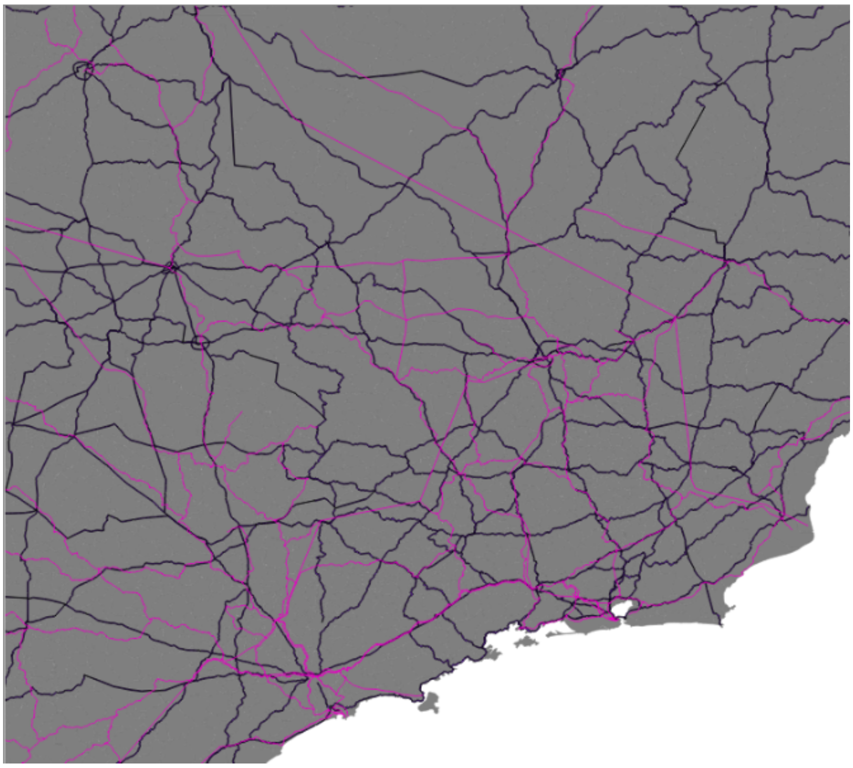
The literature has shown that standard OLS estimations of the impact of transportation networks can suffer from endogeneity problems (BAUM-SNOW; FERREIRA, 2015). While transportation infrastructure may promote accessibility, it is also true that cities expected to grow are likely candidates for receiving transport investments. These estimations are also prone to suffer from omitted variable bias, as there can be unobservable characteristics affecting the mobility of people.

To address these challenges, an established strategy in the literature is to utilize old transportation networks as a predictive tool for contemporary infrastructure. As noted by Garcia-López, Hémet e Viladecans-Marsal (2017), it is well known in the literature that modern roads are usually built near routes established by past infrastructure. Duranton e Turner (2012) argued that railroad and highway builders have to deal with similar geographical limitations. For both, the ideal path between places consists of straight lines, which strongly depends on terrain features. Thus, the presence of railways in a region can be a good indicator of a suitable terrain for future roads.

Baum-Snow (2007) used 1947 plans of the Interstate Highway System as an instrument for the network that was built later, being the first work to use construction plans to instrument for existing infrastructure. Duranton e Turner (2012) uses the same plans, along with railway networks from the 19th century and expedition routes in past centuries as instruments for current highways' location in the USA. Garcia-López, Hémet e Viladecans-Marsal (2017) uses Roman roads and 1870 railways as instruments for the Regional Express Rail (RER) infrastructure in Paris metropolitan area.

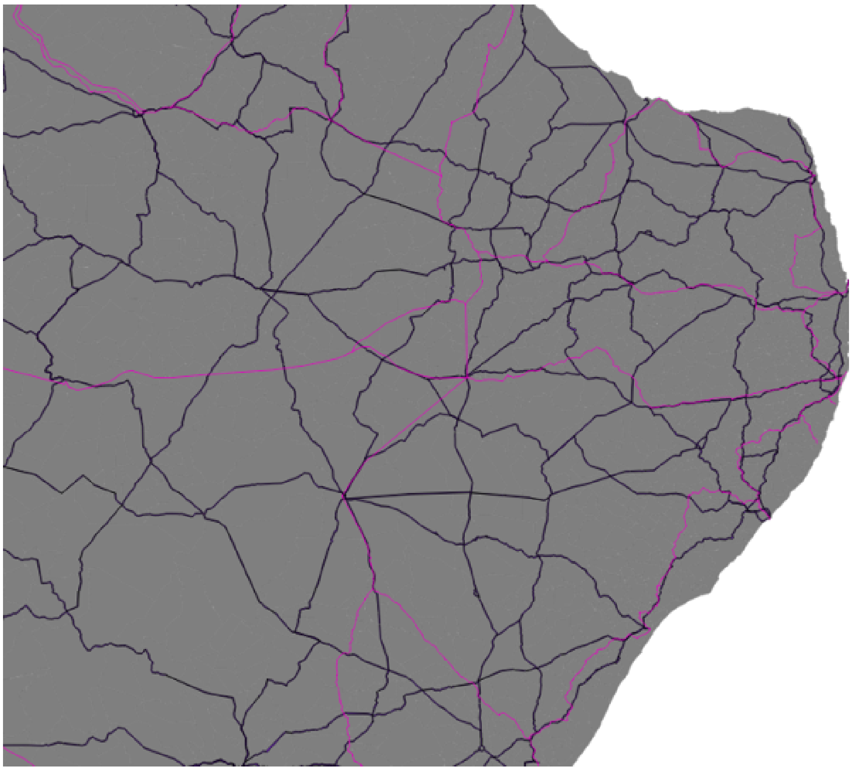
The Brazilian case seems suitable for employing this strategy. First, current highway network has positive correlation with past railway network. Figures 3 and 4 show current areas in Brazil in which several intersections between highways and railways are visible.

Figure 3 – Railways and highways in southeast region



Note: Railways are in pink while highways are in black.

Figure 4 – Railways and highways in northeast region



Note: Railways are in pink while highways are in black.

Historical data had already pointed that the initial Brazillian highways were often built following existent railways (NAZARETH, 1978). In Figures 5 and 6, we show the stock of railways in 1960 and the stock of highways in 2017, respectively. A simple correlation analysis show that 1960 railways and 2017 highways have a correlation of 0.1967. The coefficient is high enough to point that the railways are still relevant to predict the more recent Brazilian highway stock due to shared geographical requirements. In Figure 7, we present the graph of a linear regression between the log of highways length and the log of railways length. The figure shows a positive relationship between both variables.

Figure 5 – Railways in 1960

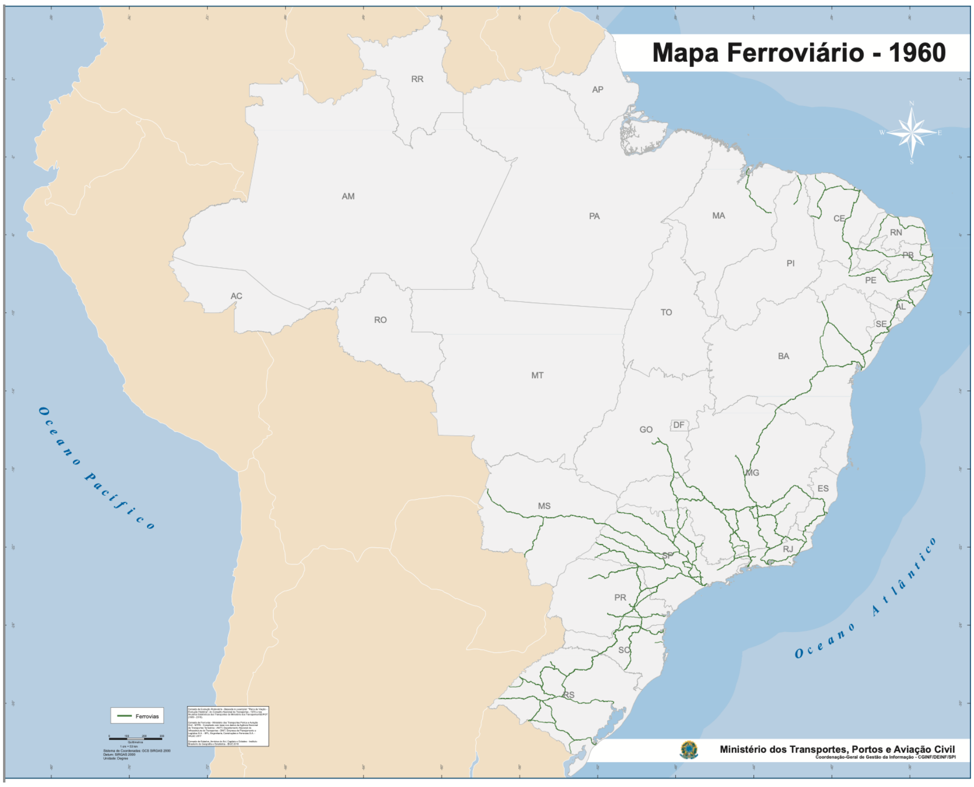


Figure 6 – Highways in 2017

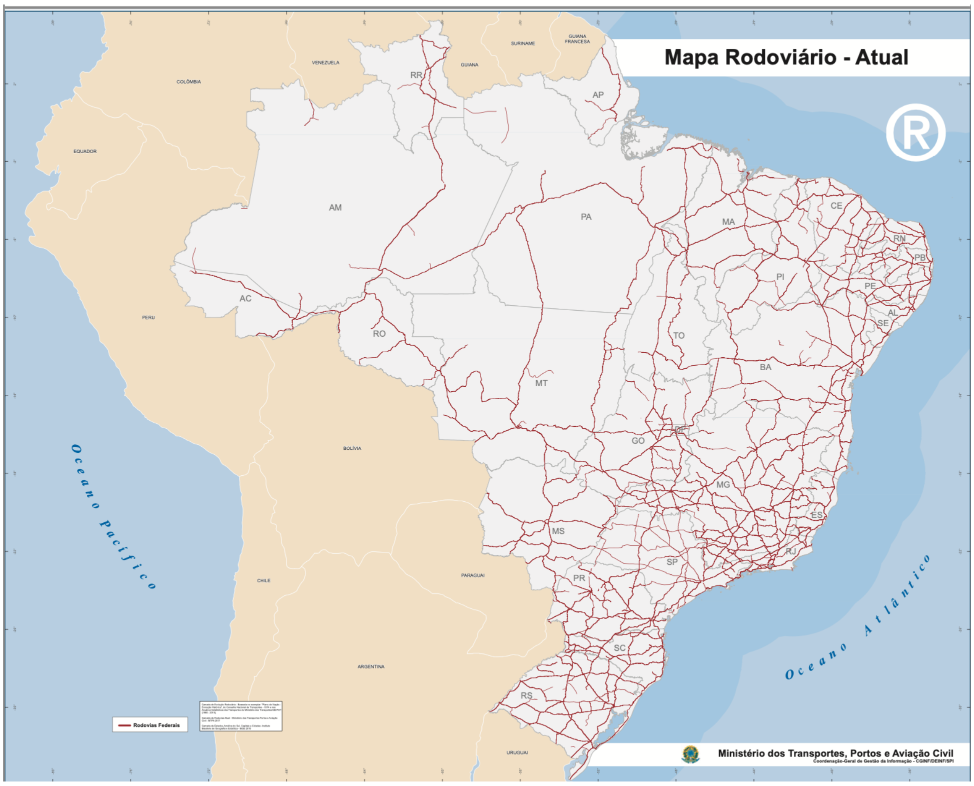
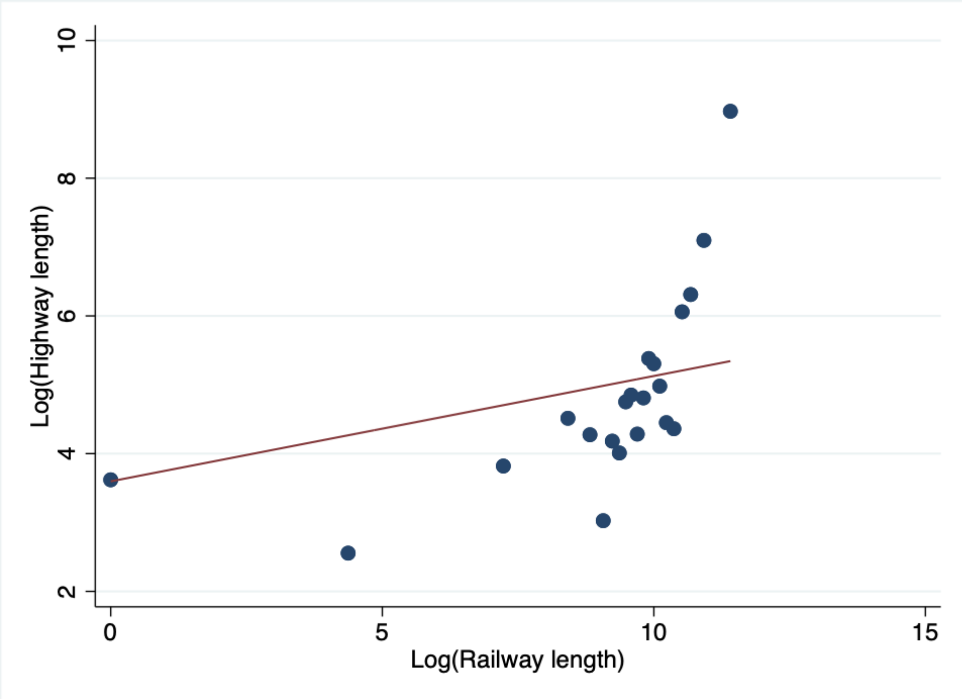


Figure 7 – Linear regression between highways and railways
Note: Binned scatterplot showing 100 points.



Second, past railway system data also satisfies the exclusion restriction. Historical data attest that past Brazilian railways were built for purposes exogenous to our variables of interest. Railway from past centuries had the purpose of supporting an export-oriented economy (NATAL, 1991), with no regard for future accessibility to services. Very few of them promoted interregional connections; so much that the early Brazilian economy has been called an "economic archipelago". Because of limited interregional transport, transporting cargo between inland cities was often hard. It usually required to first move the load from the interior to the coast, ship it to another port and then take it inland again. Because the transport even between inland cities often relied on maritime navigation, Brazilian cities resembled, in fact, a set of islands separated by the sea. From the early national railway networks in the 19th centuries to the construction of the current Brazilian highway system that began in the 1950s, there is a significant time gap through which the Brazilian economy has undergone many changes. Over this time, railways became less and less relevant to the national transportation network, while roads became predominant (NAZARETH, 1978). In contrast to early railways, the first national highways were built to connect the capitals of Brazil as part of a national integration and development policy (MORTEN; OLIVEIRA, 2018; CHEIN; PINTO, 2016). It follows that the only way past railway network can affect current development (conditional on control variables) is through the current network of highways.

Thus, our identification strategy will rely on the national railway system as of 1960 as an intrument for the highway network of 2020.

2.3.3 Data

Our dataset was obtained from four sources.

Our dependent variable came from the "COVID-19 data in Brazil: cases, deaths, and vaccination at municipal (city) level" dataset (COTA, 2020), a voluntary and public initiative that collected and aggregated official data from the Ministry of Health and State Health Departments.

The Brazilian Demographic Census is performed by the Brazilian Institute of Geography and Statistics (IBGE) every ten years. We use Census data of 2010 to obtain our control variables at the municipality.

The Ministry of Transport, Ports and Civil Aviation provides digital maps and pictures of the highways and railways of Brazil for each decade from 1960 to 2017, and the IBGE provides digital maps for the Brazilian municipalities. From these data, we were able to compute our main explanatory variable and also the instrument using Python and GIS.

We also utilize a set of geographical controls, including the computation of distances to the nearest capital and coastline using GIS. These calculations were performed by the authors using digital layers of city boundaries available from the IBGE.

In Table 8 we show some descriptive statistics for the variables used in the regression analysis.

Table 8 – Descriptive Statistics

	count	mean	std	min	median	max
Log(Mortality Rate)	5480	7.745	2.645	0.000	8.579	10.817
Dummy Highway	5480	0.506	0.500	0.000	1.000	1.000
Dummy Railway	5480	0.193	0.395	0.000	0.000	1.000
Log(distance nearest capital)	5480	12.121	0.723	8.768	12.246	13.742
Log(distance coast)	5480	12.375	1.119	7.369	12.553	14.853
Log(distance Brasilia)	5480	13.781	0.513	10.549	13.848	14.916
Income per capita	5480	560.131	258.794	125.616	532.656	2208.591
Share of services	5480	0.384	0.090	0.057	0.385	0.669
Share of formal jobs	5480	0.557	0.181	0.062	0.568	0.943
Demographic density	5480	0.000	0.001	0.000	0.000	0.013

2.4 RESULTS

2.4.1 Main results

Table 9 presents the results from our regression analysis. In columns 1 and 2 we present OLS regression results first without control variables and then with control variables. The results show that the length of highways is positively correlated with the COVID-19 mortality rate. The coefficient of 0.0486 (column 2) indicates that a 100% increase in the length of the highway extension in a city is correlated with an approximate 4.86% increase in the COVID-19 mortality rate, holding other factors constant.

However, OLS regression results may be biased due to the endogeneity of the highway network. For this reason, in column 3 we present the results from our instrumental variable regression analysis. Length of highways has a statistically significant positive coefficient of 0.142 at the 1% level. This indicates that a 100% increase in the length of the highway extension in a city is associated with an approximate 14.2% increase in the COVID-19 mortality rate. The result shows that cities with a larger highway network saw a higher mortality rate related to COVID-19 during the pandemic than other cities with smaller highway infrastructure. This finding underlines the critical role of transportation infrastructure in disease transmission and reinforces the importance of considering such factors in devising public health policies and strategies.

The first-stage F-statistic of our instrumental variable regression is 6.359, below the conventional threshold of 10, suggesting that our instrument is a weak instrument. However, the Anderson-Rubin test is significant at the 5% level, indicating that our instrument is valid and relevant.

Table 9 – Instrumental variable regression analysis

	OLS	OLS	Second Stage
	(1)	(2)	(3)
	Log(Mortality rate)	Log(Mortality rate)	Log(Mortality rate)
Log(Highway length)	0.0766***	0.0486***	0.142***
	(0.0105)	(0.00771)	(0.0432)
Controls	No	Yes	Yes
State FE	Yes	Yes	Yes
Observations	5480	5480	5480
R^2	0.104	0.157	0.131
			First Stage
			Log(Highway length)
Log(Railway length)			0.155**
			(0.0616)
Controls			Yes
State FE			Yes
Observations			5480
R^2			0.154
First stage F			6.359
Anderson-Rubin p-value			0.0320

* p<0.10, ** p<0.05, *** p<0.01. Clustered standard errors in parentheses.

2.4.2 Robustness analysis

In order to further verify the reliability of our initial findings, we conducted several robustness tests.

We first substitute the logarithm of highway length with a binary variable indicating access to a highway (Table 10). In column 1, the regressor is a binary variable that takes the value of 1 if a city is crossed by a highway, and 0 otherwise. For consistency, we also used a binary variable for railway as the instrument. In column 2, we follow Faber (2014) and use a dummy that takes the value of 1 if there is a highway within a range of 10km around the border of the city, and 0 otherwise. The coefficient of Highway access in column 1 is 1.436, significant at the 1% level. This suggests that cities in proximity to highways experience an approximately 143.6% increase in COVID-19 mortality rate compared to cities that do not contain any segment of a highway, after controlling for other factors. When the binary regressor also considers a border of 10km around the city perimeter, this percentage increases to 249.8%. This robustness test supports our main findings. The result suggests that not only the length but also the mere proximity to highways can contribute to higher COVID-19 mortality rates.

Table 10 – Robustness analysis: Highway access dummy

	(1) Log(Mortality rate)	(2) Log(Mortality rate)
Highway dummy	1.436***	
	(0.499)	
Highway dummy 10km		2.498**
		(1.235)
Controls	Yes	Yes
State FE	Yes	Yes
Observations	5480	5480
R^2	0.125	0.0264
First stage F	6.346	3.043
Anderson-Rubin p-value	0.0472	0.0472

* p<0.10, ** p<0.05, *** p<0.01. Clustered standard errors in parentheses.

The literature indicates that the highway network of Brazil was initially built to connect the capital cities of each state (MORTEN; OLIVEIRA, 2018; CHEIN; PINTO, 2016). Thus, while non-capital cities might be crossed by a highway accidentally, this is never the case for capitals. For this reason, our baseline analysis always excludes capital cities from the sample. To further explore how much the cities closest to the capitals influence our findings, in Table 11 we also exclude cities close to capitals using varying distances[2]. The results maintain significance, even when we exclude cities within a 100km radius from capital cities. The standard errors indicate that their magnitude are statistically equivalent to the baseline results. Thus, this test suggests that our findings are not driven by the cities closest to the capitals.

[2] We were motivated by the work of Faber (2014), that excludes from the sample counties that are within a 50 km commuting radius of the targeted city centers.

Table 11 – Robustness analysis: Removing cities close to capitals

	30km distance	50km distance	80km distance	100km distance
	(1)	(2)	(3)	(4)
	Log(Mortality rate)	Log(Mortality rate)	Log(Mortality rate)	Log(Mortality rate)
Log(Highway length)	0.141***	0.143***	0.134**	0.123*
	(0.0446)	(0.0444)	(0.0570)	(0.0651)
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	5358	5163	4747	4446
R^2	0.127	0.120	0.123	0.123
First stage F	5.890	6.253	6.187	5.913
Anderson-Rubin p-value	0.0514	0.0630	0.120	0.183

* p<0.10, ** p<0.05, *** p<0.01. Clustered standard errors in parentheses.

In Figure 12, columns 1 and 2, we segment our sample into cities that are located within metropolitan areas and cities outside of them. Cities included in a metropolitan region are typically involved in area-wide beneficial policies, including metropolitan transportation solutions, which could affect our results. The results remain significant and robust to this sample variation.

In Figure 12, columns 3, we exclude cities that contain an airport. SARS-CoV-2 arrived in many countries via airport travel (CANDIDO et al., 2020; DESMET; WACZIARG, 2022), including Brazil (CARMO et al., 2020; NICOLELIS et al., 2021). Thus, it is important to verify whether our results are limited to cities that contain an airport. The estimated result, however, indicates that this is not the case. The coefficient of .134 is significant at the 1% level and is statistically similar to the baseline result.

Table 12 – Robustness analysis: Metropolitan regions and airports

	Metropolitan regions	Non metropolitan	Cities without airport
	(1)	(2)	(3)
	Log(Mortality rate)	Log(Mortality rate)	Log(Mortality rate)
Log(Highway length)	0.139***	0.143**	0.134***
	(0.0360)	(0.0636)	(0.0479)
Controls	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Observations	1239	4241	5022
R^2	0.156	0.128	0.131
First stage F	6.592	4.884	7.816
Anderson-Rubin p-value	0.0238	0.181	0.0312

* p<0.10, ** p<0.05, *** p<0.01. Clustered standard errors in parentheses.

Finally, to ensure that our findings are not disproportionately influenced by a particular state or outlier data, we also employed the leave-one-out technique. This method involves running our main regression repeatedly, systematically excluding one different state in each iteration to assess the consistency of our results. The result is presented in Figure 8, and the last data point is the result from our main analysis, without any state being excluded. The outcomes show that the effect of highway length on mortality rate maintain significance even when we exclude one state at a time, reinforcing the main findings from our research and underscoring the robustness of our results.

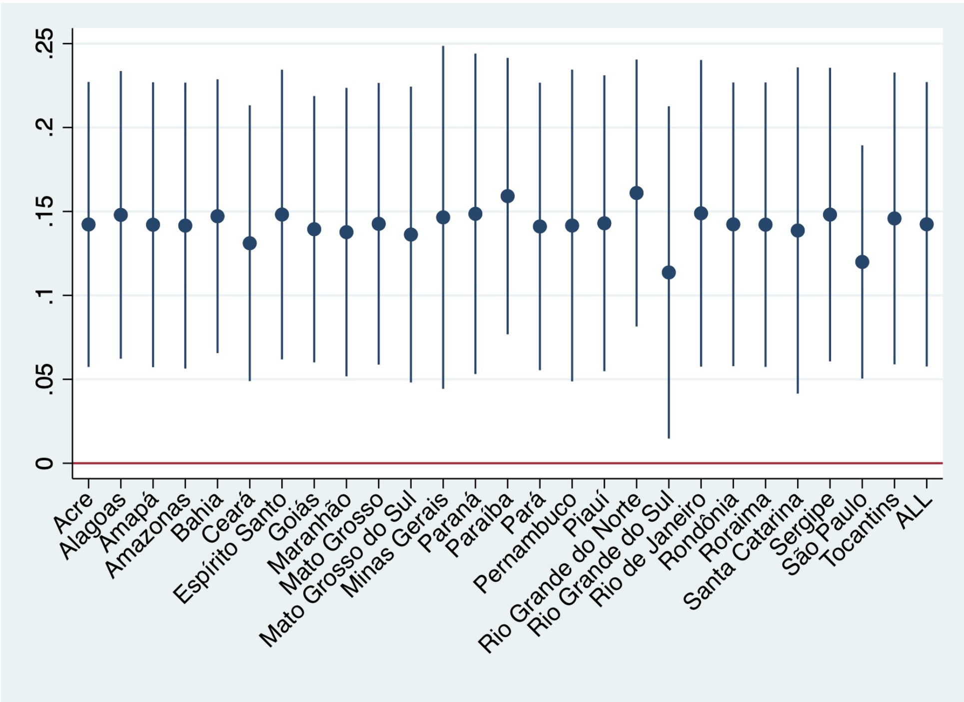


Figure 8 – Leave-one-out analysis

2.5 CONCLUSIONS

Our analysis revealed a causal positive relationship between the length of highways and the COVID-19 mortality rate in Brazil. A more extensive highway network can facilitate the spread of the virus, thereby leading to higher mortality rates.

This has important implications for public health policies, particularly in countries with extensive transportation networks like Brazil. Moreover, this outcome stresses the importance of considering transportation infrastructure when designing and implementing disease control measures. In particular, it suggests that efforts to control the spread of contagious diseases should not be confined to healthcare settings but should also consider broader societal contexts, such as transportation networks.

Future research could expand on these findings by investigating the specific mechanisms through which highways may contribute to the spread of infectious diseases, such as increased interpersonal contact or facilitated movement of infected individuals. Furthermore, our results highlight the need for similar studies in other developing countries to provide a more comprehensive understanding of the links between transportation infrastructure and disease spread.

This study provides a foundation for further investigations into this understudied area, and it underscores the need for multi-sectoral collaboration in tackling public health crises. A deeper understanding of the complex interplay between transportation infrastructure and disease transmission could guide the formulation of more effective and nuanced public health strategies.

3 COMMUTING AND SCHOOL ABSENTEEISM: EVIDENCE FROM BRAZIL

3.1 INTRODUCTION

The literature on education highlights the detrimental effects of school absenteeism on students' future outcomes (GOTTFRIED, 2014; ANSARI; PIANTA, 2019) . While academic performance is often the primary focus of research, grades alone do not fully reflect the importance of the educational experience. In addition to improving academic achievement, school is where students develop many positive non-cognitive skills and behaviors (CORNELISSEN; DUSTMANN, 2019). School absenteeism has been linked, for instance, to decreased educational engagement, social engagement, behavioral issues like alienation (NEWMANN, 1981; FINN, 1989; JOHNSON, 2005), and health-risk behaviors, including drug use (HALLFORS et al., 2002).

If school attendance influences both cognitive and non-cognitive skills development, the determinants of school absence is a relevant research topic in itself. Previous research has found several explanations for school absenteeism: school characteristics (KRUEGER, 2003), teacher's absence (BANERJEE; DUFLO, 2007), parental education (PONTILI; KASSOUF, 2007), student's personality (LOUNSBURY et al., 2004), and weather (GOODMAN, 2014).

Nevertheless, there is a crucial aspect related to absenteeism that has not been given much attention: commute time. Although the labor literature has evidence that long commuting distance to work induces absenteeism (OMMEREN; PUIGARNAU, 2011), and the education literature has found a link between travel time to school and academic performance (KOBUS; OMMEREN; RIETVELD, 2015; TIGRE; SAMPAIO; MENEZES, 2017), few previous research explored the direct relationship between commuting and absenteeism in school, especially for developing countries. Besides, as discussed on the literature review section, the available studies rarely use methods to infer causal relationships. If absenteeism is causally linked to time spent traveling to school, it is possible to propose public policies that mitigate the adverse outcomes suffered by kids who are frequently absent from school due to long commute times (ASAHI, 2014). Moreover, this research also contributes to the ongoing discussion surrounding the optimal location of schools (PIZZOLATO; SILVA, 1997; ARAYA et al., 2012; ARMAS; RAMALHINHO; REYNAL-QUEROL, 2022), as school agglomeration policies might negatively affect school accessibility, especially for low-income students (MORENO-MONROY; LOVELACE; RAMOS, 2018).

The aim of this study is to analyze the impact of travel time to school on absenteeism by using causal inference analysis in a rich survey dataset containing information on students, their families, and their schools.

Our data comprises schools in Recife, an important Brazilian city. Brazil is a developing country known for its urban areas with a high density of residents and remarkably high commuting costs (INGRAM; CARROLL, 1981; FERNÁNDEZ-MALDONADO et al., 2014; NETO; DUARTE; PÁEZ, 2015). Recife is the capital of the state of Pernambuco, located in the Northeast region of Brazil. Founded in 1537, the city is the oldest state capital in the country, featuring a well-preserved central area. It is the fourth most densely populated city in Brazil (IBGE, 2010b). The subject of commuting is specially important for Recife. Between 1992 and 2009, the city saw a gradual increase of over 5 percentage points in the proportion of home-to-work trips that take more than one hour. This increase was concerning because the city has an outdated mass transportation system (PEREIRA; SCHWANEN, 2013). The city has a bus-based public transportation system, supported by an old suburban metro system inaugurated in the 1980s. Its roadway infrastructure is also poorly suited to accommodate the recent expansion of individual transportation, which may contribute to the cost of daily commuting (SOUZA; DUARTE; NETO, 2023). The city of Recife was ranked as the most traffic-congested city in Brazil and the 24th in the world, according to the 2021 Traffic Index conducted by TomTom.[1] The transportation challenges of Recife seem to be worse for those living in the city's peripheral areas, where most of the low-income population resides (COSTA; HA; LEE, 2021).

We used an instrumental variable approach similar to the one used by Kobus, Ommeren e Rietveld (2015) and Tigre, Sampaio e Menezes (2017). Our instrument is the average distance to the two schools closest to the student's residence. As the city government of Recife induces parents to enroll their children in the closest or second closest school from home[2], the instrument should be a strong predictor of commute time and, at the same time, should not affect absenteeism except for its effect on commuting. To ensure the validity of the instrument, we employ additional control variables and perform verification tests that are discussed in Section 3.4.

Our results show that commute time has a significant impact on absenteeism. We found that, if commute time doubles, absences increases by 18.8%. We ran several robustness tests and the results remained significant with little variance. This finding is important because it suggests that public policies that reduce commute time can have a positive impact on school attendance.

This paper is organized as follows: Section 2 presents the literature review. Section 3 pre-

[1] Source (In Portuguese). Accessed on June 14, 2023.
[2] Source (In Portuguese). Accessed on June 16, 2023.

sents a theoretical framework. Section 4 describes the data and the empirical strategy. Section 5 presents the results, and section 6 concludes.

3.2 DAILY COMMUTING AND SCHOOL ATTENDANCE

3.2.1 Absenteeism and its consequences

The negative relationship between student absenteeism and academic performance has been established in the literature for many years. Studies conducted by Summers e Wolfe (1977) in Philadelphia and Monk e Ibrahim (1984) in upstate New York found that students who had a higher number of unexcused absences had a lower performance on tests. More recently, Gottfried (2009) studying data of Philapelphia students highlighted that, when absences are not excused, the negative impact on academic achievement is much greater compared to when absences are due to legitimate reasons such as health issues. The author used linear regression models with fixed-effects to analyse the data. The regression analysis studies conducted by Gottfried (2011) in Philapelphia, Gershenson, Jacknowitz e Brannegan (2015) in North Carolina, and Morrissey, Hutchison e Winsler (2014) in Florida found similar results, showing that there is a significant negative relationship between absenteeism and test scores. On a comparable study, Gottfried e Kirksey (2017) analyzed elementary school data in California with time series regressions and reported that the most critical period is the 30-day window leading up to a test. Additionaly, Gottfried (2014) studied data for American kindergarten students using linear regression and found that chronic absenteeism also hurts reading outcomes. Ansari e Pianta (2019) revealed that absenteeism in the early years of school can have long-term negative effects on academic outcomes. They analysed a sample of American children who were followed from birth through high school using correlation analysis. Finally, Klein, Sosu e Dare (2022) found that all forms of absenteeism, regardless of the reason, are negatively associated with academic achievement both during compulsory and post-compulsory schooling in Scotland. The authors used OLS linear regression with fixed-effects and first-difference models.

A number of studies have also explored the relationship between absenteeism and dropping out rates. Rumberger (1995) found that moderate to high levels of absenteeism are a significant predictor of dropout for middle school American students. Lehr, Sinclair e Christenson (2004) discovered that Minnesota students who are at risk of dropping out can be identified as early as third grade based on attendance patterns and other factors.

While academic performance is often considered the primary goal of school attendance, it is important to recognize that schools play a crucial role in children's overall development beyond academics (GOTTFRIED; ANSARI, 2021). The school environment itself is significant, as it is where children learn a range of positive non-cognitive skills and desired social behaviors. Ryan e Patrick (2001), for example, surveyed eighth-graders from two midwestern school districts regarding changes in motivation and engagement when they moved from seventh to eighth grade. They found that when students perceived their teacher as supportive, they exhibited more self-regulated learning and less disruptive behavior. Also, students' perception of being encouraged to interact with others in the classroom and to share their ideas was correlated with all indicators of motivation and engagement. West et al. (2016) found a positive correlation between school attendance and the development of qualities such as conscientiousness, self-control, grit, and growth mindset in students from fourth to eighth grade in Boston city. Cornelissen e Dustmann (2019) similarly discovered a positive effect of school attendance on student-teacher relationships, academic interest, and reduced disruptive behavior for 11-year-old students in the United Kingdom.

Inversely, several studies have shown a correlation between school absenteeism and decreased educational engagement, reduced social engagement, and behavioral problems such as alienation (NEWMANN, 1981; FINN, 1989; JOHNSON, 2005). Hallfors et al. (2002) conducted a meta-analysis with American school data. They investigated the predictors of drug use among students and found that the odds-ratio for absenteeism among seventh and eighth graders was twice as large as GPA score in predicting drug use. Eaton, Brener e Kann (2008) used logistic regression analysis on data from 8 American states and found that students who were absent without permission were twice as likely to engage in health-risk behaviors, such as tobacco use and drug use. Gottfried (2014) conducted a large-scale regression analysis on a sample of 10,740 kindergarten students in the US and found that chronic absenteeism was associated with a decline in educational engagement (as measured by approaches to learning and eagerness to learn) and a decline in social engagement (as measured by an increase in internalizing behaviors). Fuhs, Nesbitt e Jackson (2018) studied the relationship between school attendance, readiness skills, and executive function (i.e., working memory, inhibitory control, and the ability to focus attention) for preschool children in an American Midwestern city using correlation analysis. They found that children who were chronic absentees made significantly fewer gains in their executive function skills from fall to spring of their preschool year. Ansari e Pianta (2019) investigated the lasting effects of absenteeism during the earliest stages of a child's education.

Through correlation analysis on a American national dataset, they discovered that absenteeism in the first ten years of a child's education was associated with lower academic performance and less favorable social-behavioral outcomes at the age of 15. These problematic outcomes include sexual practices and risky behaviors (e.g., drinking alcohol, smoking, getting into fights). Additionally, they were more likely to continue missing school later in their educational journey. Gottfried e Ansari (2021) conducted regression analysis on a large dataset of 14,370 children in the US and found that chronically absent children demonstrated roughly 1–1.5 fewer months of gains in working memory between kindergarten and third grade compared to their peers. They also reported greater school-related stress, higher levels of victimization and social anxiety, lower levels of motivation and grit, and lower levels of school belonging.

The reviewed literature suggests that school attendance is a significant factor for both cognitive and non-cognitive development in students. However, most studied apply naive methods that are not able to identify causal relationships. Thus, it is worth highlighting two studies that have attempted to address this issue. In 2015, Kobus, Ommeren e Rietveld conducted a study on the impact of commute time on university presence and academic achievement using data from students of VU University in Amsterdam, the Netherlands. They used an instrumental variable approach, while controling for a series of socio-economic characteristics. Their instrument is the mean public transport travel time from the municipality of origin to the closest two Dutch university cities. They argue that this variable predicts the commuting time to the VU University, while having no impact on the potential students' decision to select the VU University – conditional on control variables. The municipality of origin should not directly affect university presence or academic achievement, as most parents have chosen a residence municipality ignoring the effect that this may have on their children's travel time to the university. However, the distance from the municipality of origin does influence commute times, particularly for students living with their parents or those who prefer to live in municipalities with social ties. They found a negative relationship between commute time and grades. Tigre, Sampaio e Menezes (2017), following a similar strategy, conducted a study using Brazilian data that investigated the causal impact of commuting duration on student performance. Their instrument is the average distance of current residence to the two closest schools. They argue that, as the city government asks parents to enroll their children in the closest school, and as they control for school and parenting quality, the effect of the distance from home to school should only affect achievement through commuting costs. They found that longer commutes had a negative impact on academic achievement. Neither studies, however, examine the connection between commuting and absenteeism, leaving a gap in the understanding of whether the link between commuting and grades is driven by absences' impact on academic performance or the influence of commuting on productivity due to fatigue and time loss.

3.2.2 Commuting and Absenteeism

If school attendance is important for students development and achievement, it is crucial to better comprehend the issue of absenteeism. One of the major barriers to school attendance is limited accessibility, with several studies confirming that long commute times, long distances, and inadequate transportation are significant predictors of absences.

Developed countries are not exempt from acessibility problems affecting school attendance. Dexter (1982) published a study that surveyed 155 high school students of the city of Portland regarding absenteeism. The selected sample of students included only freshmen and sophomores considered to have excesive absenteism (more than 10 absences). The author found that 34.3% of those who used buses to commute to school reported "Transportation" as a reason for their absences. More recently, Gottfried (2017) conducted a study in the US using linear regression on a national large-scale dataset and found that kindergarten students who took a school bus were less likely to be absent than those who used other modes of transportation. Similarly, García e Weiss (2018) used regression analysis and observed that a lack of adequate transportation was associated with a higher likelihood of absenteeism, especially chronic absenteeism, among 8th graders in the US. Stein e Grigg (2019) analyzed the attendance records of high school students in Baltimore and Maryland who commute to school. They used first-difference estimation method to address the issue of endogeneity and discovered that students who have to travel longer or face more complex commutes miss more days of school.

This accessibility challenge appears to be even more pronounced in developing countries. In a study conducted in Mozambique, Handa (2002) used probit regression to determine that the distance to a primary school significantly affects children's enrollment. The study found that reducing the travel time to the nearest school in 30 minutes could increase enrollment rates by 20 and 17 percentage points for both boys and girls, respectively. Similarly, Kazeem, Jensen e Stokes (2010) used logistic regression on Nigerian data and found that living 20 or more minutes from the nearest school reduces the odds of attendance by 27 percent for primary schools and 52 percent for secondary schools. Duze (2011) used descriptive statistics to analyze the impact of the distance travelled by students on school attendance in Nigeria's states of Anambra, Enugu, and Ebonyi. The author found that 65.23% of primary school students and 76.09% of secondary school students travelled more than one kilometer to school, which negatively impacted their attendance. Thembo (2011), in Western Uganda, observed a negative correlation between the distance travelled by students from their homes to school and the daily school attendance.

Most studies in this literature do not estimate causality, but two studies that use causal analysis are worth highlighting. Vuri (2010), using multivariate probit with instrumental variable, found that an additional 10 minutes of travel time to school reduced the probability of attendance by 0.2 percentage points in rural Ghana. Burde e Linden (2013) conducted a randomized controlled trial in rural northwestern Afghanistan to assess the impact of village-based schools on primary school-age children. They randomly assigned village-based schools to 13 villages among a sample of 31 villages and found that reducing the distance to school increased enrollment rates by 35% for boys and 52% for girls, virtually eliminating the gender disparity in the treatment villages. This result is particularly noteworthy as cultural norms in Afghanistan do not permit girls to travel long distances alone.

Finally, there's a notable study conducted in Brazil by Moreno-Monroy, Lovelace e Ramos (2018). In 2016, the government of the state of São Paulo intended to close selected secondary schools as part of a plan to reduce the budgetary deficit. This policy was expected to impact over 300,000 students, with many being relocated to schools far from their homes. After public opposition, the decision was not implemented. Moreno-Monroy, Lovelace e Ramos (2018) conducted a study in response to this event, using simulation to estimate the effect of centralizing public secondary schools in the Metropolitan Region of São Paulo. The authors argued that given the significant spatial disparities in public transport and schooling provision, it is crucial to consider whether providing public transport subsidies to students can offset the limited availability of public schools in certain areas. The estimations indicated that the implementation of such a policy would result in longer commuting times for students in areas with low accessibility, leading to adverse consequences in terms of attendance rates, increased likelihood of dropout, and academic performance. As young individuals residing in areas with limited accessibility would be the most affected, this policy would exacerbate issues related to inadequate local schooling, restricted public transport access, and intense competition for enrollment in high-quality schools.

Unfortunately, many studies either do not examine the direct relationship between commute distance and school attendance, or rely on simple correlation analysis without establishing causality. There is a notable scarcity of studies employing causal analysis in the context of developing countries, and to the best of our knowledge, none have been conducted for urbanized areas or specifically for Brazil. The purpose of our study is to bridge this gap by addressing these research deficiencies.

3.3 THEORETICAL FRAMEWORK

In this section, we present a simple theoretical model that describes how absences can be related to commute time. We follow the modelling presented by Kobus, Ommeren e Rietveld (2015) to a certain extent.

Suppose that a student maximize utility from school gains, P, and leisure time at home, H_l. School gains is positively related to time spent at school, which is measure by days present at school, D_u, and daily number of hours present, H_u. School gains is also positively related to hours studying at home, given by H_h. The student spends time commuting to school, given by t; and time is the only cost associated with traveling to school. Therefore, the student maximizes the function $U(P,H_l) = U(D_u,H_h,H_l)$, and his utility is subject to a time constraint $D_u(H_u + t) + H_h + H_l = M$, where M denotes the student's time budget.

We assume a Cobb-Douglas utility function given by $P = D_u^{\alpha} H_u^{\beta} H_h^{\gamma} \rightarrow PH_l^{\delta} = D_u^{\alpha} H_u^{\beta} H_h^{\gamma} H_l^{\delta}$, in which $\alpha, \beta, \gamma, \delta > 0$. It is easy to obtain that $\alpha > \beta$ is a necessary condition for the maximization problem. Therefore, the number of days elasticity of school gains is larger than the number of hours elasticity of school gains. This implies that hours of study face stronger diminishing returns than days going to school, which makes sense as students should get tired after several hours of study.

It can be shown that:

$D_u^* = \frac{M(\alpha-\beta)}{t(\alpha+\gamma+\delta)}$.

Recalling that $\alpha > \beta$, it follows that $\frac{\sigma D_u^*}{\sigma t} < 0$. This implies that, for any student, longer commute times is associated with fewer days present at school.

This simple framework illustrated our main hypothesis. We expect to find a positive relationship between commute times and school absenteeism.

3.4 EMPIRICAL STRATEGY AND DATA

3.4.1 Specification

In this section, we propose an empirical model to measure the impact of accessibility on school absenteeism. We consider the following model:

$$Absences = f(Commuting + A + T + E) \tag{3.1}$$

In which the dependent variable, *Absences*, is the total of absences of student i, who is part of the j class of the k school. The main regressors is the commute distance of the student i to the school k, given by *Commuting*. This variable was assessed using the Google API, which provides the shortest route by foot[3] taken by the student when traveling to school.

A is a vector of twenty control variables related to the students and the student's family. In T, we included three variables specifically related to the student's class. E is a vector of school-related regressors. All variables mentioned are listed in Table 13.

In this study, we employ a Poisson model estimation for our main regression analysis. The choice of the Poisson model is motivated by the nature of the outcome variable, which represents count data in the form of the number of absences. Poisson regression is specifically designed to handle count data and is well-suited for situations where the outcome variable exhibits a skewed distribution with non-negative integer values. By using the Poisson model, we can account for the inherent characteristics of count data and appropriately model the relationship between commuting and school absences.

We will also estimate this model using OLS. For this case, we will use the following specification:

$$Absences_{ijk} = \alpha + \beta Commuting_i + \Gamma A_i + \Psi T_j + \Lambda E_k + \varepsilon_{ijk} \tag{3.2}$$

In which i is the index of students, j the index of schools, and k the index of classes.

[3] 77% of students in the dataset commute to school on foot.

Table 13 – Control variables

Category	Variable	Mean	Standard Deviation
Students	Age	12.547	0.909
	Race (white)	0.187	0.390
	Gender (male)	0.506	0.500
	Has already failed a year	61.211	11.301
	Number of school days until the interview date	0.278	0.448
	Household has a computer with internet access	0.265	0.442
	Goes to school by bicycle	0.034	0.182
	Goes to school by car	0.024	0.153
Family	Considers neighborhood to be violent	0.433	0.496
	Parents supervise homework	0.940	0.238
	Mother is the main parent	0.807	0.394
	Main parent works	0.390	0.488
	Main parent concluded primary school	0.878	0.328
	Receives federal government benefit (Bolsa Família)	0.438	0.496
	Chose the school because of its quality	0.138	0.345
	Chose the school that had a vacancy	0.065	0.247
	Chose the school based on the previous school's referral	0.296	0.457
	Chose to live near school	0.024	0.154
Class	Class was interrupted this year	0.244	0.430
	High rate of absenteeism among teachers in the class	0.187	0.390
School	School offers extra-class activities	0.758	0.429
	School offers full-day classes	0.906	0.292
	Existence of a failure rate reduction program	0.892	0.310

3.4.2 Endogeneity

The standard Poisson and OLS estimations for the model above may suffer from problems related to endogeneity, as omitted variable bias. To ensure we are capturing a causal effect, we employ a similar approach to the one used by Kobus, Ommeren e Rietveld (2015) and Tigre, Sampaio e Menezes (2017). We utilize the average distance between students and the two nearest schools as an instrumental variable for the commuting distance.

The city government of Recife instructs parents to enroll their children in the school closest to their home. As a result, the average distance between the two closest schools to the residence should be a good predictor of the actual commuting distance of the student. At the same time, this variable should not affect the number of absences of students throughout the year, except for its effect on commuting. If these assumptions hold, we have a valid instrument.

However, two objections can be raised against this instrument. First, parents who are more concerned about their children's education may choose to live closer to better schools and also make an effort to ensure that their children do not miss classes. In this case, parental concern might be related both to the instrument and the endogenous variable, rendering the instrument invalid. Second, there may be heterogeneities regarding the family's socioeconomic status that affect both the locational choice of housing and school absenteeism. An example is the Bolsa Família, a federal government benefit that provides income transfers to low-income families, conditioned on the children regularly attending school. In this case, the family's socioeconomic status can determine the proximity to public schools while also influencing school absenteeism, precluding our causal identification.

To address the first objection, we include several control variables that control for parents' concern:

1. Parents supervise homework?
2. How was the child's school chosen?
3. What was the strategy adopted to get a quality school place?
4. Would you like the child to study at a different public school?
5. Why do you live in this residence?

The first four variables can identify those parents most interested in the quality of the student's education. The last variable also fulfills this function, since one of the answer options is "Close to school". In addition, we perform three robustness tests in which we use specific subsamples that remove from the sample childrens' parents that showed concern about school quality and distance.

To address the second objection, we included several variables that control for socioeconomic status: if the parent works, parent's education, and if Household has a computer with internet access. We also show that our instrumental variable is balanced regarding to socioeconomic characteristics of the family. We first correlate each decile of the IV with the following variables: parent's education, log(income), a binary indicator for Bolsa Família, and a binary indicator for formal employment. The correlations are shown in Table 14. We also run a regression of each of the variables mentioned above on the IV, and the results are shown in Table 15. The results show that the IV is not correlated with the socioeconomic characteristics of the family.

Table 14 – Correlation between IV and socioeconomic variables

	Parent's education	Log(Income)	Bolsa Família	Formal Employment
Percentile 10	0.0807	-0.0457	0.0588	0.1220
Percentile 20	0.0755	-0.0701	-0.0117	-0.0687
Percentile 30	-0.0388	-0.1137	0.0118	0.0196
Percentile 40	0.0033	-0.0429	-0.0558	0.0672
Percentile 50	-0.0554	0.0617	-0.0570	0.0977
Percentile 60	0.0239	-0.1224	0.0400	-0.0019
Percentile 70	-0.1275	0.0858	-0.0169	-0.1447
Percentile 80	-0.0438	0.0234	0.0285	-0.1794
Percentile 90	-0.1300	-0.0934	0.1174	-0.0635
Percentile 100	0.0456	-0.1193	0.0438	0.0669

Table 15 – Regression of socioeconomic variables on IV

	(1)	(2)	(3)	(4)
	Log(IV)	Log(IV)	Log(IV)	Log(IV)
Parent's education	0.00423			
	(0.00365)			
Log(Income)		-0.00576		
		(0.00380)		
Bolsa Família			-0.0248	
			(0.0270)	
Formal Employment				0.0596
				(0.0372)
N	2005	2008	2008	2008
r2	0.000681	0.000973	0.000421	0.00149

* $p<0.05$, ** $p<0.01$, *** $p<0.001$. Robust Standard errors in parentheses.

Once we control for parents' concern about the student education and socioeconomic conditions of the family, the remaning effect of the instrumental variable on school absences should be only through commuting.

3.4.3 Data

This study will use survey data provided by The Joaquim Nabuco Institute for Social Research (Fundaj), a local agency linked to the Brazilian Ministry of Education and Culture. The survey was conducted in 2017 and again in 2018 and it consisted of standardized Math and Portuguese tests, along with a series of socioeconomic questions, both objective and subjective. Interviewers collected a wide range of information on students and their school life through four questionnaires: one for students, one for the primary adult responsible for the child's academic life, one for the teachers and one for the school principal (RAPOSO et al., 2019). The questionnaire aimed at public school students ages 11 to 13. The survey was conducted in Recife, the wealthiest city in the North-Northeast region of Brazil, a relevant developing country.

In this study, we will focus our analysis on the survey data for the year 2018. The dataset is comprised of 2008 students and parents, and 87 principals and schools.

3.5 RESULTS

3.5.1 Main results

Table 16 – Poisson regression analysis

	Poisson	IV Poisson GMM	IV Poisson cf
	(1)	(2)	(3)
	Absences	Absences	Absences
Absences			
Log(API distance)	0.0698*	0.188**	0.174**
	(0.0396)	(0.0943)	(0.0847)
Observations	1786	1786	1786
χ^2	35.51		

* p<0.10, ** p<0.05, *** p<0.01. Robust standard errors in parentheses.

In this section, we present the results and discussion of our analysis, focusing on the impact of travel time to school on absenteeism. We employed various regression models to explore this relationship, accounting for potential endogeneity concerns.

Estimation results are presented in Table 16. Column 1, which utilized a Poisson model, revealed a significant positive coefficient of 0.0698 for the logarithm of distance. This suggests that an increase in travel distance of 10% is associated with an approximately 7% more absences among students. However, given the possibility of endogeneity, we proceeded to estimate instrumental variable models. At this point we introduce our instrumental variable, the average distance to the two closest schools from the student's residence. In Column 2 we present our main specification and estimation model: an IV Poisson using the generalized method of moments (GMM). The GMM allows for a flexible functional form and does not require explicit specification of the control equation. It can also handle both linear and non-linear relationships between the endogenous variables and the instruments. The coefficient for Log(API distance) increased to 0.188 and remained statistically significant at the 5% level. This strengthened effect reinforces the substantial impact of longer commute times on absences. The result indicates that doubling the travel distance to school could lead to a 18.8% increase in absences.[4] To further validate our findings, we employed an IV Poisson model with a control function in Column 3. The results were consistent with the model in Column 2, showing a coefficient of 0.174, which was also statistically significant at the 5% level. This result provides additional support for the robustness of our findings, as the relationship between travel distance and absences holds even when using a different estimation technique.

Our results demonstrate a significant positive association between travel distance to school and absenteeism. Longer commute times are linked to a higher number of absences among students, indicating that reducing travel distance can potentially improve school attendance rates. These findings hold important implications for policymakers and educators, particularly in densely populated urban areas facing transportation challenges, similar to Recife.

Effective public policies should prioritize initiatives aimed at reducing commute time. Improving transportation infrastructure, optimizing school locations, and reinforcing school choice mechanisms that consider proximity to students' residences are potential strategies to consider. By addressing the challenges associated with long travel times, policymakers can mitigate the adverse effects on students' attendance and educational outcomes.

[4] The average distance to school is 2,031.368 meters and the average number of absences is 4.049 days.

3.5.2 Robustness tests

3.5.2.1 Alternative methods

Alternative methods can offer valuable insights and provide a comprehensive evaluation of the research findings. The main method employed in this study is IV Poisson, which it effective in addressing endogeneity issues and accommodating count data. However, to ensure the robustness of the results, it is important to consider alternative methods. Results are presented in Table 17.

Ordinary Least Squares is the default model for regression analysis, making it a suitable benchmark against which to compare the IV Poisson results (column 1). 2SLS (Two-Stage Least Squares) extends OLS by incorporating instrumental variables to address endogeneity (column 2). Tobit ML (Maximum Likelihood) is specifically designed for censored dependent variables, making it relevant when dealing with truncated data (column 3). Lastly, Tobit 2-step combines the strengths of Tobit models and instrumental variable techniques to account for potential endogeneity and censoring simultaneously (column 4). Notably, Tobit models assume a continuous dependent variable, which may not be suitable for count data, as is the case in our study. However, the obtained results are still useful for comparative analysis. Additionally, it is worth mentioning that a log transformation was applied to the explained variables in Table 17 to ensure comparability with the Poisson estimations.

The results show that the main findings (Table 16) are robust to alternative methods.

Table 17 – Alternative regression methods

	OLS	2SLS	Tobit ML	Tobit 2-step
	(1)	(2)	(3)	(4)
	Log(Absences)	Log(Absences)	Log(Absences)	Log(Absences)
main				
Log(API distance)	0.0624***	0.125**	0.225***	0.225***
	(0.0215)	(0.0520)	(0.0773)	(0.0756)
Observations	1786	1786	1786	1786
R^2	0.0307	0.0262		
First stage F		394.9		
Anderson-Rubin p-value		0.0160		

3.5.2.2 Alternative regressors

Table 18 displays the results of the robustness tests conducted using alternative regressors in the IV Poisson framework. The four columns represent different specifications of the IV Poisson model (1)-(4). Column 1 shows our prefered especification: commuting is measured as the walking distance from the student's home to school obtained via Google API using IV Poisson with control function (also shown in Column 2 of Table 16). In a alternative specification, we used the number of walking minutes, also obtained via Google API (Column 2). In Column 3, we used the number of minutes to commute from home to school as reported by the student's parent. Finally, in Column 4, we used the Euclidean distance between the school and the student's home.

The estimations show that the main findings are robust to alternative specifications of the commuting variable, as shown by the similar magnitudes and statistical significance in columns 2 and 4. The specification that utilizes the parent's self-reported commuting time (column 2) is the one that diverges the most, likely due to the subjective nature of self-reported data, which makes it more susceptible to measurement errors. Nonetheless, the result remains positive and significant.

Table 18 – Alternative regressor variables

	IV Poisson GMM (1) Absences	IV Poisson GMM (2) Absences	IV Poisson GMM (3) Absences	IV Poisson GMM (4) Absences
Absences				
Log(API distance)	0.188**			
	(0.0943)			
Log(API time)		0.198**		
		(0.0997)		
Log(guardian time)			0.410*	
			(0.212)	
Log(euclidian distance)				0.167*
				(0.0894)
Observations	1786	1786	1756	1737

* p<0.10, ** p<0.05, *** p<0.01. Robust standard errors in parentheses.

3.5.2.3 Alternative samples

Even though we included several control variables and estimated the regression via two-staged least squared to obtain precise and causal estimations, it is also important to perform robustness tests to assess the reliability of our findings. Thus, after the obtaining the main results, we repeated the regression analysis using specific subsamples (Table 19).

First, we will redo the analysis excluding those who answered that they chose the school based on quality (column 1). If the parent is more concerned about the child's education than the average parent, this preoccupation will affect both the travel time to school and absences. By removing those concerned parents and obtaining similar results, we can make a stronger case that we are capturing a causal link between commute time and school absenteeism.

Second, we also filter our sample to keep only those that affirmed that the school choice was either "referral from previous school" or "where there was vacancy" (column 2).

Third, we will remove those who affirmed that lived where they live for the reason of being near school. We strongly believe that this reason is confounding variables that may influence students' school attendance. By removing it, we can further confirm that our findings were not guided by other reasons other than commute time (column 3).

Table 19 – Alternative samples

	IV Poisson GMM	IV Poisson GMM	IV Poisson GMM
	(1)	(2)	(3)
	Absences	Absences	Absences
Absences			
Log(API distance)	0.226**	0.513***	0.182*
	(0.104)	(0.199)	(0.0945)
Observations	1550	664	1739

* p<0.10, ** p<0.05, *** p<0.01. Robust standard errors in parentheses.

3.6 CONCLUSIONS

Overall, the regression results provide evidence of a significant positive relationship between travel time to school and absenteeism. The coefficients consistently suggest that longer commute times are associated with a higher number of absences among students. These findings support the hypothesis that commute time plays a crucial role in students' attendance and highlight the importance of reducing travel time to promote better school attendance rates.

It is worth noting that the results are based on a dataset comprising schools in Recife, a densely populated Brazilian city facing transportation challenges. Therefore, the findings may have particular relevance for urban areas with similar characteristics. The robustness of the results is further supported by the statistical significance and consistency across different models.

The regression analysis provides important insights for policymakers and educators. The results suggest that public policies aimed at reducing commute time to school may have a positive impact on addressing absenteeism issues. By improving transportation infrastructure, optimizing school locations, and reinforcing school choice mechanisms that consider proximity to students' residences, it may be possible to mitigate the adverse effects of long travel times on students' attendance.

However, it is important to acknowledge the limitations of the study. The analysis is based on observational data, and although instrumental variable techniques help address endogeneity concerns, there may still be unobserved factors that influence both travel time and absences.

In conclusion, the results indicate a significant relationship between travel time to school and absenteeism in Recife, Brazil. The findings underscore the need for targeted interventions and policies to reduce commute time and improve school attendance. Further research exploring similar relationships in different settings would contribute to a more comprehensive understanding of the impact of commute time on students' educational experiences.

REFERÊNCIAS

ADDA, J. Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data*. *The Quarterly Journal of Economics*, v. 131, n. 2, p. 891–941, maio 2016. ISSN 0033-5533, 1531-4650. Disponível em: <https://academic.oup.com/qje/article/131/2/891/2607116>.

ALONSO, W. *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press, 1964. ISBN 978-0-674-73085-4. Disponível em: <https://hup.degruyter.com/view/title/320970>.

ANSARI, A.; PIANTA, R. C. School absenteeism in the first decade of education and outcomes in adolescence. *Journal of School Psychology*, v. 76, p. 48–61, out. 2019. ISSN 00224405. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S002244051930055X>.

ARAYA, F.; DELL, R.; DONOSO, P.; MARIANOV, V.; MARTÍNEZ, F.; WEINTRAUB, A. Optimizing location and size of rural schools in Chile: F. Araya et al. *International Transactions in Operational Research*, v. 19, n. 5, p. 695–710, set. 2012. ISSN 09696016. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/j.1475-3995.2012.00843.x>.

ARMAS, J. D.; RAMALHINHO, H.; REYNAL-QUEROL, M. Improving the accessibility to public schools in urban areas of developing countries through a location model and an analytical framework. *PLOS ONE*, v. 17, n. 1, p. e0262520, jan. 2022. ISSN 1932-6203. Disponível em: <https://dx.plos.org/10.1371/journal.pone.0262520>.

ASAHI, K. The Impact of Better School Accessibility on Student Outcomes. *SERC DISCUSSION PAPER*, 2014.

BANERJEE, A.; DUFLO, E.; QIAN, N. On the Road: Access to Transportation Infrastructure and Economic Growth in China. p. 51, 2012.

BANERJEE, A. V.; DUFLO, E. The Economic Lives of the Poor. *Journal of Economic Perspectives*, v. 21, n. 1, p. 141–167, jan. 2007. ISSN 0895-3309. Disponível em: <https://pubs.aeaweb.org/doi/10.1257/jep.21.1.141>.

BASTIAANSSEN, J.; JOHNSON, D.; LUCAS, K. Does better job accessibility help people gain employment? The role of public transport in Great Britain. *Urban Studies*, v. 59, n. 2, p. 301–322, fev. 2022. ISSN 0042-0980, 1360-063X. Disponível em: <http://journals.sagepub.com/doi/10.1177/00420980211012635>.

BAUM-SNOW, N. Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, v. 122, n. 2, p. 775–805, 2007. Disponível em: <http://www.jstor.org/stable/25098858>.

BAUM-SNOW, N.; FERREIRA, F. Causal Inference in Urban and Regional Economics. In: *Handbook of Regional and Urban Economics*. Elsevier, 2015. v. 5, p. 3–68. ISBN 978-0-444-59533-1. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/B9780444595171000015>.

BAUM-SNOW, N.; GLAESER, E. L.; ROSENTHAL, S. S. The Spread and Consequences of COVID-19 for Cities: An Introduction. *Journal of Urban Economics*, v. 127, p. 103428, jan. 2022. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119022000055>.

BEGHIN, N. Notes on inequality and poverty in Brazil: current situation and challenges. *From Poverty to Power: how active citizens and effective states can change the world*, 2008.

BERGANTINO, A. S.; MADIO, L. Travel-to-work. Which factors matter? An analysis on regional labor markets in the Uk. *Working papers SIET*, 2016.

BJARNASON, T. The effects of road infrastructure improvement on work travel in Northern Iceland. *Journal of Transport Geography*, v. 41, p. 229–238, dez. 2014. ISSN 09666923. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0966692314002142>.

BOUSSAUW, K.; NEUTENS, T.; WITLOX, F. Relationship between Spatial Proximity and Travel-to-Work Distance: The Effect of the Compact City. *Regional Studies*, v. 46, n. 6, p. 687–706, jun. 2012. ISSN 0034-3404, 1360-0591. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/00343404.2010.522986>.

BRITO, F.; SOUZA, J. d. Expansão urbana nas grandes metrópoles: o significado das migrações intrametropolitanas e da mobilidade pendular na reprodução da pobreza. *São Paulo em Perspectiva*, v. 19, n. 4, p. 48–63, dez. 2005. ISSN 0102-8839. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-88392005000400003&lng=pt&tlng=pt>.

BURDE, D.; LINDEN, L. L. Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools. *American Economic Journal: Applied Economics*, v. 5, n. 3, p. 27–40, jul. 2013. ISSN 1945-7782, 1945-7790. Disponível em: <https://pubs.aeaweb.org/doi/10.1257/app.5.3.27>.

CANDIDO, D. D. S.; WATTS, A.; ABADE, L.; KRAEMER, M. U. G.; PYBUS, O. G.; CRODA, J.; OLIVEIRA, W. D.; KHAN, K.; SABINO, E. C.; FARIA, N. R. Routes for COVID-19 importation in Brazil. *Journal of Travel Medicine*, v. 27, n. 3, p. taaa042, maio 2020. ISSN 1195-1982, 1708-8305. Disponível em: <https://academic.oup.com/jtm/article/doi/10.1093/jtm/taaa042/5809508>.

CARAZZA, L. E. B. *Three essays on urban economics: evidences from Brazil*. Tese (doctoralThesis), out. 2016. Accepted: 2017-04-27T15:00:47Z Publisher: Universidade Federal de Pernambuco. Disponível em: <https://repositorio.ufpe.br/handle/123456789/18669>.

CARMO, R. F.; NUNES, B. E. B. R.; MACHADO, M. F.; ARMSTRONG, A. C.; SOUZA, C. D. F. Expansion of COVID-19 within Brazil: the importance of highways. *Journal of Travel Medicine*, v. 27, n. 5, p. taaa106, ago. 2020. ISSN 1195-1982, 1708-8305. Disponível em: <https://academic.oup.com/jtm/article/doi/10.1093/jtm/taaa106/5863927>.

CASSEL, S. H.; MACUCHOVA, Z.; RUDHOLM, N.; RYDELL, A. Willingness to commute long distance among job seekers in Dalarna, Sweden. *Journal of Transport Geography*, v. 28, p. 49–55, abr. 2013. ISSN 09666923. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0966692312002657>.

CHAUVIN, J. P. *Why Does COVID-19 Affect Some Cities More than Others?: Evidence from the First Year of the Pandemic in Brazil*. [S.l.], 2021. Disponível em: <https://publications.iadb.org/en/node/30766>.

CHAUVIN, J. P.; GLAESER, E.; MA, Y.; TOBIO, K. What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics*, v. 98, p. 17–49, mar. 2017. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119016300067>.

CHEIN, F.; PINTO, C. C. d. X. Roads, Urbanization and Development: some empirical evidences from Brazilian experience. In: *Urbanization and Development*. [S.l.]: PAA, 2016. p. 48.

CHEN, H.; VOIGT, S.; FU, X. Data-Driven Analysis on Inter-City Commuting Decisions in Germany. *Sustainability*, v. 13, n. 11, p. 6320, jun. 2021. ISSN 2071-1050. Disponível em: <https://www.mdpi.com/2071-1050/13/11/6320>.

CNT. *Pesquisa CNT indica piora da qualidade das rodovias*. 2017.

CNT. *Transporte rodoviário: desempenho do setor, infraestrutura e investimentos*. 2017.

COELHO, M. T. P.; RODRIGUES, J. F. M.; MEDINA, A. M.; SCALCO, P.; TERRIBILE, L. C.; VILELA, B.; DINIZ-FILHO, J. A. F.; DOBROVOLSKI, R. Global expansion of COVID-19 pandemic is driven by population size and airport connections. *PeerJ*, v. 8, p. e9708, ago. 2020. ISSN 2167-8359. Disponível em: <https://peerj.com/articles/9708>.

COOK, R. L.; ROYCE, R. A.; THOMAS, J. C.; HANUSA, B. H. What's driving an epidemic? The spread of syphilis along an interstate highway in rural North Carolina. *American Journal of Public Health*, v. 89, n. 3, p. 369–373, mar. 1999. ISSN 0090-0036, 1541-0048. Disponível em: <https://ajph.aphapublications.org/doi/full/10.2105/AJPH.89.3.369>.

COOMBES, M.; CASADO-DÍAZ, J. M. The evolution of Local Labour Market Areas in contrasting region. p. 25, 2005.

CORNELISSEN, T.; DUSTMANN, C. Early School Exposure, Test Scores, and Noncognitive Outcomes. *American Economic Journal: Economic Policy*, v. 11, n. 2, p. 35–63, maio 2019. ISSN 1945-7731, 1945-774X. Disponível em: <https://pubs.aeaweb.org/doi/10.1257/pol.20170641>.

COSTA, C.; HA, J.; LEE, S. Spatial disparity of income-weighted accessibility in Brazilian Cities: Application of a Google Maps API. *Journal of Transport Geography*, v. 90, p. 102905, jan. 2021. ISSN 09666923. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0966692320309820>.

COTA, W. *Monitoring the number of COVID-19 cases and deaths in Brazil at municipal and federative units level*. [S.l.], 2020. Disponível em: <https://preprints.scielo.org/index.php/scielo/preprint/view/362/version/371>.

DESMET, K.; WACZIARG, R. JUE Insight: Understanding spatial variation in COVID-19 across the United States. *Journal of Urban Economics*, v. 127, p. 103332, jan. 2022. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119021000140>.

DEXTER, R. L. An investigation of the causes of absenteeism in an inner-city school. 1982.

DUARTE, L. B. ACESSIBILIDADE AO EMPREGO E RESULTADOS NO MERCADO DE TRABALHO. *UFPE*, p. 161, 2020.

DURANTON, G.; TURNER, M. A. Urban Growth and Transportation. *The Review of Economic Studies*, v. 79, n. 4, p. 1407–1440, out. 2012. ISSN 1467-937X, 0034-6527. Disponível em: <https://academic.oup.com/restud/article-lookup/doi/10.1093/restud/rds010>.

DUZE, C. Average Distance Travelled To School by Primary and Secondary School Students in Nigeria and Its Effect on Attendance. *African Research Review*, v. 4, n. 4, set. 2011. ISSN 2070-0083, 1994-9057. Disponível em: <http://www.ajol.info/index.php/afrrev/article/view/69236>.

EATON, D. K.; BRENER, N.; KANN, L. K. Associations of Health Risk Behaviors With School Absenteeism. Does Having Permission for the Absence Make a Difference? *Journal of School Health*, v. 78, n. 4, p. 223–229, abr. 2008. ISSN 0022-4391, 1746-1561. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/j.1746-1561.2008.00290.x>.

FABER, B. Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies*, v. 81, n. 3, p. 1046–1070, jul. 2014. ISSN 1467-937X, 0034-6527. Disponível em: <https://academic.oup.com/restud/article-lookup/doi/10.1093/restud/rdu010>.

FANG, L.-Q.; WANG, L.-P.; VLAS, S. J. D.; LIANG, S.; TONG, S.-L.; LI, Y.-L.; LI, Y.-P.; QIAN, Q.; YANG, H.; ZHOU, M.-G.; WANG, X.-F.; RICHARDUS, J. H.; MA, J.-Q.; CAO, W.-C. Distribution and Risk Factors of 2009 Pandemic Influenza A (H1N1) in Mainland China. *American Journal of Epidemiology*, v. 175, n. 9, p. 890–897, maio 2012. ISSN 1476-6256, 0002-9262. Disponível em: <https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwr411>.

FERNÁNDEZ-MALDONADO, A. M.; ROMEIN, A.; VERKOREN, O.; PESSOA, R. P. P. Polycentric Structures in Latin American Metropolitan Areas: Identifying Employment Sub-centres. *Regional Studies*, v. 48, n. 12, p. 1954–1971, dez. 2014. ISSN 0034-3404, 1360-0591. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/00343404.2013.786827>.

FINN, J. D. Withdrawing from School. *Review of Educational Research*, p. 26, 1989.

FUHS, M. W.; NESBITT, K. T.; JACKSON, H. Chronic Absenteeism and Preschool Children's Executive Functioning Skills Development. *Journal of Education for Students Placed at Risk (JESPAR)*, v. 23, n. 1-2, p. 39–52, abr. 2018. ISSN 1082-4669, 1532-7671. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/10824669.2018.1438201>.

GARCÍA, E.; WEISS, E. Student Absenteeism: Who Misses School and How Missing School Matters for Performance. *Economic Policy Institute*, 2018.

GARCIA-LÓPEZ, M.-À.; HÉMET, C.; VILADECANS-MARSAL, E. Next train to the polycentric city: The effect of railroads on subcenter formation. *Regional Science and Urban Economics*, v. 67, p. 50–63, nov. 2017. ISSN 01660462. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0166046216303271>.

GERSHENSON, S.; JACKNOWITZ, A.; BRANNEGAN, A. Are Student Absences Worth the Worry in U.S. Primary Schools? *Education Finance and Policy*, v. 12, n. 2, p. 137–165, dez. 2015. ISSN 1557-3060, 1557-3079. Disponível em: <http://www.mitpressjournals.org/doi/10.1162/EDFP_a_00207>.

GIBBONS, S.; LYYTIKÄINEN, T.; OVERMAN, H. G.; SANCHIS-GUARNER, R. New road infrastructure: The effects on firms. *Journal of Urban Economics*, v. 110, p. 35–50, mar. 2019. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119019300105>.

GLAESER, E. L.; GORBACK, C.; REDDING, S. J. JUE Insight: How much does COVID-19 increase with mobility? Evidence from New York and four other U.S. cities. *Journal of Urban Economics*, v. 127, p. 103292, jan. 2022. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119020300632>.

GOODMAN, J. *Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time*. Cambridge, MA, 2014. w20221 p. Disponível em: <http://www.nber.org/papers/w20221.pdf>.

GOTTFRIED, M. A. Excused Versus Unexcused: How Student Absences in Elementary School Affect Academic Achievement. *Educational Evaluation and Policy Analysis*, v. 31, n. 4, p. 392–415, dez. 2009. ISSN 0162-3737, 1935-1062. Disponível em: <http://journals.sagepub.com/doi/10.3102/0162373709342467>.

GOTTFRIED, M. A. The Detrimental Effects of Missing School: Evidence from Urban Siblings. *American Journal of Education*, v. 117, n. 2, p. 147–182, fev. 2011. ISSN 0195-6744, 1549-6511. Disponível em: <https://www.journals.uchicago.edu/doi/10.1086/657886>.

GOTTFRIED, M. A. Chronic Absenteeism and Its Effects on Students' Academic and Socioemotional Outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, v. 19, n. 2, p. 53–75, abr. 2014. ISSN 1082-4669, 1532-7671. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/10824669.2014.962696>.

GOTTFRIED, M. A. Linking Getting to School With Going to School. *Educational Evaluation and Policy Analysis*, v. 39, n. 4, p. 571–592, dez. 2017. ISSN 0162-3737, 1935-1062. Disponível em: <http://journals.sagepub.com/doi/10.3102/0162373717699472>.

GOTTFRIED, M. A.; ANSARI, A. Detailing New Dangers: Linking Kindergarten Chronic Absenteeism to Long-Term Declines in Executive Functioning. *The Elementary School Journal*, v. 121, n. 3, p. 484–503, mar. 2021. ISSN 0013-5984, 1554-8279. Disponível em: <https://www.journals.uchicago.edu/doi/10.1086/712426>.

GOTTFRIED, M. A.; KIRKSEY, J. J. "When" Students Miss School: The Role of Timing of Absenteeism on Students' Test Performance. *Educational Researcher*, v. 46, n. 3, p. 119–130, abr. 2017. ISSN 0013-189X, 1935-102X. Disponível em: <http://journals.sagepub.com/doi/10.3102/0013189X17703945>.

GREEN, D. A.; MORISSETTE, R.; SAND, B. M.; SNODDY, I. Economy-Wide Spillovers from Booms: Long-Distance Commuting and the Spread of Wage Effects. *Journal of Labor Economics*, v. 37, n. S2, p. S643–S687, jul. 2019. ISSN 0734-306X, 1537-5307. Disponível em: <https://www.journals.uchicago.edu/doi/10.1086/703362>.

GUMIERO, R. G. Dimensões do desenvolvimento do complexo industrial Portuário de Suape : política econômica e política social. maio 2015. Accepted: 2016-10-04T18:58:38Z Publisher: Universidade Federal de São Carlos. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/7685>.

HADDAD, E. A.; BARUFI, A. M. B. From rivers to roads: Spatial mismatch and inequality of opportunity in urban labor markets of a megacity. *Habitat International*, v. 68, p. 3–14, out. 2017. ISSN 01973975. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0197397516307433>.

HALLFORS, D.; VEVEA, J. L.; IRITANI, B.; CHO, H.; KHATAPOUSH, S.; SAXE, L. Truancy, Grade Point Average, and Sexual Activity: A Meta-Analysis of Risk Indicators for Youth Substance Use. *Journal of School Health*, v. 72, n. 5, p. 205–211, maio 2002. ISSN 00224391, 17461561. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/j.1746-1561.2002.tb06548.x>.

HAM, M. v. *Job access, workplace mobility and occupational achievement*. Delft: Eburon, 2002. ISBN 978-90-5166-885-8.

HANDA, S. Raising primary school enrolment in developing countries The relative importance of supply and demand. *Journal of Development Economics*, 2002.

HARRIS, J. E. Critical Role of the Subways in the Initial Spread of SARS-CoV-2 in New York City. *Frontiers in Public Health*, v. 9, p. 754767, dez. 2021. ISSN 2296-2565. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpubh.2021.754767/full>.

HEIDRICH, Á. L. Território, integração socioespacial, região, fragmentação e exclusão social. *Território e desenvolvimento: diferentes abordagens. Francisco Beltrão: UNIOESTE*, p. 37–66, 2004.

HISSA-TEIXEIRA, K. Uma análise da estrutura espacial dos indicadores socioeconômicos do nordeste brasileiro (2000-2010). *EURE (Santiago)*, v. 44, n. 131, p. 101–124, jan. 2018. ISSN 0250-7161. Publisher: Pontificia Universidad Católica de Chile. Facultad de Arquitectura, Diseño y Estudios Urbanos. Instituto de Estudios Urbanos y Territoriales. Disponível em: <http://www.scielo.cl/scielo.php?script=sci_abstract&pid=S0250-71612018000100101&lng=es&nrm=iso&tlng=pt>.

HOLL, A. Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain. *Regional Science and Urban Economics*, v. 34, n. 3, p. 341–363, maio 2004. ISSN 01660462. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0166046203000590>.

HOLL, A. Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal*. *Journal of Regional Science*, v. 44, n. 4, p. 693–712, nov. 2004. ISSN 0022-4146, 1467-9787. Disponível em: <http://doi.wiley.com/10.1111/j.0022-4146.2004.00354.x>.

HORTENCIO, V. E. F.; MILANI, A. M. R. Celso Furtado, política e planejamento econômico: revisitando a luta para a superação do subdesenvolvimento do Nordeste. *REVISTA ECONOMIA POLÍTICA DO DESENVOLVIMENTO*, v. 9, n. 22, p. 69–85, set. 2019. ISSN 1984-0756, 1984-0756. Disponível em: <https://192.168.5.14/index.php/repd/article/view/8751>.

IBGE. *Censo 2010*. 2010.

IBGE. *Indicadores sociais municipais: uma análise dos resultados do universo do censo demográfico 2010*. 2010.

INGRAM, G. K.; CARROLL, A. The spatial structure of Latin American cities. *Journal of Urban Economics*, v. 9, n. 2, p. 257–273, mar. 1981. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/0094119081900449>.

JOHNSON, G. M. Student Alienation, Academic Achievement, and WebCT Use. *Educational Technology & Society*, p. 11, 2005.

KAZEEM, A.; JENSEN, L.; STOKES, C. S. School Attendance in Nigeria: Understanding the Impact and Intersection of Gender, Urban-Rural Residence, and Socioeconomic Status. *Comparative Education Review*, v. 54, n. 2, p. 295–319, maio 2010. ISSN 0010-4086, 1545-701X. Disponível em: <https://www.journals.uchicago.edu/doi/10.1086/652139>.

KERR, A. Tax(i)ing the Poor? Commuting Costs in South African Cities. *South African Journal of Economics*, v. 85, n. 3, p. 321–340, set. 2017. ISSN 00382280. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/saje.12161>.

KLEIN, M.; SOSU, E. M.; DARE, S. School Absenteeism and Academic Achievement: Does the Reason for Absence Matter? *AERA Open*, v. 8, p. 233285842110711, jan. 2022. ISSN 2332-8584, 2332-8584. Disponível em: <http://journals.sagepub.com/doi/10.1177/23328584211071115>.

KOBUS, M. B.; OMMEREN, J. N. V.; RIETVELD, P. Student commute time, university presence and academic achievement. *Regional Science and Urban Economics*, v. 52, p. 129–140, maio 2015. ISSN 01660462. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0166046215000216>.

KRUEGER, A. B. Economic Considerations and Class Size. *The Economic Journal*, v. 113, n. 485, p. F34–F63, fev. 2003. ISSN 0013-0133, 1468-0297. Disponível em: <https://academic.oup.com/ej/article/113/485/F34/5079544>.

LEHR, C. A.; SINCLAIR, M. F.; CHRISTENSON, S. L. Addressing Student Engagement and Truancy Prevention During the Elementary School Years: A Replication Study of the Check & Connect Model. *Journal of Education for Students Placed at Risk (JESPAR)*, v. 9, n. 3, p. 279–301, jul. 2004. ISSN 1082-4669, 1532-7671. Disponível em: <http://www.tandfonline.com/doi/abs/10.1207/s15327671espr0903_4>.

LI, B.; MA, L. JUE insight: Migration, transportation infrastructure, and the spatial transmission of COVID-19 in China. *Journal of Urban Economics*, v. 127, p. 103351, jan. 2022. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119021000334>.

LI, X.; CHEN, H.; SHI, Y.; SHI, F. Transportation Equity in China: Does Commuting Time Matter? *Sustainability*, v. 11, n. 21, p. 5884, out. 2019. ISSN 2071-1050. Disponível em: <https://www.mdpi.com/2071-1050/11/21/5884>.

LOUNSBURY, J. W.; STEEL, R. P.; LOVELAND, J. M.; GIBSON, L. W. An Investigation of Personality Traits in Relation to Adolescent School Absenteeism. *Journal of Youth and Adolescence*, v. 33, n. 5, p. 457–466, out. 2004. ISSN 0047-2891, 1573-6601. Disponível em: <http://link.springer.com/10.1023/B:JOYO.0000037637.20329.97>.

MANGRUM, D.; NIEKAMP, P. JUE Insight: College student travel contributed to local COVID-19 spread. *Journal of Urban Economics*, v. 127, p. 103311, jan. 2022. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119020300826>.

MATA, D. da; DEICHMANN, U.; HENDERSON, J.; LALL, S.; WANG, H. Determinants of city growth in Brazil. *Journal of Urban Economics*, v. 62, n. 2, p. 252–272, set. 2007. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119006000878>.

MCQUAID, R.; GREIG, M.; ADAMS, J. Unemployed Job Seeker Attitudes towards Potential Travel-to-Work Times. *Growth and Change*, v. 32, n. 3, p. 355–368, jan. 2001. ISSN 0017-4815, 1468-2257. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/0017-4815.00163>.

MILLS, E. S. An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, v. 57, n. 2, p. 197–210, 1967.

MITZE, T.; KOSFELD, R. The propagation effect of commuting to work in the spatial transmission of COVID-19. *Journal of Geographical Systems*, v. 24, n. 1, p. 5–31, jan. 2022. ISSN 1435-5930, 1435-5949. Disponível em: <https://link.springer.com/10.1007/s10109-021-00349-3>.

MONK, D. H.; IBRAHIM, M. A. Patterns of Absence and Pupil Achievement. *American Educational Research Journal*, v. 21, n. 2, p. 295–310, jan. 1984. ISSN 0002-8312, 1935-1011. Disponível em: <http://journals.sagepub.com/doi/10.3102/00028312021002295>.

MOREIRA, M. J. Tests with correct size when instruments can be arbitrarily weak. *Journal of Econometrics*, v. 152, n. 2, p. 131–140, out. 2009. ISSN 03044076. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0304407609000190>.

MORENO-MONROY, A. I.; LOVELACE, R.; RAMOS, F. R. Public transport and school location impacts on educational inequalities: Insights from São Paulo. *Journal of Transport Geography*, v. 67, p. 110–118, fev. 2018. ISSN 09666923. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0966692316303453>.

MORRISSEY, T. W.; HUTCHISON, L.; WINSLER, A. Family income, school attendance, and academic achievement in elementary school. *Developmental Psychology*, v. 50, n. 3, p. 741–753, mar. 2014. ISSN 1939-0599, 0012-1649. Disponível em: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0033848>.

MORTEN, M.; OLIVEIRA, J. The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City. p. 65, 2018.

MOURA, R.; BRANCO, M. L. G. C.; FIRKOWSKI, O. L. C. d. F. Movimento pendular e perspectivas de pesquisas em aglomerados urbanos. *São Paulo em Perspectiva*, v. 19, n. 4, p. 121–133, dez. 2005. ISSN 0102-8839. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-88392005000400008&lng=pt&tlng=pt>.

MOURA, R.; DELGADO, P.; COSTA, M. A. MOVIMENTO PENDULAR E POLÍTICAS PÚBLICAS: ALGUMAS POSSIBILIDADES INSPIRADAS NUMA TIPOLOGIA DOS MUNICÍPIOS BRASILEIROS. In: *Brasil em desenvolvimento 2013: estado, planejamento e políticas públicas*. Brasília: Instituto de Pesquisa Econômica Aplicada (IPEA), 2013. ISBN 978-85-7811-185-4 978-85-7811-184-7 978-85-7811-183-0. OCLC: 874844918.

MUTH, R. F. CITIES AND HOUSING; THE SPATIAL PATTERN OF URBAN RESIDENTIAL LAND USE. 1969. Disponível em: <https://trid.trb.org/view/545388>.

NARITOMI, J.; SOARES, R. R.; ASSUNÇÃO, J. J. Institutional Development and Colonial Heritage within Brazil. *The Journal of Economic History*, v. 72, n. 2, p. 393–422, maio 2012. ISSN 0022-0507, 1471-6372. Disponível em: <https://www.cambridge.org/core/product/identifier/S0022050712000071/type/journal_article>.

NATAL, J. L. A. *Transporte, ocupação do espaço e desenvolvimento capitalista no Brasil: historia e perspectivas*. Tese (Doutor em Economia) — Universidade Estadual de Campinas, Campinas, 1991. Disponível em: <http://acervus.unicamp.br/index.asp?codigo_sophia=36962>.

NAZARETH, P. B. d. Transporte e desenvolvimento urbano. *Revista de Administração Pública*, v. 12, n. 2, p. 229–288, 1978. Disponível em: <https://bibliotecadigital.fgv.br/ojs/index.php/rap/article/view/7461>.

NETO, R. D. M. S.; AZZONI, C. R. Non-Spatial Government Policies and Regional Income Inequality in Brazil. *Regional Studies*, v. 45, n. 4, p. 453–461, abr. 2011. ISSN 0034-3404. Publisher: Routledge _eprint: https://doi.org/10.1080/00343400903241485. Disponível em: <https://doi.org/10.1080/00343400903241485>.

NETO, R. S.; DUARTE, G.; PÁEZ, A. Gender and commuting time in São Paulo Metropolitan Region. *Urban Studies*, v. 52, n. 2, p. 298–313, fev. 2015. ISSN 0042-0980, 1360-063X. Disponível em: <http://journals.sagepub.com/doi/10.1177/0042098014528392>.

NETO, R. S.; MOURA, K. Commuting time, public exposure and victimization: Evidence from Brazilian metropolitan regions. *Papers in Regional Science*, v. 98, n. 2, p. 1159–1175, abr. 2019. ISSN 1056-8190, 1435-5957. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/pirs.12382>.

NEWMANN, F. Reducing Student Alienation in High Schools: Implications of Theory. *Harvard Educational Review*, v. 51, n. 4, p. 546–564, dez. 1981. ISSN 0017-8055, 1943-5045. Disponível em: <http://hepgjournals.org/doi/10.17763/haer.51.4.xj67887u87l5t66t>.

NICOLELIS, M. A. L.; RAIMUNDO, R. L. G.; PEIXOTO, P. S.; ANDREAZZI, C. S. The impact of super-spreader cities, highways, and intensive care availability in the early stages of the COVID-19 epidemic in Brazil. *Scientific Reports*, v. 11, n. 1, p. 13001, jun. 2021. ISSN 2045-2322. Disponível em: <https://www.nature.com/articles/s41598-021-92263-3>.

ÖHMAN, M.; LINDGREN, U. Who are the long-distance commuters? Patterns and driving forces in Sweden. *Cybergeo*, ago. 2003. ISSN 1278-3366. Disponível em: <http://journals.openedition.org/cybergeo/4118>.

OJIMA, R.; JR, E. M.; PEREIRA, R. H. M. O estigma de morar longe da cidade: repensando o consenso sobre as “cidades-dormitório” no Brasil. *São Paulo*, v. 12, n. 24, p. 23, 2010.

OLIVIERI, C.; FAGEDA, X. Urban mobility with a focus on gender: The case of a middle-income Latin American city. *Journal of Transport Geography*, v. 91, p. 102996, fev. 2021. ISSN 09666923. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0966692321000491>.

OMMEREN, J. N. van; PUIGARNAU, E. Gutiérrez-i. Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics*, v. 41, n. 1, p. 1–8, jan. 2011. ISSN 01660462. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0166046210000633>.

OSTER, E. ROUTES OF INFECTION: EXPORTS AND HIV INCIDENCE IN SUB-SAHARAN AFRICA. *Journal of the European Economic Association*, v. 10, n. 5, p. 1025–1058, out. 2012. ISSN 15424766. Disponível em: <https://academic.oup.com/jeea/article-lookup/doi/10.1111/j.1542-4774.2012.01075.x>.

PEREIRA, R. H. M.; SCHWANEN, T. Tempo de deslocamento casa-trabalho no Brasil (1992-2009): Diferenças entre regiões metropolitanas, níveis de renda e sexo. 2013.

PIZZOLATO, N.; SILVA, H. F. The Location of Public Schools: Evaluation of Practical Experiences. *International Transactions in Operational Research*, v. 4, n. 1, p. 13–22, jan. 1997. ISSN 0969-6016, 1475-3995. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/j.1475-3995.1997.tb00058.x>.

PONTILI, R. M.; KASSOUF, A. L. Fatores que afetam a freqüência e o atraso escolar, nos meios urbano e rural, de São Paulo e Pernambuco. *Revista de Economia e Sociologia Rural*, v. 45, n. 1, p. 27–47, mar. 2007. ISSN 0103-2003. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-20032007000100002&lng=pt&tlng=pt>.

RAPOSO, I. P. d. A.; MENEZES, T. A. d.; LIMA, R. C. d. A.; NÓBREGA, R. Z. da. Peer effects and scholastic achievement: spatial models estimates using the student friendship network at the classroom level. *Economia Aplicada*, v. 23, n. 1, p. 5–24, 2019.

REDDING, S. J.; TURNER, M. A. Transportation Costs and the Spatial Organization of Economic Activity. p. 58, 2014.

REIS, E.; PIMENTEL, M.; ALVARENGA, A. I.; SANTOS, M. Áreas mínimas comparáveis para os períodos intercensitários de 1872 a 2000. *Rio de Janeiro: Ipea/Dimac*, v. 40, 2008.

ROWE, F.; BELL, M. The drivers of long-distance commuting in Chile: The role of spatial distribution of economic activities. p. 31, 2018.

RUMBERGER, R. W. Dropping Out of Middle School: A Multilevel Analysis of Students and Schools. *American Educational Research Journal*, v. 32, n. 3, p. 583–625, set. 1995. ISSN 0002-8312, 1935-1011. Disponível em: <http://journals.sagepub.com/doi/10.3102/00028312032003583>.

RYAN, A. M.; PATRICK, H. The Classroom Social Environment and Changes in Adolescents' Motivation and Engagement During Middle School. *American Educational Research Journal*, v. 38, n. 2, p. 437–460, jun. 2001. ISSN 0002-8312, 1935-1011. Disponível em: <http://journals.sagepub.com/doi/10.3102/00028312038002437>.

SANDOW, E. Commuting behaviour in sparsely populated areas: evidence from northern Sweden. *Journal of Transport Geography*, v. 16, n. 1, p. 14–27, jan. 2008. ISSN 09666923. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0966692307000506>.

SANDOW, E. *On the road: social aspects of commuting long distances to work*. Umeå: Department of Social and Economic Geography, Umeå University, Sweden, 2011. (GERUM, 2011:2). ISBN 978-91-978344-6-9.

SANDOW, E.; WESTIN, K. Preferences for commuting in sparsely populated areas: The case of Sweden. *Journal of Transport and Land Use*, v. 2, p. 21, 2010.

SHEARMUR, R. Travel From Home: An Economic Geography of Commuting Distances in Montreal. *Urban Geography*, v. 27, n. 4, p. 330–359, jun. 2006. ISSN 0272-3638, 1938-2847. Disponível em: <http://www.tandfonline.com/doi/abs/10.2747/0272-3638.27.4.330>.

SILVA, A. C. J. da; BAGATTOLLI, C. Policy mix e governança multinível em políticas de desenvolvimento regional no Nordeste: o caso do setor automotivo em Pernambuco. *Redes. Revista do Desenvolvimento Regional*, v. 25, n. 3, p. 1212–1232, 2020. Publisher: Universidade de Santa Cruz do Sul.

SILVA, H. de; JOHNSON, L.; WADE, K. Long distance commuters in Australia: A socio-economic and demographic profile. p. 18, 2011.

SILVA, J. D. S.; FREITAS, C. O. de. Análise dos determinantes da migração pendular nas regiões metropolitanas do sudeste brasileiro. *Economia & Região*, v. 8, n. 2, p. 29, ago. 2020. ISSN 2317-627X. Disponível em: <http://www.uel.br/revistas/uel/index.php/ecoreg/article/view/37560>.

SOUCH, J. M.; COSSMAN, J. S.; HAYWARD, M. D. Interstates of Infection: Preliminary Investigations of Human Mobility Patterns in the COVID-19 Pandemic. *The Journal of Rural Health*, v. 37, n. 2, p. 266–271, mar. 2021. ISSN 0890-765X, 1748-0361. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/jrh.12558>.

SOUZA, A. L. D.; DUARTE, L. B.; NETO, R. D. M. S. Acessibilidade espacial ao emprego como determinante da participação, empregabilidade e informalidade: Um estudo empírico para a Região Metropolitana do Recife. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 16, n. 2, p. 184–213, abr. 2023. ISSN 2447-7990. Disponível em: <https://revistaaber.org.br/rberu/article/view/850>.

STEIN, M. L.; GRIGG, J. A. Missing Bus, Missing School: Establishing the Relationship Between Public Transit Use and Student Absenteeism. *American Educational Research Journal*, v. 56, n. 5, p. 1834–1860, out. 2019. ISSN 0002-8312, 1935-1011. Disponível em: <http://journals.sagepub.com/doi/10.3102/0002831219833917>.

STRANO, E.; VIANA, M. P.; SORICHETTA, A.; TATEM, A. J. Mapping road network communities for guiding disease surveillance and control strategies. *Scientific Reports*, v. 8, n. 1, p. 4744, mar. 2018. ISSN 2045-2322. Disponível em: <https://www.nature.com/articles/s41598-018-22969-4>.

SUMMERHILL, W. R. Big social savings in a small laggard economy: railroad-led growth in Brazil. *The Journal of Economic History*, v. 65, n. 1, p. 72–102, 2005. Publisher: Cambridge University Press.

SUMMERS, A. A.; WOLFE, B. L. Do Schools Make a Difference? *The American Economic Review*, v. 67, n. 4, p. 639–652, 1977. ISSN 00028282.

THEMBO, K. N. School accessibility and students' attendance in secondary Schools in Kasese district, Uganda. *Kyambogo University*, 2011.

TIGRE, R.; SAMPAIO, B.; MENEZES, T. THE IMPACT OF COMMUTING TIME ON YOUTH'S SCHOOL PERFORMANCE. *Journal of Regional Science*, v. 57, n. 1, p. 28–47, jan. 2017. ISSN 00224146. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1111/jors.12289>.

TURNER, T.; NIEMEIER, D. Travel to work and household responsibility: new evidence. *Transportation*, v. 24, p. 397–419, 1997. Publisher: Springer.

United Nations. *Foreign direct investment in latin america and the caribbean 2012*. Place of publication not identified: United Nations Pubns, 2013. OCLC: 858846020. ISBN 978-92-1-121834-3.

VANDYCK, T.; PROOST, S. Inefficiencies in regional commuting policy*: Inefficiencies in regional commuting policy. *Papers in Regional Science*, p. no–no, jun. 2012. ISSN 10568190. Disponível em: <http://doi.wiley.com/10.1111/j.1435-5957.2012.00437.x>.

VURI, D. The Effect of Availability of School and Distance to School on Children's Time Allocation in Ghana: The Effect of Availability of School and Distance to School. *LABOUR*, v. 24, p. 46–75, dez. 2010. ISSN 11217081. Disponível em: <http://doi.wiley.com/10.1111/j.1467-9914.2010.00499.x>.

WAN, L.; WAN, Q. High-speed railway and the intercity transmission of epidemics: Evidence from COVID-19 in China. *Economic Modelling*, v. 114, p. 105934, set. 2022. ISSN 02649993. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0264999322001808>.

WEST, M. R.; KRAFT, M. A.; FINN, A. S.; MARTIN, R. E.; DUCKWORTH, A. L.; GABRIELI, C. F. O.; GABRIELI, J. D. E. Promise and Paradox: Measuring Students' Non-Cognitive Skills and the Impact of Schooling. *Educational Evaluation and Policy Analysis*, v. 38, n. 1, p. 148–170, mar. 2016. ISSN 0162-3737, 1935-1062. Disponível em: <http://journals.sagepub.com/doi/10.3102/0162373715597298>.

WHEATON, W. C. Commuting, congestion, and employment dispersal in cities with mixed land use. *Journal of Urban Economics*, v. 55, n. 3, p. 417–438, maio 2004. ISSN 00941190. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0094119004000087>.

ZASLAVSKY, R.; GOULART, B. N. G. d. Migração pendular e atenção à saúde na região de fronteira. *Ciência & Saúde Coletiva*, v. 22, n. 12, p. 3981–3986, dez. 2017. ISSN 1678-4561, 1413-8123. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232017021203981&lng=pt&tlng=pt>.

ZENOU, Y. Urban unemployment, agglomeration and transportation policies. *Journal of Public Economics*, v. 77, n. 1, p. 97–133, jul. 2000. ISSN 00472727. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0047272799000833>.