CAF - WORKING PAPER #2022/11
First version: September 7, 2021 (current)

# Evaluating the Impact of Buenos Aires's *Metrobus* on Within-City Spatial Sorting

**Pablo Ernesto Warnes**[1]

[1]Assistant Professor, Department of Economics, Aalto University School of Business and Helsinki Graduate School of Economic.
pablo.warnes@aalto.fi

How do improvements in the urban transport infrastructure affect the spatial sorting of residents with different levels of income and education within a city? What are the welfare effects of improving urban transit once we take into account these patterns of spatial sorting? In this paper, I study the effects of the construction of a bus rapid transit system (BRT) on the spatial reorganization of residents within the city of Buenos Aires, Argentina. To do so, I leverage an individual level panel data set of more than two million residents with which I can describe intra-city migration patterns. I first find reduced form evidence that the construction of the BRT increased the spatial segregation between high and low-skilled residents within the city. I then develop a dynamic quantitative spatial equilibrium model of a city with heterogeneous workers that allows me to quantify the welfare effects of this BRT system while taking into account these spatial sorting patterns. With this quantitative framework, I can measure the average welfare gains for residents that were living near the BRT lines before these were built.

**KEYWORDS**
Transport infrastructure, BRT, spatial sorting, urban economics

CAF - DOCUMENTO DE TRABAJO #2022/11

Esta versión: 9 de junio de 2022

# Evaluando el impacto del Metrobus de Buenos Aires sobre la distribución espacial de los residentes al interior de la ciudad

**Pablo Ernesto Warnes**[1]

[1] Assistant Professor, Department of Economics, Aalto University School of Business and Helsinki Graduate School of Economic.
pablo.warnes@aalto.fi

¿Cómo afectan las mejoras en la infraestructura de transporte urbano a la distribución espacial de los residentes con diferentes niveles de ingresos y educación dentro de una ciudad? ¿Cuáles son los efectos sobre el bienestar de mejorar el tránsito urbano una vez que tomamos en cuenta estos patrones de distribución espacial? En este artículo se estudian los efectos de la construcción de un sistema de autobuses de tránsito rápido ( BRT, por sus siglas en inglés) en la reorganización espacial de los residentes dentro de la Ciudad de Buenos Aires, Argentina. Para hacerlo, se utiliza una base de datos de panel a nivel individual de más de dos millones de residentes. Estos datos permiten describir los patrones de migración dentro de la ciudad. En un primer paso, se encuentra evidencia de forma reducida de que la construcción del BRT aumentó la segregación espacial dentro de la ciudad entre los residentes de alta y baja calificación. Luego, se desarrolla un modelo espacial cuantitativo de equilibrio general dinámico de una ciudad con trabajadores heterogéneos. Este modelo permite cuantificar los efectos sobre el bienestar del BRT teniendo en cuenta los patrones de distribución espacial ya mencionados. A su vez, este marco cuantitativo permite medir las ganancias de bienestar promedio para los residentes por lugar de residencia previo a la construcción del BRT.

**KEYWORDS**

Infraestructura de transporte, BRT. distribución espacial, economía urbana

## 1 | INTRODUCTION

How do improvements in the urban transport infrastructure affect the spatial sorting of residents with different levels of income and education within the city? What are the welfare effects of these improvements in transport infrastructure once we take into account these patterns of spatial sorting? Given the world's rapid urban growth of the last decades,[1] it is no surprise that many cities (especially in the developing world) are investing large sums in improving their public transport infrastructure.[2] The literature on public transport improvements has shown that these projects tend to increase population density, land rents, and output in the neighborhoods that are close to the improved infrastructure.[3] However, one concern that arises is that these improvements in public transport may displace incumbent low-skilled residents through rent increases, a process often referred to as "transit-induced gentrification." As a consequence, these low-skilled incumbent residents might suffer welfare losses.[4] In this paper, I study how the construction of a bus rapid transit system[5] (BRT) in the city of Buenos Aires, Argentina, affected the spatial sorting of high- and low-skilled residents,[6] as well as the welfare consequences of these transport infrastructure improvements for the incumbent residents of the neighborhoods near the new transit lines.

A distinguishing feature of this paper is that it leverages individual level panel data on the addresses of more than two million residents. By using a panel data set at the individual level, I can follow residents as they move within the city, which allows me to describe intra-city migration patterns at a very fine spatial scale. I use these data to contribute to the literature on the effects of urban transit infrastructure in three ways. First, I document the heterogeneous response by high- and low-skilled residents in their intra-city migration decisions to the improvements in market access generated by the new BRT system. Second, I develop one of the first dynamic quantitative spatial equilibrium models of a city that combines commuting and migration decisions taken by forward-looking agents. Third, I use the dynamic quantitative spatial equilibrium framework with heterogeneous workers to estimate the welfare effects for high- and low-skilled incumbent residents in the city. The crucial difference with previous welfare estimates of transport infrastructure improvements is that, by relying on individual level panel data, I can account for the changes in migration decisions as a consequence of these infrastructure improvements. In doing so, I can differentiate between the welfare gains of the residents that were already living in affected areas before the improvements, and the welfare gains of the residents who live in the affected areas after a new steady state equilibrium is reached.

Between 2011 and 2017, the city of Buenos Aires opened six lines of a BRT network, the

[1] The world's urban population accounted for 34% of total global population in 1960, 55% in 2018, and is expected to account for 68% of total global population by 2050 (UN, 2018).

[2] According to Hannon et al. (2020), more than $1.4 trillion will be spent in new light-rail and metro projects in cities around the world between 2019 and 2025.

[3] See Redding and Turner (2015), Berg et al. (2017), and Roberts et al. (2020) for surveys of the literature on the effects of transport infrastructure improvements.

[4] The phenomenon referred to as "transit-induced gentrification" has been extensively studied in the urban planning and transportation engineering literature (Chapple and Loukaitou-Sideris, 2019, Dawkins and Moeckel, 2016, Chapple and Loukaitou-Sideris, 2019, Bardaka et al., 2018, Tehrani et al., 2019), with varying results, and often not accounting for the endogeneity concerns that might arise from the placement of the transport infrastructure (see Padeiro et al. 2019 for a review of the literature).

[5] A bus rapid transit system is a bus-based public transport system that includes dedicated roadways for buses and platform-level boarding, and assigns priority to buses at intersections where buses may interact with other traffic.

[6] I define a resident as low-skilled if she has no post-secondary education, and as high-skilled if she has at least some post-secondary education.

*Metrobus*. This network spans 51 kilometers and transports close to a million passengers a day. Although Buenos Aires already had a robust network of buses and subway lines that provided coverage for most of the city's neighborhoods, the new BRT lines resulted in an average increase in speed of 30% for the buses that went through those avenues (Kamrowska-Zaluska, 2017). These increases in transport speed led to increases in market access for commuters located in neighborhoods that were originally less well connected to locations in the city with high employment density. The effects of this change in market access are the main focus of this study.

In order to analyze the sorting patterns of residents by skill type, I leverage a uniquely detailed data source: the electoral register for all election years between 2011 and 2017 for the city of Buenos Aires and the surrounding suburbs, which allows me to track individuals as they change their residential location within the city. Since voting is compulsory in Argentina, the government maintains a register with residential addresses for all eligible voters. I have processed this register for the city of Buenos Aires and the surrounding suburbs within the Buenos Aires Province. With these data, I have geocoded[7] all addresses for more than three million eligible voters that reside in the city or surrounding suburban areas. I use this register, along with other data sources, such as census data and household surveys, to implement a reduced form analysis that relies on a historical IV identification strategy. In order to capture the full general equilibrium effects of the changes in the transport network on every location of the city, I construct a measure of commuter market access developed by Ahlfeldt et al. (2015) and extended to a setting with heterogeneous workers by Tsivanidis (2019).[8] I then examine how changes in this measure of market access affected the change in the share of high-skilled workers[9] in each neighborhood, the change in high- and low-skilled population, the change in population density in each neighborhood, and the change in residential floorspace prices. I address the endogeneity concerns regarding the placement of the BRT lines by constructing an instrumental variable for the changes in market access that relies on the placement of a tram system built in the first half of the twentieth century, that was dismantled by 1963. The identifying assumption behind this strategy[10] is that the placement of a tram system that had been decided prior to 1938 should not be correlated with changes in unobservable confounders between 2011 and 2017. Through this reduced form analysis, I find that, contrary to the hypothesis of transport-induced gentrification, the *Metrobus* led to increases in the high-skilled share mostly in neighborhoods that already had a relatively high level of high-skilled share, while neighborhoods in the lowest quintile of the distribution of high-skilled share saw a decrease in their high-skilled share as a consequence of the improvements in transport infrastructure. For the same increase in market access, a neighborhood at the eightieth percentile of the initial distribution of high-skilled share saw a proportional increase in its share of high-skilled residents six times higher than a neighborhood in the twentieth percentile. These patterns of spatial sorting between high- and low- skilled workers led to an increase in the segregation by skill type in the city.

Having described these sorting patterns between high- and low-skilled workers in a reduced form setting, I develop a dynamic model of a city in which residents who are forward-looking must make migration and commuting decisions each period. On the production side, I assume that each location produces a location-specific variety of the

[7] Geocoding is the process of assigning geographical coordinates to an address.

[8] This measure summarizes the total impact of the transit network on firms and residents, and can be derived from a wide class of models, as long as they imply a gravity equation for commuting flows (Tsivanidis, 2019).

[9] The share of high-skilled residents is defined as the population of high-skilled residents divided by the total population.

[10] This type of historical route IV approach was developed by Duranton and Turner (2012) and has been used frequently in the transportation economics literature (Redding and Turner, 2015).

consumption good under perfect competition. To do so, they must employ both types of labor (high- and low-skilled). These assumptions lead to labor demand being spatially distributed across the city, which in turn motivates the commuting decisions of the residents. This model draws from the dynamic structure used by Caliendo et al. (2019) and Balboni (2019) to study questions of regional economics, and from the commuting structure present in the static models of a city developed by Ahlfeldt et al. (2015) and Tsivanidis (2019), which allow for residents to live and work in different locations. I rely on the individual level migration data constructed from the electoral register to estimate key parameters of the model. I develop a numerical solution algorithm for the model, which is based on Caliendo et al. (2019). Having solved the model, I show that it predicts very similar patterns of spatial sorting as a response to the transport improvements as the ones found by the reduced form analysis.

Using this framework, I estimate the welfare impact of implementing this BRT system in the city of Buenos Aires on the residents that were already living near the BRT lines before they were put in place, which I refer to as incumbent residents. Welfare gains are measured as the difference in aggregate expected utility in the period prior to the implementation of the new transport infrastructure between a scenario where the changes in the transport infrastructure are implemented and a scenario where no change in the transport infrastructure is implemented. I find that welfare gains were significantly higher for high-skilled workers, who gained 1% on average, while low skilled-workers gained 0.6% on average. These gains were also distributed very unevenly between neighborhoods. Suburban neighborhoods located in outside the city proper saw the largest gains, while within the city proper, neighborhoods with the highest share of high-skilled residents saw the lowest gains. Indeed, low-skilled residents in these neighborhoods actually saw a decrease in welfare due to the construction of the BRT system. Suggesting that the larger increases in population and, consequently, in housing prices produced by the transport improvements in neighborhoods with higher high-skilled share offset a larger fraction of the welfare gains obtained by increases in market access for the residents that were already living there.

Finally, I use the dynamic quantitative framework to study how different counterfactual configurations of the BRT system would have produced different results in terms of spatial sorting and welfare gains. To do so, I solve for the full equilibrium path of endogenous variables under different assumptions on which BRT lines were built. I then compare, in different parts of the city, the welfare gains for incumbents in these counterfactual scenarios to the welfare gains for incumbents obtained from having built the entire BRT system. I find that building any line by itself would have led to significantly lower welfare gains for the high-skilled residents that were living close to the line before it was built. This, however, is not the case with the low-skilled residents, who, counter-intuitively, in fact are strictly better off in the case where only one line is built in the neighborhoods with the highest high-skill share. This is likely because in the case where all lines are built, housing prices increase in multiple locations, which reduces the value of the outside option of moving, however, when only one line is built, low-skilled residents can at least move out to other locations that are relatively cheaper, and so, do not lose welfare on average.

This paper contributes to several strands of literature. There is a large body of work that studies the impact of transport infrastructure improvements on the spatial distribution of residents and economics activity (Allen and Arkolakis, 2019, Baum-Snow, 2007, Duranton and Turner, 2012, Gonzalez-Navarro and Turner, 2018, Baum-Snow et al., 2018, Gibbons and Machin, 2005, Glaeser et al., 2008, Tsivanidis, 2019, Pathak et al., 2017). This paper contributes to this body of research by providing a general equilibrium framework that incorporates intra-city migration decisions, and using this framework to study the welfare effects for incumbent residents of the areas that were close to the transport improvements.

There is also a substantial body of work, especially in recent years, that develops quantitative spatial equilibrium models of cities that incorporate commuting (Allen et al., 2016, Severen, 2018, Heblich et al., 2020, Tsivanidis, 2019). These studies, however, lack migration data at the level of spatial aggregation necessary to study intra-city migration within the context of these models. As a consequence, these papers have not included agents that must make migration decisions in a dynamic setting. Because I can track individuals within the city as they change residential location, I am able to develop a dynamic quantitative spatial equilibrium model, where residents choose their residential location in a forward-looking manner.

This paper also contributes to the growing body of work that develops dynamic quantitative spatial equilibrium models for regions or countries (Caliendo et al., 2019, Balboni, 2019, Morten and Oliveira, 2018, Desmet et al., forthcoming, Bryan and Morten, 2019). Within this body of work, this paper is the first to incorporate commuting, which allows for workers to live and work in different locations.

A fourth strand of literature to which this paper contributes is the large body of research on the determinants of gentrification (Almagro and Domínguez-Iino, 2019, David et al., 2017, Couture et al., 2018, Dragan et al., 2019, Su, 2020). This paper contributes to this body of research by evaluating the hypothesis that transport infrastructure improvements might have led to gentrification of neighborhoods that were initially populated predominantly by low-skilled residents. I do not find that the BRT system implemented in Buenos Aires led to gentrification (defined as an increase in the share of high-skilled residents) in the neighborhoods that had an initially low level of high-skill share. I do, however, find that a more localized transport improvement, targeted only to neighborhoods with low high-skilled share, would have led to gentrification.

Finally, I contribute to the growing body of literature that studies the effects of improvements in transport infrastructure in developing countries[11] (Baum-Snow and Turner, 2017 Fajgelbaum and Redding, 2018, and Tsivanidis, 2019) by examining the impact of the BRT system implemented in the city of Buenos Aires. This is particularly important, given that most of the urban growth in the next thirty years will likely take place in developing countries (UN, 2018), and therefore, understanding the effects of urban transport infrastructure improvements in the specific context of developing countries has become increasingly relevant.

The paper is organized as follows. In Section 2, I provide a brief background on the city of Buenos Aires, and on the BRT project that will serve as the main focus of this paper; in Section 3, I describe the main sources of data used in this paper; in Section 4, I conduct a reduced form analysis to study the effects of the BRT system on key outcomes; in Section 5, I develop a dynamic quantitative spatial equilibrium model that allows me to study the welfare effects of the transport infrastructure improvements, as well as analyze counterfactual scenarios; in Section 6, I explain the estimation procedure for the parameters of the model; in Section 7, I analyze the welfare effects of the transport infrastructure improvements using the estimated model; in Section 8, I examine different counterfactual scenarios where only a subset of the BRT lines were built; finally, Section 9 concludes.

[11] See Bryan et al. (2020) for a review of the literature on urban economics in developing countries.

## 2 | BACKGROUND

The city of Buenos Aires is the capital and largest city of Argentina, with 2.9 million inhabitants.[12] Representing 28% of the country's GDP, it is the economic, political, and cultural center of Argentina. In this section, I will provide a brief overview of the city's demographic composition, as well as of the spatial distribution of economic activity and residents by skill type. I will also analyze the commuting patterns of residents of the city prior to the construction of the BRT system. I will then describe the BRT system that was put in place between 2011 and 2017 in the city, which will be the main focus of my analysis.

### 2.1 | City Structure

Economic activity in the City of Buenos Aires, measured by employment, is highly concentrated around the city's central business district (CBD). Alves et al. (2018) find that 50% of the formal employment in the city is located within a 4 kilometer radius of the CBD, in an area that constitutes less than 19% of the surface area of the city. From Figure 1 we can see that most of the employment is concentrated in 4 of the 15 communes, with 31% of employment located in commune 1, where the CBD is located.

We can then separate workers between those with at least some post-secondary education (which I will call high-skilled), and those with at most a completed secondary education (which I will call low-skilled). Figure 2 shows the fraction of workers by skill type that work in each commune. We can see that high-skilled workers are more spatially concentrated in commune 1, around the CBD, while low-skilled workers are more spatially dispersed. This is consistent with high-skilled workers being employed more intensively in industries that tend to be more spatially concentrated around the CBD.

In terms of the spatial distribution of residents by residential location, we can see in Figure 3 that there is a clear decreasing gradient in the high-skilled share—defined as the share of residents that are high-skilled in a location—from the north-east to the south-west. This exact same pattern can be seen by looking at Figure 4, that shows the average housing prices (in terms of US dollars per square meter) by district. We can see that the most expensive districts are in the north-eastern side of the city, while the least expensive parts of the city are in the southern and south-western districts.

### 2.2 | Commuting in Buenos Aires

With over 130 bus lines running through the city and six subway lines connecting the periphery of the city with the CBD (see Figure 5), Buenos Aires already had a robust urban transportation network prior to the construction of the Metrobus. As a result, almost 80% of trips within the city done in 2010 were done either by public transport (bus, subway or train), or by walking or biking. Table 1 shows the percentage of trips done in 2010 within the city by mode of transport by both high-skilled and low-skilled residents. We can see that although high-skilled residents used cars at a slightly higher rate, the majority of trips for both high- and low-skilled residents were done by public transit, walking, or biking (82% of trips for low-skilled residents, and 74% of trips for high-skilled residents).

[12] The entire metropolitan area surrounding the city of Buenos Aires constitutes the third largest metropolitan area in Latin America, with a population of 14.8 million, and represents 33% of the countries population.

### 2.3 | *Metrobus*: The BRT System of the City of Buenos Aires

Between 2011 and 2018, the city of Buenos Aires built a system of bus rapid transit (BRT) lines that spans 50.5 km (31.4 mi). The main objective of this infrastructure project was to improve commute times by bus between the residential areas in the periphery of the city and the neighborhoods with the highest employment density. Each line in the system is composed of two to four dedicated bus lanes, with physical separation from regular traffic, and platform-level boarding. By 2016, the system was composed of five lines that functioned full time.[13] Figure 6 shows the placement of these five lines in the city.

The first line of the Metrobus, the *Juan B. Justo* Line, was inaugurated in May 2011. This line is estimated to transport around 100,000 passengers a day, and reduce commuting costs along the Juan B. Justo Avenue by 40% according to government estimates[14]. The *9 de Julio* Line and the *Sur* Line were inaugurated in 2013. The *9 de Julio* Line joins the two main train stations of the city, and goes across the CBD. This line transports an estimated 250,000 passengers a day, and reduces bus travel time by 50% along the *9 de Julio* Avenue according to the government's estimates. The *Sur* Line serves the southern part of the city (which includes most of the lowest income neighborhoods in the city). The *Metrobus Sur* Line also serves approximately 250,000 passengers a day, it has reduced travel times by 15% according to the government's estimates, and has increased ridership by 30%. The *Cabildo* Line and the *25 de Mayo* Line were inaugurated in 2015. The *Cabildo* Line is estimated to serve around 200,000 passengers and was designed to connect the periphery of the city with the subway line D (see Figure 2 for a map of both the subway lines and the Metrobus lines). The *San Martin* line was inaugurated in 2016, and is estimated to have decreased transport times for buses along the line by 20%.

## 3 | DATA

The two main geographic units used in the analysis are the census tracts (*radio censal*) for the city of Buenos Aires, and the voting districts (*circuito electoral*) for the city of Buenos Aires. The city is divided into census tracts, which are the smallest geographical unit at which the data from the national census is reported. Census tracts in the city have an average surface area of 0.05 square kilometers (0.02 square miles), and an average population of 813 people. Voting districts are much larger, and divide the city into 167 regions, with an average surface area of 1.21 square kilometers (0.47 square miles), and an average population of 14952.

For the purpose of this analysis, I have included as the region of study all census tracts and districts that are separated by at most three kilometers to a BRT stop (measured in terms of the distance to the centroid of the tract or district). This results in a total of 3731 census tracts and 201 districts, most of which are located within the administrative boundaries of the City of Buenos Aires, but some of which extend outside the city borders and into the surrounding suburbs that are part of the larger metropolitan area. Although the three kilometer limit is a fairly arbitrary measure, considering that the effects of the BRT decay with respect to the distance to the BRT, it is unlikely that an inclusion of a wider area of the metropolitan region would significantly alter the results.

As my main source of data on residential location at the individual level, I use the electoral registries for the years 2011, 2013, 2015, and 2017. Voting is compulsory for all

[13] A sixth line, the *25 de Mayo* line operates only during rush hour, was built on a highway, and runs mid-distance and long-distance buses that connect the suburban metropolitan region to the city center. Since this line does not have any intermediary stops within the city proper, I will not include it in my main analysis.

[14] This would translate to a reduction of 44 minutes a day, or 7 days per year for the average commuter of this line, according to government estimates.

Argentine citizens between the ages of 18 and 70, and the polling place for each citizen is decided based on their residence, so the government keeps an up-to-date record of every person's address, which is updated for every election. I have obtained the electoral register for the city of Buenos Aires for the years 2011, 2013, 2015, and 2017. These registers include the National ID number, name, sex at birth, birth year, residential address as free text, and a description of occupation or profession for every person that is eligible to vote. I have cleaned and geocoded (assigned GPS coordinates) almost all addresses[15] in the city of Buenos Aires (which account for approximately 2.5 million people).

One potential concern with these data arises from the fact that residence changes are self-reported, which might imply under-reporting of changes in residence or a lag in the change of residential addresses. Although I cannot rule out this possibility, and in fact it is highly likely that these data does not reflect the exact residential location of every citizen at each point in time, I show that the cross-sectional distribution of population in 2011 for the city of Buenos Aires that is implied in the electoral register data is highly correlated with the census data from the 2010 census. Figure 7 shows a binscatter plot that compares the number of voters from the electoral registry data that were geocoded in each census tract in 2011 with the population in each census tract in 2010, obtained from the 2010 national census. As we can see in this figure, there is a positive correlation between the 2010 census population and the number of voters geocoded from the electoral registry, with the 2010 census population being above the geocoded population for every bin average, which is consistent with the number of eligible voters being a fraction of the total population in each census tract.

I have obtained a restricted version of the the Annual Household Survey for the city of Buenos Aires (*Encuesta Anual de Hogares*) that details the employment location of each household member surveyed that is employed for every year between 2010 and 2019. With these data, I estimate the fraction of high-skilled and low-skilled workers that work in each commune of the city. I then employ the land use data census for the year 2011 to build a measure of employment by district within each commune, under the assumptions that employment is distributed in proportion to the fraction of land used for commercial and productive purposes within each district, and that the skill-share is constant across districts within a commune.[16]

I use ArcMap's Network Analysis tool to construct a model of the city's transportation network before and after the BRT lines were built. I then calculate the minimum commute time between census tracts using Dijkstra's algorithm to measure the changes in commute times after the BRT was put in place.[17] I also use the 2010 mobility survey for the City of Buenos Aires (*Encuesta de Mobilidad Domiciliaria*) to calculate average travel speeds by mode before the BRT was put in place, as well as to estimate the semi-elasticity of commuting with respect to travel times for high- and low-skilled residents in the model.

Regarding floorspace prices, I use the ask price for a subset of all listings from a major online marketplace from 2009 to 2017, as well as additional information such as the number of bedrooms, surface area, address, and GPS coordinates. With these data, I estimate the average housing price per square meter by district and by census tract. For floorspace prices in the suburbs, which are located in the Buenos Aires Province, I rely on Secretaría de Planeamiento (2012) and Secretaría de Planeamiento (2017) that report the average price per

[15] One limitation with these data is that I require a full address (street name and number) in order to assign geographical coordinates to a citizens residential location. Therefore, I cannot locate citizens that live in disadvantaged low-income settlements that do not have a formal address. Because of this limitation, I can only geo-locate approximately 95% of the electoral register. See Appendix A for more information on this process.

[16] See Appendix A for details.

[17] See Appendix A for more information on how the commute times are calculated.

square meter of land in each county (*partido*), which is a significantly larger area than each district. For this reason, my current regressions using housing prices should be interpreted with caution, given the discrepancy in the data used for the city proper and the surrounding suburban districts.

As mentioned above, I employ the land use data from the 2011 land use map (*Relevamiento de Usos del Suelo*) produced by the government of the city of Buenos Aires. This data set contains land use information on every parcel of land in the city and, crucially, allows me to distinguish between land used for commercial and productive purposes from land used for residential purposes or other uses.

Finally, I use data from the 2010 national census,[18] as well as from the National Household Survey (*Encuesta Permanente de Hogares*) to supplement the previously described sources of data.

### 3.1 | Defining High- and Low-Skilled Residents

The electoral registry data includes a free text description of the occupation of each citizen. I processed this information and matched each description to a one-digit occupation group in the International Standard Classification of Occupations (ISCO).[19] Table 2 shows the ten one-digit occupation groups of the ISCO.

I define a high-skilled worker as a worker who has at least some post-secondary education. From the National Household Survey for the city of Buenos Aires in 2010, we can calculate the fraction of workers within each one-digit ISCO occupation group that is high-skilled in the city of Buenos Aires according to this classification. We can see in Figure 8 that for five of the occupation groups,[20] the majority of workers in those occupations are low-skilled. For the remaining five occupation groups,[21] the majority of the workers in those occupations are high-skilled in the city of Buenos Aires. Therefore, I classify as a high-skilled worker every resident of the city categorized as being in an occupation group where the majority of the workers in that occupation are high-skilled, and I classify as a low-skilled worker every resident of the city in an occupation group where the majority of the workers in that occupation are low-skilled. Figure 9 compares the share of high-skilled residents by census tract calculated using the electoral register data from 2011 to the share of high-skilled residents by census tract obtained from the 2010 census.[22] We can see in this graph that there is a positive and significant correlation between both measures.

## 4 | REDUCED FORM ANALYSIS

### 4.1 | Commuter Market Access: Using the Model to Measure the Impact of the BRT System

In order to capture the full general equilibrium effects of the changes in the transport network due to the construction of the *Metrobus* BRT on every location of the city, I construct a measure of commuter market access developed by Ahlfeldt et al. (2015) and Tsivanidis (2019). This measure summarizes the total impact of the transit network on firms and

[18] The following census was scheduled to occur in 2020, but has been postponed due to the emergency sanitary measures put in place as a consequence of the COVID-19 pandemic.

[19] See the Appendix A for more information on this matching process.

[20] These are: armed forces, elementary occupations, plant and machine operators, and assemblers, craft and related trades workers, and service and sales workers.

[21] Which are: managers, professionals, technicians and associate professionals, clerical support workers, and skilled agricultural, forestry and fishery workers.

[22] Where high-skilled is defined as having at least some post-secondary education.

residents, and can be derived from a wide class of models, as long as they imply a gravity equation for commuting flows (Tsivanidis, 2019). In Section 5, I derive this measure within the dynamic framework developed in this paper. I show that the commuter market access (CMA) for a resident of skill-type $g$ living in a residential location (district or census tract) $n$ at time $t$ can be defined as

$$CMA_{gn,t} = \sum_j \left( \frac{w^g_{j,t}}{d_{nj,t}} \right)^{\theta_g}.$$

Where $d_{nj,t} = \exp(\kappa\tau_{ij,t})$ are the commute costs associated with commuting between $i$ and $j$, with $\tau_{ij}$ being the minimum commute time between $i$ and $j$ calculated using Dijkstra's algorithm, and $\kappa$ being the semi-elasticity of commute costs to commute times; $w_{gj,t}$ is the model-consistent wage for a worker of type $g$ that works in workplace location $j$ at time $t$, and can be calculated with information on the spatial distribution of employment population and residential population by skill type at a given year (given parameter estimates); and $\theta_g$ is the semi-elasticity of commute shares with respect to commute costs ($d_{nj,t}$).

Given parameter estimates[23] for $\theta_g$ and $\kappa$, we can calculate these measures of market access for 2011 by using data on employment population at the district level[24] and residential population at the census tract level, as well as commute times calculated through a commuting model before the BRT was put in place. This will result in a vector of market access $\{CMA_{gn,2011}\}_{g\in\{h,l\},n\in I}$ for every residential location $n$ and both skill types $g$ in 2011. Following Tsivanidis (2019), I then calculate $\{CMA_{gn,2017}\}_{g\in\{h,l\},n\in I}$ by fixing the employment and residential population at 2011 levels and by changing the commute costs to reflect the new commute times with the full BRT put in place. I fix the employment and residential population at 2011 levels in order to abstract from endogenous changes in population that might bias the estimation.[25] I then calculate the change in the logged market access, $\Delta_{2011-2017}CMA_{gn,t}$, as the measure that will allow me to capture the full general equilibrium effects of changes in the transport network on the outcomes analyzed in this section.

Figures 10 and 11 show the market access measure estimated for every census tract in the city in 2011. We can clearly see two things in these maps: first, market access is higher near the subway lines and closer to the CBD, where most subway lines converge; second, although the market access measure for high- and low-skilled workers are clearly not identically distributed across census tracts, there is a strong correlation between these measure. In fact, the estimated Pearson correlation coefficient between these two measures at the census tract level in 2011 is 0.99.

In Figures 12 and 13, we can see that the change in commuter market access for both high- and low-skilled workers decreases with distance to the new BRT lines, and increases with distance to the CBD. This is consistent with the fact that places closer to the BRT saw a larger proportional increase in their commute speeds, while places further from the CBD were able to benefit relatively more from these increases in speed, since they translated into larger decreases in commute times towards the major areas of employment. Moreover, the changes in market access for high- and low-skilled workers are highly correlated. In fact, the estimated correlation coefficient between changes in market access for high- and

[23] See Section6 for details on how these parameters are estimated.

[24] Due to data limitations, in the current version of this paper, I will assume that all residents living in the suburbs will commute to work to the city proper. This is of course not a realistic assumption, but in order to improve on this assumption I would need sufficiently geographically detailed data on employment for the Buenos Aires Province.

[25] In Appendix B I show that this assumption does not affect the main results from the reduced form analysis.

low-skilled workers is 0.98. Due to this high degree of correlation between both measures, I conduct the main reduced form analysis by regressing changes in the outcomes of interest on changes in the commuter market access measure for high-skilled workers. All the main results hold when changes in market access for low-skilled workers are used instead.[26]

### 4.2 | Identification Strategy: Instrumenting BRT Placement With Historical Tramway

The main identification concern when studying the effects of transport infrastructure is that the placement of this infrastructure is clearly not randomly allocated. In particular, one might worry that the BRT lines were placed such that they served neighborhoods that had specific unobserved characteristics, such as a secular population trend (either decreasing or increasing), that would bias the OLS results. In order to address these concerns, I use a historical route IV approach[27] based on Duranton and Turner (2012).

In the early twentieth century, Buenos Aires had a vast system of electric trams that connected the city center to the suburbs of the city. In the 1920s, privately operated buses started competing with the trams by running parallel to the tramways and often outpacing the tramcars in order to lure passengers that were waiting for the trams (Singh, 2018). Due in part to this competition and in part to the construction of five subway lines in the following decades, the trams saw a steady decrease in ridership between the late 1930s and the 1960s. In 1963, the tramway system was finally dismantled. The bus lines, however, continued operating on essentially the same routes that the old tramway system followed. When the Metrobus BRT lines were built, they were chosen so as to follow existing bus routes, with the added restriction that they had to be built on large two-way avenues that allowed for the construction of boarding platforms and exclusive lanes for buses.

Given that many of the bus routes were historically determined by the placement of the tram system, I construct an instrument based on the tramway routes that were built before 1938 and that ran through large two-way avenues. Figure 14 shows the placement of these tramway routes and the *Metrobus* BRT system, we can see that there is a clear spatial correlation between the two. The instrument is constructed by calculating commute times assuming that the BRT was built following the tramway routes, using those times to calculate a measure of market access for each residential location, and then calculating the change in market access between 2011 and this hypothetical scenario. This hypothetical change in market access is used to instrument for the change in market access calculated using the actual placement of the BRT. The identifying assumption is that the placement of the tramway routes decided prior to 1938 is not correlated with contemporary changes in unobservable variables that might correlate with the placement of the BRT and affect the outcomes of interest directly.

In Table 3, we can see the results from the first stage regressions of the 2SLS IV estimation for every specification that is used in the following section. We can see that the historical IV instrument has indeed a very strong first stage, with Cragg-Donald Wald F statistics that range from 1470 to 277. With these values of the F-statistics we can confidently reject the null hypothesis of weak instruments in the test proposed by Stock and Yogo (2005).[28] We can also gather from these regressions that the coefficient for the regression of observed changes in market access on instrumented changes in market access is positive and significant, as

[26] See Appendix B for reduced form results using changes in the commuter market access measure for low-skilled workers.

[27] This type of identification strategy has been used frequently in the transportation economics literature (Redding and Turner, 2015)

[28] The null hypothesis of weak instruments can also be rejected for the F-test proposed by Sanderson and Windmeijer (2016).

expected given the spatial correlation between the tram lines and the BRT lines.

### 4.3 | Main Specification and Results

In this section, I analyze the effects of improving the transport infrastructure through the construction of the BRT system in Buenos Aires on the share of high-skilled residents[29] in each census tract, the average floorspace price in each census tract, the population density, and the population density by skill type in each census tract. In order to do so, I estimate the following equation through OLS and by instrumenting the change in market access (defined in Subsection 4.1) with a historical IV (defined in Subsection 4.2):

$$\Delta Y = \alpha + \beta_1 \Delta CMA + \beta_2 \text{HS share}_0 + \beta_3 \Delta CMA \times \text{HS share}_0 + \Lambda + \epsilon. \quad (1)$$

Where $\Delta Y$ is the change in the outcome of interest; $\Delta CMA$ is the change in the commuter market access as defined in Subsection 4.1 between 2011 and 2017 due to the construction of the BRT; HS $\text{share}_0$ is intended to capture the initial differences in the high-skilled share between census tracts, and is calculated as the average high-skilled share for all census tracts that are contiguous to a given census tract[30]; $\Lambda$ is a set of potential controls, such as fixed effects at the neighborhood level; and $\epsilon$ is the error term in each regression.

Table 4 shows the results of estimating equation 1 both by OLS and IV for the change in the high-skilled share by census tract between 2011 and 2017. We can first see from these results that while the IV estimation in column 2 results in a positive and significant coefficient for the change in market access, the OLS estimation produces, however, produces a negative coefficient. This result suggests that the placement of the BRT might have favored areas where the high-skilled share had a decreasing secular trend, resulting in a negative bias of the coefficient. From the IV estimation, we find that a 1% change in market access leads to a 0.27% change in the high-skilled share. In columns 3 and 4, we can see that once we include an interaction term between the initial high-skilled share and the change in market access, the IV and the OLS estimates become very similar to one another. Both of these estimates show that, although the high-skilled share increased on average as a consequence of the increase in market access, this effect was not equal across neighborhoods. In census tracts with a higher initial high-skilled share, the high-skilled share increased more as a consequence of the increase in market access. In fact, using the results from column 3 we see that the derivative of $\Delta \log(\text{CMA})$ with respect to $\Delta \log(\text{HS share})$ is equal to 0.1 for a census tract at the twentieth percentile of the initial high-skilled share distribution,[31] and equal to 0.8 for a census tract at the eightieth percentile of the distribution of initial high-skilled share. Finally, these results are robust to controlling for commune or county[32] fixed effects (although the significance level decreases due to loss of power in the estimation), which suggests that these results are not driven by differential effects across neighborhoods. Taken as a whole, these results imply that the BRT increased segregation by skill type within

---

[29] The share of high-skilled residents in a census tract $i$ at time $t$ is defined as $\frac{L^h_{i,t}}{L^h_{i,t}+L^l_{i,t}}$ where $L^g_{i,t}$ is the total population of skill type $g$ residing in $i$ at time $t$.

[30] I use this measure in order to avoid mechanical correlations between the initial high-skilled share and the change in high-skilled share. All coefficients are not statistically different at 5% level when using the actual high-skilled share instead of this measure in the regressions.

[31] Recall that this is not the actual initial high-skilled share distribution, but the average distribution for all the census tracts contiguous to a given census tract in 2011.

[32] Every census tract is included in a commune within the city proper or in a county within the suburbs located in the Greater Buenos Aires area.

the city, by increasing the high-skilled share at a higher rate in census tracts that had an initially higher level of high-skilled share.

In Table 5, we can see that the same patterns observed in the effect of market access on the high-skilled share can be observed in the effect of market access on floorspace prices. However, it is worth noting that the floorspace price data used for census tracts in the Buenos Aires Province suburbs was obtained from a different data source at a much more aggregate level. This results should be interpreted with caution for the time being.

Table 6 shows the results of estimating equation 1 by IV for the changes in total population, population of high-skilled residents, and population of low-skilled residents by census tract. We can see that total population, as well as high-skilled population increased more in census tracts with a higher initial high-skill share. A 1% increase in market access led to an increase of the high-skilled population of 0.9% on average in a census tract at the twentieth percentile of the initial high-skilled share distribution, while the same same 1% increase in market access led to an increase of the high-skilled population of 1.18% on average in a census tract at the same twentieth percentile. The fact that the high-skilled population seems to have reacted more than the total population, and than the low-skilled population, may explain why the high-skilled share increase more with an increase in market access in places that had an initially higher share of high-skilled residents.

## 5 | MODEL

Most quantitative spatial equilibrium models of a city rely on comparative statics for welfare analysis (e.g. Tsivanidis, 2019, Ahlfeldt et al., 2015). This analysis can tell us the average welfare gains for residents living at a specific location after the changes of interest have occurred, and a new static equilibrium has been reached. We know, however, that residents are mobile within a city, and can change their residential location as a response to changes in the city's fundamentals (such as transport infrastructure improvements). Therefore, in order to analyze the welfare gains for residents living in each location before the improvements in transport infrastructure were put in place, I develop a model that can incorporate migration decision by forward-looking agents explicitly. This model draws from the dynamic structure of Caliendo et al. (2019) and Balboni et al. (2020), and from Ahlfeldt et al. (2015) and Tsivanidis (2019) in terms of the commuting structure, that allows workers to work and live in different locations within the city.

### 5.1 | Setup

A city is composed of $I$ residence locations (indexed by $i$ or $n$), and $J$ workplace locations, (indexed by $j$ or $s$). I assume that there is a fixed mass $L$ of atomistic residents in the city[33] that are infinitely lived, and who must decide how much to consume, where to work, and where to live each period. These residents can either be high-skilled or low-skilled (indexed by $g \in \{h, l\}$). Each period, residents start in a residence location, which will depend on their choices from the previous period. They then observe a vector of idiosyncratic match productivities for each workplace location. They take wages in each location as given, as well as commute costs from their residence location. With this information, they choose a workplace location for that period. Once they know their workplace location, and have obtained a wage income discounted by the commute cost of commuting from their residence

[33] This is a closed city model. This assumption has been made mostly due to data limitations, since I have limited information on migration patterns for individuals between the city and the rest of the country (or the world), as well as for modeling convenience.

location to their workplace location, they choose how much to consume of the freely traded consumption bundle and of the floorspace (housing services) available in their residence location. Having consumed the consumption bundle and the housing services, they observe a vector of idiosyncratic preference shocks for each residence location for the next period. With this information, they choose a residence location for the next period. If they decide to leave their current location, they pay an origin-destination specific moving cost.

Production in each workplace location is characterized by a static problem. Each location produces a variety of the final consumption good[34] under perfect competition, employing both types of labor (high- and low-skilled).

I assume a fixed supply of floorspace in each residence location, which, along with aggregate floorspace demand by each skill-type in each location, determines the floorspace price in that location for that period. I also assume that all floorspace is owned by an absentee landlord to whom all rents are payed.

## 5.2 | Resident's Problem

### 5.2.1 | Consumption Problem

Given a residential location $i$ and a workplace location $j$, a resident of skill type $g \in \{h, l\}$ will have an intra-period utility function at time $t$ defined as

$$U^{g}_{ij,t} = \left(\frac{c^{g}_{ij,t}}{\alpha}\right)^{\alpha} \left(\frac{H^{g}_{R,ij,t}}{1-\alpha}\right)^{1-\alpha}.$$

Where $c^{g}_{ij,t}$ is the level of consumption of the freely traded numeraire consumption good for a resident of skill type $g$, and $H^{g}_{R,ij,t}$ is the level of housing services or residential floorspace consumed by a resident of skill type $g$.

I assume that a resident of $i$ that works in workplace location $j$ earns a wage income $\frac{w^{g}_{j,t}\varepsilon_{j,t}}{d_{ij,t}}$, where $w^{g}_{j,t}$ is the competitive wage payed at workplace location $j$ for a worker of skill type $g$, $\varepsilon^{g}_{j,t}$ is the idiosyncratic match-productivity shock for location $j$, and $d_{ij,,t}$ is the commuting cost[35] between residential location $i$ and workplace location $j$. Residents will also receive a lump-sum housing income $\psi^{g}_{i,t}$, which represents the fraction of all housing rents that are payed to this resident, and is calculated as

$$\psi^{g}_{i,t} = o_{ig}\frac{E_{i,t}}{L_{Ri,t}},$$

where $o_{ig}$ is the share of homeowners of type $g$ living in residential location $i$, $E_{i,t}$ is the total expenditure on housing in $i$ at time $t$, and $L_{Ri,t}$ is the total population in $i$ at time $t$. In section D.1, I explore two alternative assumptions regarding home ownership. Mainly, I first assume that both high- and low-skilled residents recieve the same lump sum payment, which depends on the total home ownership rate in each residential location (regardless of skill type). I then also explore the results of the model under the assumption that all residents are renters and rents are payed to an absentee landlord. I find that, the main

[34]Consumers will consume a CES bundle of these varieties.

[35]Following Ahlfeldt et al. (2015), I define $d_{ij,t} = \exp(\kappa\tau_{ij,t})$, where $\tau_{ij,t}$ is the average commute time in minutes between $i$ and $j$ at time $t$, and $\kappa$ is the semi-elasticity of commute costs with respect to commute times. I make the simplifying assumption that there is only one mode of transport, which includes walking, and public transportation. As discussed in Section 2, the vast majority of trips within the city for both high- and low-skilled workers were done either walking o by taking public transport.

conclusions derived from this model hold under all three possible assumptions.

The consumer's intra-period problem can be then be expressed as

$$\max_{\{c^g_{i,t}, H^g_{R,ij,t}\}} C^g_{ij,t} = \left(\frac{c^g_{ij,t}}{\alpha}\right)^{\alpha}\left(\frac{H^g_{R,ij,t}}{1-\alpha}\right)^{1-\alpha} \quad \text{subject to } c^g_{ij,t} + r_{Ri,t}H^g_{R,ij,t} = \frac{w^g_{j,t}\varepsilon^g_{j,t}}{d_{ij,t}} + \psi^g_{i,t}.$$

Where $r_{Ri,t}$ is the price of a square meter of residential floorspace in location $i$. I assume that the housing rents are payed to an absentee landlord who does not consume goods or services in the city.

The solution of the consumer's intra-period problem implies

$$c^{g*}_{ij,t} = \alpha\left[\frac{w^g_{j,t}\varepsilon^g_{j,t}}{d_{ij,t}} + \psi^g_{i,t}\right], \tag{2}$$

$$H^{g*}_{R,ij,t} = \frac{(1-\alpha)}{r_{Ri,t}}\left[\frac{w^g_{j,t}\varepsilon^g_{j,t}}{d_{ij,t}} + \psi^g_{i,t}\right]. \tag{3}$$

Which implies an indirect utility function

$$C^{g*}_{ij,t} = \frac{w^g_{j,t}\varepsilon^g_{j,t}r^{\alpha-1}_{Ri,t}}{d_{ij,t}} + \psi^g_{i,t}r^{\alpha-1}_{Ri,t}.$$

### 5.2.2 | Dynamic Problem

The agent's full dynamic problem at time $t$ can be expressed as the following Bellman equation:

$$v^g_{n,t} = \max_{\{i,j\}}\left\{C^{g*}_{nj,t} + \beta E_t[v^g_{i,t+1}] - \mu^g_{ni} + \eta^g_{i,t}\right\} \tag{4}$$

Where $i$ is the residential location that the agent chooses for $t+1$, and $j$ is the workplace location that the agent chooses for period $t$, $\mu^g_{ni}$ is the moving cost required to move from residential location $n$ to location $i$, which is assumed to be constant over time, and $\eta^g_{i,t}$ is the idiosyncratic preference shock on residential location choices for period $t+1$, which is observed in period $t$.

Note that the maximization problem is separable between choosing the optimal workplace at time $t$, and choosing the optimal residence location for next period $i$. Therefore, we can write

$$v^g_{n,t} = \max_{\{j\}}\left\{\frac{w^g_{jt}\varepsilon^g_{jt}r^{\alpha-1}_{Rn,t}}{d_{nj}} + \psi^g_{n,t}r^{\alpha-1}_{Rn,t}\right\} + \max_{\{i\}}\left\{\beta E_t[v^g_{i,t+1}] - \mu^g_{ni} + \eta^g_{i,t}\right\}.$$

Taking expectations with respect to the joint distribution of both idiosyncratic shocks ($\varepsilon_j$ and $\eta_i$), we get

$$E_0 v^g_{n,t} = E_0\left[\max_{\{j\}}\left\{\frac{w^g_{j,t}\varepsilon^g_{j,t}}{d_{nj,t}}r^{\alpha-1}_{Rn,t} + \psi^g_{n,t}r^{\alpha-1}_{Rn,t}\right\}\right] + E_0\left[\max_{\{i\}}\left\{\beta E_t[v^g_{i,t+1}] - \mu^g_{ni} + \eta^g_{i,t}\right\}\right].$$

Let $V^g_{n,t} = E_0 v^g_{n,t}$, then it follows from the Markovian structure of this dynamic problem that $E_t[v^g_{n,t+1}] = V^g_{n,t+1}$, therefore,

$$V^g_{n,t} = E_0 \left[ \max_{\{j\}} \left\{ \frac{w^g_{j,t} \varepsilon^g_{j,t}}{d_{nj,t}} r^{\alpha-1}_{Rn,t} + \psi^g_{n,t} r^{\alpha-1}_{Rn,t} \right\} \right] + E_0 \left[ \max_{\{i\}} \left\{ \beta V^g_{i,t+1} - \mu^g_{ni} + \eta^g_{i,t} \right\} \right]. \quad (5)$$

We can solve both expectations by using the fact that $\varepsilon^g_{j,t}$ is distributed extreme value type I, and $\eta^g_{i,t}$ is distributed extreme value type II (see Appendix C). By solving these expectations we obtain

$$V^g_{n,t} = \tilde{T}_g \left[ \Phi^{\frac{1}{\theta_g}}_{Rgn,t} + \psi^g_{n,t} \right] r^{\alpha-1}_{Rn,t} + \nu_g \ln \sum_{i=1}^{I} \exp\left( \beta V^g_{i,t+1} - \mu^g_{i,n} \right)^{\frac{1}{\nu_g}}. \quad (6)$$

Where $\Phi_{Rgi,t} = \sum_j \left( \frac{w^g_{j,t}}{d_{ij,t}} \right)^{\theta_g}$, $\theta_g$ and $\tilde{T}_g$ are the shape and scale parameters of the CDF of $\varepsilon^g_j$, and $\nu_g$ is the scale parameter of the CDF of $\eta^g_i$.

### 5.3 | Neighborhood Migration Flows

Let $m_{in,t}$ be the fraction of agents that are residing in location $n$ at the beginning of period $t$, and move to location $i$ by the end of period $t$. From the distribution of idiosyncratic preferences $\eta_{n,t}$ we can obtain the following equation:[36]

$$m^g_{in,t} = \frac{\left[ \exp(\beta V^g_{i,t+1} - \mu^g_{in}) \right]^{\frac{1}{\nu_g}}}{\sum_{m=1}^{N} \left[ \exp(\beta V^g_{m,t+1} - \mu^g_{mn}) \right]^{\frac{1}{\nu_g}}} \quad (7)$$

### 5.4 | Labor Supply

As in Tsivanidis (2019), the probability of a worker that lives in $i$ at time $t$ decides to work in location $j$ is:

$$\Pi_{j|tig} = \frac{(w^g_{j,t}/d_{ij,t})^\theta}{\sum_s \left( \frac{w^g_{s,t}}{d_{is,t}} \right)^\theta} = \frac{(w^g_{j,t}/d_{ij,t})^\theta}{\Phi_{Rgi,t}}. \quad (8)$$

Therefore, labor supply at time $t$ for workplace location $j$ will be

$$L^g_{Fj,t} = \sum_{i=1}^{I} \Pi_{j|tig} L^g_{Ri,t}, \quad (9)$$

where $L^g_{Ri,t}$ is the residential population at time $t$ in location $i$ of workers of type $g$.

[36] See Appendix C for derivation.

### 5.5 | Residential Population

Let $m^g_{in,t}$ be the fraction of agents of type $g$ that are residing in location $n$ at the beginning of period $t$, and move to location $i$ by the end of period $t$. It must be then that

$$L^g_{Ri,t} = \sum_{n=1}^{I} m^g_{in,t} L^g_{Rn,t-1} \tag{10}$$

### 5.6 | Production

I will assume there are J varieties of the consumption good, differentiated by location of production, and supplied in a competitive market. Consumers have CES preferences over each variety, with elasticity of substitution $\sigma > 1$. Producers solve a static problem each period $t$, where they choose their demand for low-skill and high-skill effective units of labor, in order to maximize their profits in that period. I assume firms produce using an Cobb-Douglas production function such that the output of the variety of the final good produced in location $j$ at time $t$ will be:

$$q_j = A_j(\tilde{L}^l_{jt})^{\rho_j}(\tilde{L}^h_{jt})^{(1-\rho_j)}$$

Where $A_j$ is the exogenous TFP from producing in j, $\tilde{L}^h_{Fj,t}$ is the demand at time t in location j for effective units of high-skill labor, and $\tilde{L}^l_{Fj,t}$ is the demand at time t in location j for effective units of low-skill labor. I allow $\rho_j$ to vary by workplace location in order to capture the fact that firms in different parts of the city may have different skill intensities.[37] Taking first order conditions of the producers problem, one finds that

$$w^h_{j,t}\tilde{L}^h_{Fj,t} = (1-\rho_j)X_{jt}, \tag{11}$$

$$w^l_{j,t}\tilde{L}^l_{Fj,t} = \rho_j X_{jt}. \tag{12}$$

Where $X_{jt}$ is the total expenditure on variety j at time t. From the CES demand for varieties of the consumption good, we know that in equilibrium $X_{jt} = p^{1-\sigma}_{jt}X$, where $X = \sum_{i=1}^{I} \alpha \sum_{g\in\{h,l\}} \bar{y}_{igt}$ is the total expenditure on consumption in the economy, and $\bar{y}_{igt}$ is the mean income of a worker of type g in location i at time t. Perfect competition will imply that the price of each variety is equal to its marginal cost at each time t, $p_{jt} = A^{-1}_j(w^l_{jt})^{\rho_j}(w^h_{jt})^{(1-\rho_j)}$. Combining these results we can express[38] labor demand at each location j at time t for each skill type g as a function of the wage vector $\mathbf{w}_t = \{\{w^h_{jt}\}^J_{j=1}, \{w^l_{jt}\}^J_{j=1}\}$,

$$\tilde{L}^g_{Fj,t} = f_{jgt}(\mathbf{w}_t) \tag{13}$$

Finally, we must relate the effective units of labor of each skill type to the physical units of labor of that skill type ($L^g_{Fjt}$). Let $\bar{\varepsilon}^g_{jt}$ be the average worker-match productivity for a worker of skill type g, at time t in workplace location j. This average productivity can be

[37] The vector of $\rho_j$ for all j will be calibrated using the skill share of employment observed in 2010 for each workplace location.

[38] See Appendix C for the derivation.

computed as[39]

$$\bar{\varepsilon}^{g}_{jt} = T_g \sum_{i}^{I} \pi_{j|tig}^{-\frac{1}{\theta_g}} \frac{1}{d_{ijt}} \frac{\pi_{j|tig} L^{g}_{Rit}}{\sum_{n}^{I} \pi_{j|tng} L^{g}_{Rnt}}. \tag{14}$$

It follows then from the definition of the worker-match productivities that

$$L^{g}_{Fjt} = \frac{\tilde{L}^{g}_{Fjt}}{\bar{\varepsilon}^{g}_{jt}} \ \forall\, g \in \{h, l\}.$$

### 5.7 | Labor Market Clearing

From equation (9) we obtain the labor supply in workplace $j$ at time $t$ of skill type $g$, given a residential population distribution $\{L^{g}_{Ri,t}\}_{i=1}^{I}$ of skill type $g$. At the same time, from equation (13) we know that labor demand can be expressed as a function of the wage vector $\bar{w}_t$. Therefore, labor market clearing implies

$$L^{g}_{Fjt} = \frac{f_{jgt}(\mathbf{w}_t)}{\bar{\varepsilon}^{g}_{jt}} = \sum_{i=1}^{I} \frac{(w^{g}_{j,t}/d_{ij,t})^{\theta}}{\sum_j \left(\frac{w^{g}_{j,t}}{d_{ij,t}}\right)^{\theta}} L^{g}_{Ri,t}. \tag{15}$$

### 5.8 | Floorspace Market Clearing

Each residential location has a fixed amount of floorspace $\bar{H}_{Ri}$. Let $\bar{y}^{g}_{i,t}$ be the expected income of a resident of $i$ of skill type $g$ at time $t$, where the expectation is taken with respect to the distribution of match productivities. From equation (3) we know that the aggregate demand for housing in location $i$ is

$$H^{S}_{Ri} = \sum_{g\in\{h,l\}} \frac{L^{g}_{Ri,t}\bar{y}^{g}_{i,t}}{r_{Ri,t}}(1-\alpha).$$

Solving for $\bar{y}^{g}_{i,t}$ we find that[40]

$$H^{S}_{Ri} = \sum_{g\in\{h,l\}} \frac{L^{g}_{Ri,t}\left(\tilde{T}_g \Phi_{Rgi,t}^{\frac{1}{\theta_g}} + \psi^{g}_{n,t}\right)}{r_{Ri,t}}(1-\alpha).$$

Market clearing for residential floorspace requires that the supply for residential floorspace in location $i$ ($\bar{H}_{Ri}$) be equal to the demand for residential floorspace in location $i$, for all residential location, at every time $t$, which implies that

$$r_{Ri,t} = \frac{\sum_{g\in\{h,l\}} L^{g}_{Ri,t}\left(\tilde{T}_g \Phi_{Rgi,t}^{\frac{1}{\theta_g}} + \psi^{g}_{n,t}\right)(1-\alpha)}{\bar{H}_{Ri}}. \tag{16}$$

[39] See Appendix C for derivation.
[40] See Appendix C of this derivation.

### 5.9 | Equilibrium Definitions

In this model, the endogenous state of the economy at time $t$ is determined by the vector of residential labor allocation $\{L^{g}_{Ri,t}\}$, and the vector of workplace labor allocation $\{L^{g}_{Fj,t}\}$. The time varying fundamentals of this economy are the pairwise transport time[41] from each residential location to each workplace location, $t_{ij}$. The constant fundamentals are the pairwise moving costs $\mu^{g}_{in}$, and the workplace productivities $A_j$. The parameters in this model are assumed constant and are: the semi-elasticity of commute costs with respect to commute times ($\kappa$), the residential location migration elasticities for each worker type ($\nu_g$), the parameters that govern the distribution of idiosyncratic workplace match-productivities for each worker type ($T_g$ and $\theta_g$), the time discount factor ($\beta$), the parameters that determine the shape of the production function at each workplace location j, $\{\rho_{jg}\}$, the parameter that determines the shape of the utility function ($\alpha$), and the elasticity of substitution of the CES consumption bundle ($\sigma$). Following Caliendo et al. (2019) I will now define three equilibrium concepts in this model (for a given value of the model parameters), a *temporary equilibrium*, a *sequential equilibrium*, and a *stationary equilibrium*.

**Definition 1** *Given a vector of residential location population $\{L^{g}_{Ri,t}\}$ at time $t$, a pairwise transport time matrix $\{d_{ij}\}$, and a vector of workplace productivities $\{A_j\}$, a* ***temporary equilibrium*** *is defined as a vector of wages for each worker type $g \in \{h, l\}$, $\{w^{g}_{j,t}\}$, and residential floorspace rents $\{r_{Ri,t}\}$ that satisfy the equilibrium conditions of the static sub-problem, which is determined by the labor market clearing condition expressed in equation (15) for each worker type $g \in \{h, l\}$, and the residential floorspace market clearing condition expressed in equation (16).*

Note that in this simple framework, at any time $t$, once the residential location vector is determined, the static problem can be subsumed into finding the vector of wages that equalizes labor demand and labor supply in each workplace location for each worker type, and the vector of rents that equalizes floorspace demand and floorspace supply in each residential location. Once all the wages and rents are determined for period $t$, individual consumption of the numeraire consumption good and of residential floorspace follows from the solution to the worker's consumption problem, and the market for the final consumption good must clear by Walras's Law. Having defined the temporary equilibrium, I now proceed to define the sequential competitive equilibrium for this model given a path of exogenous fundamentals, and given parameter values. Let $m^{g}_{t} = \{m^{g}_{in,t}\}^{I,I}_{i=1,n=1}$, $L^{g}_{Rt} = \{L^{g}_{Ri,t}\}^{I}_{i=1}$, $L^{g}_{Ft} = \{L_{Fj,t}\}^{J}_{j=1}$, $w^{g}_{t} = \{w^{g}_{j,t}\}^{J}_{j=1}$, and $V^{g}_{t} = \{V^{g}_{i,t}\}^{I}_{i=1}$ be the residential location migration shares, residential population, workplace population, wages, and lifetime expected utilities for each worker type $g \in \{h, l\}$ respectively. Following Caliendo et al. (2019) I define the sequential competitive equilibrium in this model as follows.

**Definition 2** *Given an initial distribution of residential population for each worker type $L^{g}_{R0}$, an initial distribution of employment for each worker type $L^{g}_{F0}$, and a known sequence of time-varying and non time-varying fundamentals ($\{d_{ij,t}\}^{I,J,\infty}_{i=1,j=1,t=1}$, $\{A_j\}^{J}_{j=1}$, and $\{\mu^{g}_{in}\}^{I,I}_{i=1,n=1}$), a* ***sequential competitive equilibrium*** *is a sequence of $\{L^{g}_{Rt}, L^{g}_{Ft}, m^{g}_{t}, w^{g}_{t}, V^{g}_{t}\}^{\infty}_{t=0,g\in\{h,l\}}$ that solves equations (6), (7), (10), and the temporary equilibrium at each time $t$.*

We can now define a stationary equilibrium for this model as follows:

**Definition 3** *A* ***stationary equilibrium*** *of this model is a sequential competitive equilibrium such that all the elements of the vector $\{L^{g}_{Rt}, L^{g}_{Ft}, m^{g}_{t}, w^{g}_{t}, V^{g}_{t}\}^{\infty}_{t=0}$ are constant for all $t$ and for all worker skill type $g$.*

[41] The pairwise transport cost $d_{ij}$ is calculated as $\exp(\kappa t_{ij})$, where $\kappa > 0$ is the semi-elasticity of commute costs with respect to commute times.

Note that, as always, in a stationary equilibrium all the fundamentals must be constant for all $t$, since a change in fundamentals would lead to changes in the endogenous variables of the model through time. Also, note that in a stationary equilibrium individual residents may still move from one residential location to another, or from one workplace location to another, in so far as the aggregate populations in each location, and the migration flows between locations for each worker skill type stay constant.

### 5.10 | Model Solution

In this section I will describe how to solve for the full transitional dynamics of this model. The solution relies heavily on Caliendo et al. (2019), with some modifications to allow for the fact that in this model workers do not necessarily work in the same location as where they reside, as well as for the fact that in this model there are two types of workers (high-skilled workers and low-skilled workers). This numerical solution method relies on the fact that we can express the equilibrium conditions in first differences, and by employing this "dynamic hat algebra" we do not have to take a stance on the level of the fundamentals of this economy. Additionally, this approach does not require us to assume that the economy is in a steady state equilibrium at our initial period $t = 0$.

### 5.11 | Equations in First Difference

In this section I will express the equations that govern the transition dynamics of this model (equations (6), (7), (10), and (16)) in first differences. For any time dependent variable $y_t$, denote $\dot{y}_{t+1} \equiv \frac{y_{t+1}}{y_t}$. I will also denote $u^g_{n,t} \equiv \exp(V^g_{n,t})$ to simplify the notation. We can now express the transition dynamics in first difference as follows[42]:

$$\dot{u}^g_{n,t+1} = \left[\frac{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1} + \psi^g_{n,t+1} r^{\alpha-1}_{Rn,t+1}\right)}{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t} r^{\alpha-1}_{Rgn,t} + \psi^g_{n,t} r^{\alpha-1}_{Rn,t}\right)}\right] \left[\sum_{k=1}^{I} m^g_{kn,t} \left(\dot{u}^g_{k,t+2}\right)^{\frac{\beta}{\nu_g}}\right]^{\nu_g}, \tag{17}$$

$$\dot{m}^g_{in,t+1} = \frac{\left(\dot{u}^g_{i,t+2}\right)^{\frac{\beta}{\nu_g}}}{\sum_{k=1}^{I} m^g_{kn,t} \left(\dot{u}^g_{k,t+2}\right)^{\frac{\beta}{\nu_g}}}, \tag{18}$$

$$\dot{r}_{Ri,t+1} = \frac{\sum_{g\in\{h,l\}} L^g_{Ri,t+1}\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgi,t+1} + \psi^g_{n,t+1}\right)}{\sum_{g\in\{h,l\}} L^g_{Ri,t}\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgi,t} + \psi^g_{n,t}\right)}. \tag{19}$$

### 5.12 | Solution Algorithm

I will now summarize the solution algorithm used to calculate the solution to this dynamic discrete choice model.

**1.** Initiate the algorithm at $t = 0$ with a guess for a path of $\{\{\{\dot{u}^{g0}_{t+1}\}^{I}_{n=1}\}^{\infty}_{t=0}\}_{g\in\{h,l\}}$, such

[42] See Appendix C for the derivation of these equations.

that $\dot{u}^{g}_{n,T+1} = 1$ for all T large enough, and for all $n$. Take as given $\{L^{g}_{R0}, L^{g}_{F0}, m^{g}_{-1}, \{d_{ij,t}\}^{I,J,\infty}_{i=1,j=1,t=1}, \{A_j\}^{J}_{j=1}, \{\bar{H}_{Ri}\}^{I}_{i=1}, \{r_{Ri,0}\}^{I}_{i=1}\}$.

2. For all $t \geqslant 0$, use $\{\{\dot{u}^{g0}_{t+1}\}^{I}_{n=1}\}^{\infty}_{t=0}$ to solve for $\{m^{g}_{t}\}^{\infty}_{t=1}$ using equation (18).
3. Use $L^{g}_{R0}$, $\{m_t\}^{\infty}_{t=1}$, and equation (10) to solve for $\{L^{g}_{Rt}\}^{\infty}_{t=0}$.
4. Use $\{L^{g}_{Rt}\}$ and $L^{g}_{F0}$ to estimate model consistent wages for period zero ($t = 0$), $\{w^{g}_{j,0}\}^{J}_{j=1}$ using the labor market clearing condition for each skill type[43], (15).
5. Use $\{L^{g}_{Rt}\}$, $\{w^{g}_{j,0}\}^{J}_{j=1}$, $L^{g}_{F0}$, and vector equation implied by (15) to solve forward[44] for the model consistent $\{L^{g}_{Ft}\}^{\infty}_{t=1}$ and $\{\{w^{g}_{j,t}\}^{J}_{j=1}\}^{\infty}_{t=0}$.
6. Use $\{L^{g}_{Rt}\}$, $\{\Phi_{Rgt}\}$, and equation (19) to solve for $\{\dot{r}_{Rn,t+1}\}$. Then use $\{r_{Ri,0}\}^{I}_{i=1}$ and $\{\dot{r}_{Rn,t+1}\}$ to solve for $\{r_{Rn,t}\}$.
7. For each t, use $\{\{w^{g}_{j,t+1}\}, \{m^{g}_{t}\}, \{d_{ij,t}\}, \{\dot{r}_{Rn,t+1}\}$, and $\{\dot{u}^{g}_{i,t+2}\}\}$ to calculate backwards $\{\dot{u}^{g}_{n,t+1}\}$ using equation (17) for each skill type $g$. This will result in a new sequence $\{\dot{u}^{g1}_{n,t+1}\}^{\infty}_{t=0}$.
8. Verify if $\{\dot{u}^{g1}_{n,t+1}\}^{\infty}_{t=0} \approx \{\dot{u}^{g0}_{n,t+1}\}^{\infty}_{t=0}$, if not, then start the again from step 1, with $\{\dot{u}^{g1}_{n,t+1}\}^{\infty}_{t=0}$ as your new guess. If indeed $\{\dot{u}^{g1}_{n,t+1}\}^{\infty}_{t=0} \approx \{\dot{u}^{g0}_{n,t+1}\}^{\infty}_{t=0}$, then $\{\dot{u}^{g1}_{n,t+1}\}^{\infty}_{t=0}$ is the solution.

## 5.13 | Welfare Gains

The expected utility at time $t$ in location $n$ for a worker of skill type $g$ can be written as[45]

$$V^{g}_{n,t} = \sum_{s=t}^{\infty} \beta^{s-t} \left[ \left( \tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,s} + \psi^{g}_{n,t+1} \right) r^{\alpha-1}_{Rn,s} - \nu_g \ln(m^{g}_{nn,s}) \right]. \tag{20}$$

If we define $\hat{V}^{g}_{n,t}$ as the expected utility under a counterfactual evolution of the model's fundamentals, then we can define the compensating variation in consumption ($\delta^{g}_{n,t}$) as the constant increase in consumption that must be made each period under the counterfactual so as to make an agent of skill type $g$ living in location $n$ at time $t$ indifferent between the counterfactual world and the realized equilibrium. In other words, $\delta^{g}_{n,t}$ is such that

$$V^{g}_{n,t} = \sum_{s=t}^{\infty} \beta^{s-t} \left[ \delta^{g}_{n,t} + \left( \tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,s} + \psi^{g}_{n,t+1} \right) r^{\alpha-1}_{Rn,s} - \nu_g \ln(\hat{m}^{g}_{nn,s}) \right].$$

Solving for $\delta^{g}_{n,t}$, this implies that

$$\delta^{g}_{n,t} = (1-\beta) \sum_{s=t}^{\infty} \beta^{s-t} \left[ \tilde{T}_g \left( \Phi^{\frac{1}{\theta_g}}_{Rgn,s} r^{\alpha-1}_{Rn,s} - \hat{\Phi}^{\frac{1}{\theta_g}}_{Rgn,s} \hat{r}^{\alpha-1}_{Rn,s} \right) + (\psi^{g}_{n,t+1} r^{\alpha-1}_{Rn,s} - \hat{\psi}^{g}_{n,t+1} \hat{r}^{\alpha-1}_{Rn,s}) - \nu_g \ln \left( \frac{m^{g}_{nn,s}}{\hat{m}^{g}_{nn,s}} \right) \right]. \tag{21}$$

By calculating a population-weighted average of the expected utility in each residential location at time $t$, we obtain a measure of average welfare at time $t$,

$$W_t = \sum_{g \in \{h,l\}} \sum_{n=1}^{I} \frac{L^{g}_{n,0}}{\sum_{g \in \{h,l\}} \sum_{i=1}^{I} L^{g}_{i,0}} V^{g}_{n,t}. \tag{22}$$

[43]See Appendix C for an explanation on how to obtain the model-consistent wages as a function of residential and workplace population for each skill type.
[44]See Appendix C for a detailed explanation of this algorithm, as well as a proof of convergence and uniqueness.
[45]See Appendix C for a derivation of this equation.

Where we weight the expected utility in each residential location by the share of the population of that skill type living in that location at time $t = 0$. In the same way, we can calculate the average welfare in the counterfactual world as

$$\hat{W}_t = \sum_{g\in\{h,l\}} \sum_{n=1}^{I} \frac{L^g_{n,0}}{\sum_{g\in\{h,l\}} \sum_{i=1}^{I} L^g_{i,0}} \hat{V}^g_{n,t}. \tag{23}$$

With these measures, we can calculate the proportional gains in average welfare at time t with respect to the counterfactual as simply

$$\Delta W_t = \frac{W_t}{\hat{W}_t} - 1.$$

We can repeat this same exercise by skill type to obtain the average expected utility in each residential location at time t for each skill level g as

$$W^g_t = \sum_{n=1}^{I} \frac{L^g_{n,0}}{\sum_{i=1}^{I} L^g_{i,0}} V^g_{n,t}, \tag{24}$$

and the average welfare by skill type in the counterfactual as

$$\hat{W}^g_t = \sum_{n=1}^{I} \frac{L^g_{n,0}}{\sum_{i=1}^{I} L^g_{i,0}} \hat{V}^g_{n,t}. \tag{25}$$

Comparing these two measures by skill type we can obtain the proportional gains in average welfare at time t for each skill type with respect to the counterfactual as

$$\Delta W^g_t = \frac{W^g_t}{\hat{W}^g_t} - 1. \tag{26}$$

## 6 | PARAMETER ESTIMATION

There are 10 key parameters that must be known in order to solve this model, $\beta$, $\alpha$, $\sigma$, $T_h$, $T_l$, $\kappa$, $\theta_h$, $\theta_l$, $\nu_h$, and $\nu_l$. I assume the bi-annual discount factor, $\beta$, to be 0.92 (implying approximately a 4% real annual interest rate). I also assume the parameter that governs the intra-period utility function, $\alpha$, to be equal to 0.7, which results in a expenditure share in housing of 0.3. I assume that the elasticity of substitution for the CES aggregator of the consumption bundle, $\sigma$ is equal to 6. This number is chosen based on the choice by Tsivanidis (2019) for his Armington elasticity of substation, who, in turn, chooses that number based on the median estimates from Feenstra et al. (2018). I fix $T_l$ to be 1, and calibrate $T_h$ such that the average wage premium in the city at $t = 0$ is equal to the observed wage premium calculated from the National Household Survey for 2010 (See Appendix C for details), which results in $T_h$ being equal to 1.14. I take $\kappa = 0.01$ from Ahlfeldt et al. (2015).

In Subsection 6.1 I explain how I estimate the semi-elasticities of commute flows with respect to commute costs for each skill type ($\theta_h$, and $\theta_l$) by relying on the 2010 commuting survey for the city of Buenos Aires (*ENMODO*). Finally, in Subsection 6.2 I explain how I estimate $\nu_h$ and $\nu_l$, which are the multiplicative inverse of the intra-city migration elasticities with respect to expected income for each skill type.

### 6.1 | Estimating Commuting Semi-Elasticities

In order to calculate $\theta_h$, and $\theta_l$, I use data from the commuting survey for the city of Buenos Aires from 2010 (*ENMODO*). This commuting survey divided the city into 21 zones, and provides individual level data on commuting patterns for a sample of 22,500 households in the metropolitan region of Buenos Aires. Following Ahlfeldt et al. (2015), I take logs of equation 8, and add an error term that reflects the measurement error resulting from the different methods employed to measure commute times in the model and in the survey. This results in

$$\pi^g_{ij} = -\theta_g \times \kappa\tau^g_{ij} + \gamma^g_i + \zeta^g_j + e^g_{ij}, \tag{27}$$

where $\pi^g_{ij} = \log(\Pi_{j|gi0})$ is the natural logarithm of the commuting probability between $i$ and $j$ at time $t = 0$ for a skill type $g$, $\tau^g_{ij}$ is the commute time from $i$ and $j$ for a skill type $g$, $\gamma^g_i$ is a fixed effect at the origin level for a skill type $g$, $\zeta^g_j$ is a fixed effect at the destination level for a skill type $g$, and $e^g_{ij}$ is the error term. Aggregating this equation to the zone level results in an approximate equation (ignoring the Jensen inequality term) of the form

$$\pi^g_{IJ} = -\theta_g \times \kappa\tau^g_{IJ} + \gamma^g_I + \zeta^g_J + e^g_{IJ}. \tag{28}$$

I will estimate this equation using the bilateral commuting flows from the commuting survey in 2010. I estimate this equation using both a linear fixed effects estimator, as well as with a Poisson Pseudo Maximum Likelihood estimator, to address the concerns regarding the granularity of the commuting data (Dingel et al., 2020). Table 7 shows the results of these estimations for each skill type. From columns 2 and 4 we can see that $\theta_h \times \kappa = -0.029$, and $\theta_l \times \kappa = -0.038$. Using the fact that $\kappa = 0.01$, we find that $\theta_h = 2.9$, and $\theta_l = 3.8$. These results are very similar to the ones estimated by Tsivanidis (2019), who finds a $\theta_h$ equal to 2.7 and a $\theta_l$ equal to 3.3 using commuting data from Bogota.

### 6.2 | Estimating Intra-City Migration Elasticities

I will estimate the intra-city migration elasticities by skill type ($1/\nu_h$ and $1/\nu_l$) by adapting the estimation method used in Caliendo et al. (2019), which is in itself an adapted version of Artuç et al. (2010). I show that we can derive an estimating equation from this model that relates current differences in migration flows to future differences in expected income, as well as future differences in migration flows. I then estimate this equation using an IV-GMM PPML estimator (Windmeijer and Santos Silva, 1997), using past migration flows by skill type to instrument for future migration flows, and lagged differences in commuter market access to instrument for the differences in commuter market access. The estimating equation[46] (which closely resembles the specification in Caliendo et al. (2019) but with expected income instead of future wages) is

$$\log\left(\frac{m^g_{in,t}}{m^g_{nn,t}}\right) = -\mu^g_{in}\frac{1-\beta}{\nu_g} + \frac{\beta}{\nu_g}\left[\tilde{T}_g\left(\Phi^{\frac{1}{\theta_g}}_{Rgi,t+1} r^{\alpha-1}_{Ri,t+1} - \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)\right] + \beta\log\left(\frac{m^g_{in,t+1}}{m^g_{ii,t+1}}\right) + \omega^g_{t+1}. \tag{29}$$

[46] See Appendix C for the derivation of this equation.

Following Caliendo et al. (2019) and Artuç et al. (2010), I estimate this equation under the assumption that the average moving costs are equal for all location pairs $(i, j)$, so that $\mu^g_{ij} = \mu \, \forall (i, j) \in N \times N$, which allows me to define $\tilde{C}_g = -\mu \frac{1-\beta}{\nu_g}$, and obtain the following equation:

$$\log\left(\frac{m^g_{in,t}}{m^g_{nn,t}}\right) = \tilde{C}_g + \frac{\beta}{\nu_g}\left[\tilde{T}_g\left(\Phi^{\frac{1}{\theta_g}}_{Rgi,t+1} r^{\alpha-1}_{Ri,t+1} - \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)\right] + \beta \log\left(\frac{m^g_{in,t+1}}{m^g_{ii,t+1}}\right) + \omega^g_{t+1}. \quad (30)$$

Where $\tilde{C}_g$ is a constant and $\omega^g_{t+1}$ is a random term that includes any unexpected shock at time $t + 1$ that affects the relative migration flows at time $t + 1$. Estimating this equation for each skill type using bi-annual intra-city migration flows constructed from the individual-level data on residential locations obtained from the electoral register leads to an estimate of $\nu_h$ equal to 1.4, and a $\nu_l$ equal to 1.8. This implies that high-skilled workers have a higher intra-city migration elasticity than low-skilled workers (since the migration elasticity is the multiplicative inverse these parameters).

## 7 | WELFARE RESULTS

Having estimated the necessary elasticities and key parameters in the model, I use the algorithm describe in Subsection 5.12 to solve numerically for the sequential competitive equilibrium (as defined in Subsection 5.9) that converges to a stationary equilibrium in 25 periods, which is equal to 50 years, since I choose each period to be equal to two years, in order to match the bi-annual individual-level migration data. I use data on employment and residential population by district[47] in 2011 to initialize the employment and population vectors, I then use the migration matrix estimated from individual level data at the district level between 2011 and 2013 to initialize the migration matrix in the model at $t = 0$. Finally, I calculate minimum commute times using ArcMaps's Network Analyst Tool for each period by adding sequentially each line that appeared in each bi-annual period, with these commute times I can calculate $\{d_{ij,t}\}$ for every $t = 0, 1, 2, 3, ...$[48] When estimating a counterfactual scenario, I will use different commute times, calculated in the same way, under different assumptions regarding the changes in the transport infrastructure. These changes in the vector of time-varying fundamentals will lead to different sequential competitive equilibria and stationary equilibria.

In this section I will use the estimated quantitative model to analyze the impact on resident's welfare of the construction of the BRT system in Buenos Aires. One of the key contributions of this dynamic framework is that it allows me to quantify the average welfare gains for residents based on where they were living before the BRT system was built (which I refer to as "incumbents"), as opposed to doing comparative statics on the *ex post* equilibrium. In the following subsection I study the welfare gains by skill type for incumbents, and how these gains differ for different parts of the city. I then study the aggregate welfare gains by skill type in the entire city, as well as the average gains for people living close to the BRT lines at $t = 0$ and at the *ex post* stationary equilibrium.

[47] In order to alleviate concerns regarding the sparsity of the migration matrix at the census tract level, I solve the model at a more aggregate spatial level, the electoral district. The main results hold when the model is solved at the census tract level (see Appendix D).

[48] I will assume that $d_{ij,t} = d_{ij,3}$ for all $t \geqslant 4$.

### 7.1 | Welfare Gains for Incumbents

Let $V^g_{n,0}$ be the expected utility of a resident of skill type $g$ living in location $i$ at time $t = 0$ in the scenario where the full BRT system will be built, as calculated in equation 20. Let us also calculate $\hat{V}^g_{n,0}$ as the expected utility at time $t = 0$ under the assumption that no BRT line will be built. Since residents are forward-looking (and have perfect foresight), these expected utilities will include the expected future utility flows under each scenario, and will, therefore, be different. By calculating $V^g_{n,0}/\hat{V}^g_{n,0} - 1$ we can measure the welfare gains in terms of expected utility for residents according to their place of residence before the BRT was put in place.

When analyzing the spatial distribution of the average welfare gains for incumbents by district (Figure 15), there are three findings that stand out. First, welfare gains were higher for both high- and low-skilled incumbents living closer to the BRT lines and further from the CBD, since these are the locations where commuter market access increased the most. This is especially true for the suburbs outside the city proper. Second, welfare gains for high- and low-skilled incumbents are highly spatially correlated. In fact, the Pearson correlation coefficient between high- and low-skilled gains by district is 0.96. Third, welfare gains for both skill types seem larger for incumbents living near the southernmost BRT line than for incumbent living near the northernmost BRT line, with incumbents living near the other BRT lines seeing an intermediate level of welfare gains on average.

From Figure 3 we see that different BRT lines cross through neighborhoods with different levels of high-skilled share at baseline. In particular, census tracts near (less than one kilometer) from the first, second, and fifth lines built have an average high-skilled share that is close the median,[49] while census tracts near the third line built (the southernmost line) have, on average, much lower levels of high-skilled share,[50] and the census tracts near the fourth line built (the northernmost line) have the highest high-skilled share on average (Table 8).[51] With this information, I group the lines intro three subsets: the southern line (third line to be built) constitutes one group, and represents the line built along neighborhoods with the lowest levels of high-skilled share in 2010; the first, second, and fifth lines to be built constitute the second group, and represent the lines that were built along neighborhoods with a medium level of high-skilled share; and the northern line (fourth to be built) constitutes the third group, and represents the line built along neighborhoods with the highest high-skilled share in the city.

We can now compare the average welfare gains for incumbent residents living near[52] each group of lines (Figure 16), and divide the welfare gains for the residents that live further away between those living in the suburbs (in the Buenos Aires Province) and those living in the city proper. From Figure 16 we can see that the largest gains are perceived by high-skilled residents of the suburbs living further than one kilometer from the BRT lines. We also see that high-skilled residents in general gain substantially more than low-skilled residents at the same distance to a line. This difference is particularly stark in the case of the line with the high high-skilled share neighborhoods, where low-skilled residents are worse off due to the BRT lines. These welfare losses are consistent with the fact that a lower share of low-skilled workers are homeowners in more affluent neighborhoods, where the share of high-skilled workers is high. Since this is the case, when market access increases, this drives home prices to increase, which in turn displaces non-home-owning low-skilled residents.

[49] The median census tract had a high-skilled share of 0.5 according to the 2010 census.

[50] The average census tract that is at most at one kilometer from the southern line has a high-skilled share of 0.31, placing it at the thirteenth percentile of the distribution of high-skilled share in the city in 2010.

[51] The average census tract near the northern line had a high-skilled share of 0.63 in 2010, placing it at the seventy-third percentile of the distribution of high-skilled share in the city.

[52] At most at one kilometer from a BRT line.

As an alternative way of studying the difference in welfare gains by initial level of high-skilled share, I divide all residential locations into four groups according to the four quartiles of the initial high-skilled share distribution. Figure 17 shows the average welfare gains for high- and low-skilled incumbent residents for each quartile of the distribution of initial high-skilled shares. As we can see, the largest gains were perceived by high-skilled residents that were initially living in the bottom quartile of the distribution of high-skilled shares. For both high- and low-skilled residents, welfare gains are decreasing in the initial share of high-skilled residents.

There are four main forces contributing to these welfare results. First, I find that high-skilled workers have a lower elasticity of commuting with respect to commute costs, which is consistent with previous findings in the literature (Tsivanidis, 2019). Because of this lower elasticity, high-skilled workers bare a larger incidence of any changes in market access, which in this case is a force that tends to increase the welfare gains for high-skilled residents relative to low-skilled residents.

A second factor that shapes the geographical patterns of welfare gains is the fact that, because of the specific placement of the BRT lines, as well as the distribution of high- and low-skilled population in the city, the locations that saw the highest increases in market access due to the construction of the BRT had, on average, lower levels of high-skilled share. Since larger increases in market access tend to translate into larger welfare gains, this correlation between increases in market access and the initial level of high-skilled share will result in higher welfare gains for incumbent residents of locations with lower initial levels of high-skilled share, on average.

Finally, there are two factors related to the housing market that contribute to shaping the spatial distribution of welfare gains. First, as we can see from Table 9, there is a negative correlation between home ownership rates and the initial level of high-skilled share. This is true both for high- and low-skilled residents. Since higher home ownership rates imply that the incumbent residents will, on average, gain more from increases in real estate prices, this correlation implies that, for the same level of rent increase, locations with lower high-skilled share will, on average, see higher welfare gains. A second factor related to the housing market is that, as shown in the reduced form analysis, housing prices increased relatively more in locations with higher initial levels of high-skilled share, which, on average, tended to reduce the welfare gains from the construction of the BRT system.

### 7.2 | Aggregate Welfare Results

Estimating $\Delta W_t^g$ from equation 26 at $t = 0$ for high- and low-skilled residents, I find that high-skilled workers gained significantly more than low-skilled workers on average from the BRT. The aggregate welfare gains at $t = 0$ from the construction of the BRT system were 1% for high-skilled residents and 0.6% for low-skilled residents. Additionally, I find that the *Metrobus* resulted in a net present value gain (net of construction costs) of $13.26 billion measured in 2010.[53] This is equivalent to a constant increase in the 2010 GDP of the City of Buenos Aires of 0.4%.

[53] See Appendix D for details on this calculation. These calculations do not take into account the suburbs because of a lack of output data at the sub-provincial level. As such, they should be interpreted as a lower bound of the GDP gains from the construction of the Metrobus BRT.

# 8 | COUNTERFACTUALS: LOCAL TRANSPORT INFRASTRUCTURE IMPROVEMENTS

In this section I study what would have happened to spatial sorting between high- and low-skilled residents if only a subset of lines were built, as opposed to the full BRT system. I construct three counterfactual scenarios: in the first scenario, only the southern line, that goes through the districts with the lowest high-skilled share in 2010, is built; in the second scenario, the three lines that go through districts that are close to the median value of the high-skilled distribution in 2010 (the first, second, and fifth lines to be built) are the only lines to be built; in the third scenario, only the northern line, that serves the districts with the highest high-skilled share on average, is built. I study the consequences of building these three different subsets of BRT lines on spatial sorting, and on welfare for incumbents that live near each group of BRT lines.

## 8.1 | Welfare Gains in Each Counterfactual Scenario

Interestingly, when we compare the gains for incumbents that were living near the southern line (Figure 18) between the scenario where only that line was built, and the scenario were all BRT lines were built, we can see that the welfare gains are 2.4 times larger for high-skilled residents and 1.7 time larger for low-skilled residents in the case where all the BRT lines are built. By comparison, the welfare gains for low-skilled incumbent residents living near the northern BRT line (with a high high-skilled share) would have been significantly better off if only that line had been built (Figure 20). They would have gained 0.08% as opposed to losing 0.18% in welfare when the entire system is built. This difference is likely due to the fact that when the entire system is built, housing prices increase in areas further away from the northern BRT line, which makes it more costly for low-skilled workers to move to another part of the city. By only building one line, the fraction of low-skilled workers that end up being displaced by the increases in housing prices have a relatively higher value option of moving, while at the same time, the residents that stay benefit from the increased market access. Finally, we can see that in the case where only the BRT lines that run through neighborhoods with a medium level of high-skilled share on average were built, the high-skilled incumbents would have been worse off by about 24%, but the low-skilled incumbents would have gone from a modest 0.1% increase in welfare to a decrease in welfare of 0.05%.

# 9 | CONCLUSION

In this paper, I make three contributions to the understanding of the effects of improving the urban transit infrastructure. First, I show that migration response to these improvements by high- and low- skilled workers depend on the initial characteristics of the neighborhoods that are being targeted by these improvements. For the same level of increase in market access, neighborhoods with an initial proportion of high-skilled share at the eightieth percentile of the distribution see increases in the high-skilled share that are, on average, eight times those of neighborhoods at the twentieth percentile of the distribution. This differential effect led to an increase in spatial segregation in the city.

Second, I develop a dynamic quantitative spatial equilibrium framework that allows us to quantify the welfare effects of these transport projects on the residents that were living near these improvements before they were put in place. This is particularly significant because transport infrastructure projects are a form of place based policy, and as such, policy makers chose the placement of these projects in order to target the residents that are living in the neighborhoods that are closer to these transport improvements. Therefore, being able to quantify the welfare gains of these incumbent residents is of direct policy relevance. In the case of Buenos Aires, I find that the BRT system increased welfare significantly more for high-skilled residents than for low-skilled residents. I also find that the surrounding suburbs saw the larges gains from the Metrobus, which is consistent with the fact that the BRT lines connect the buses leading from the suburbs to the city center. As a consequence, residents commuting from further away by bus will see larger gains in terms of time saved, which will translate in larger gains in terms of increases in market access.

Third, I study the counterfactual welfare effects for incumbent residents living near the BRT lines of more localized transport infrastructure improvements. Due to budgetary and political considerations, governments might decide to prioritize transport improvements in poorer areas, where the share of low-skilled workers is higher. I show that, in the case of Buenos Aires, building a single line that ran through the neighborhoods with the highest proportion of low-skilled workers would have still benefited incumbent residents living near the line, but the welfare gains would had been significantly lower.

## 10 | FIGURES

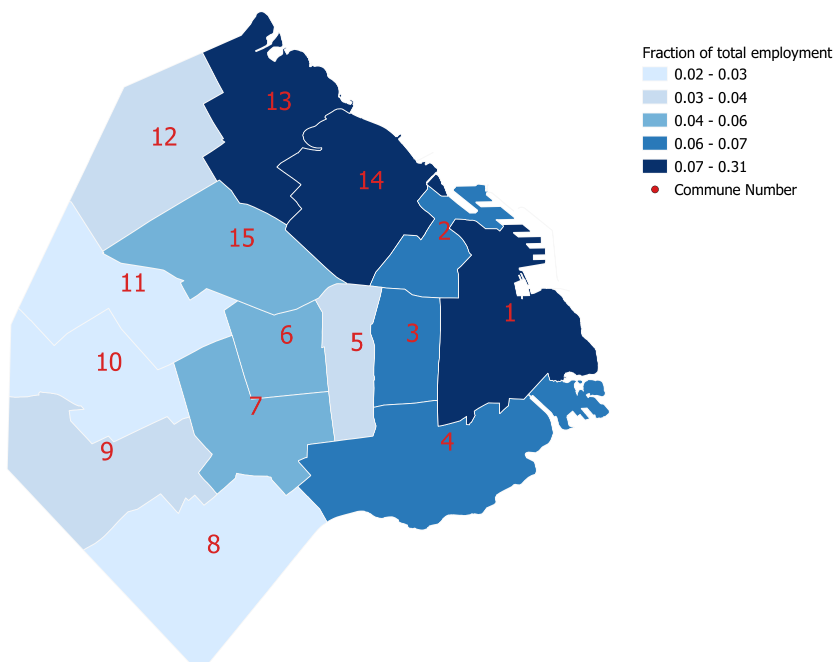


FIGURE 1 Fraction of total employment population by commune.

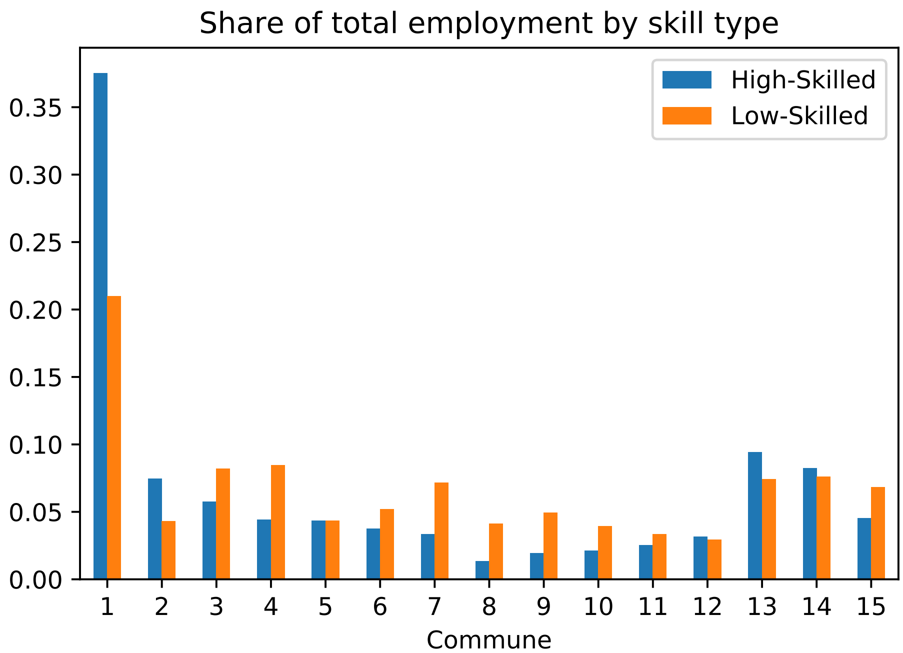


FIGURE 2 Share of total employment by skill type for each commune.

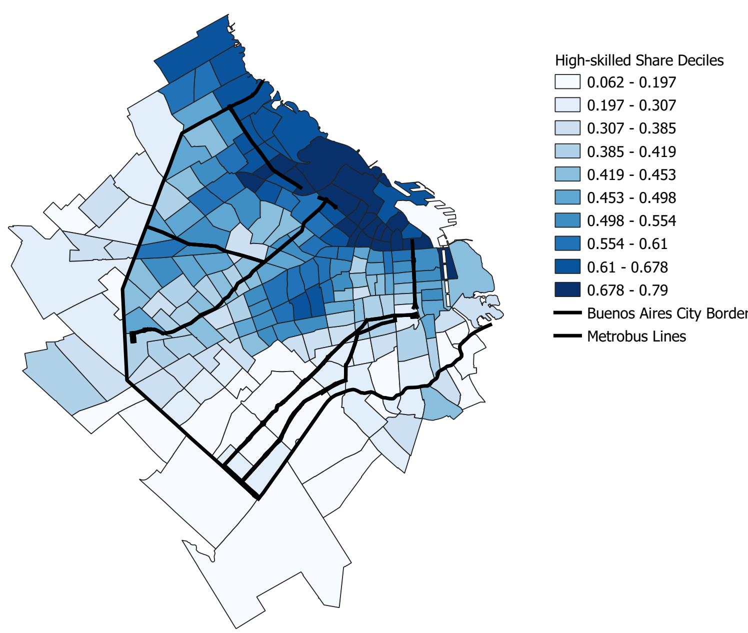


FIGURE 3 Share of high-skilled residents living in each district in 2011.

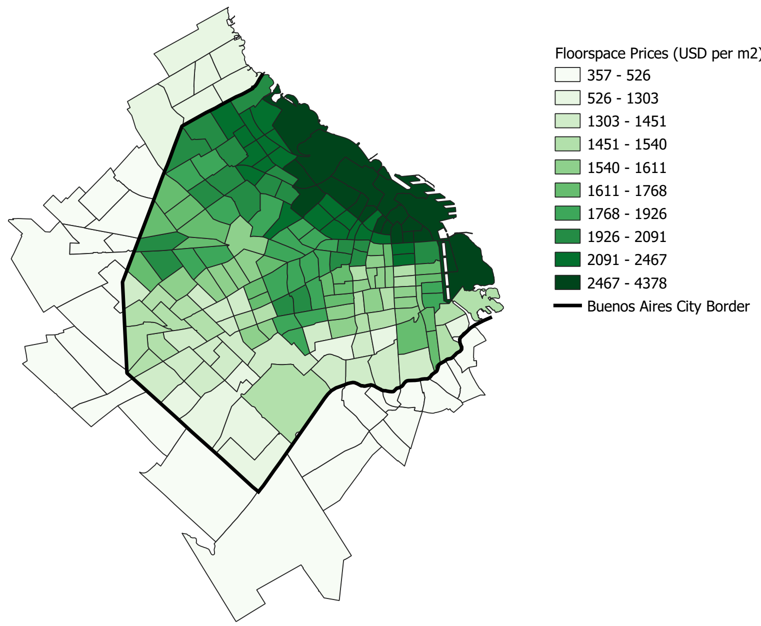


FIGURE 4 Percentiles of floorspace prices, calculated as the average sale price per square meter, by district in 2011.


FIGURE 5 Subway lines built up to 2010 in the city of Buenos Aires.


FIGURE 6 *Metrobus* lines built between 2011 and 2017 in the city of Buenos Aires, numbered by the order in which they were constructed.

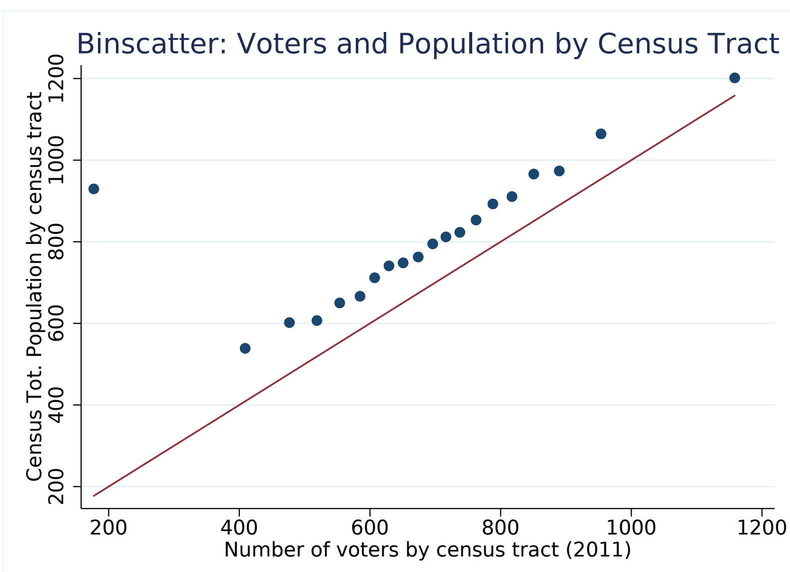


FIGURE 7 Binscatter of the 2010 census population by census tract compared to the number of geocoded voters by census tract for the city of Buenos Aires. The red line is a 45 degree line.

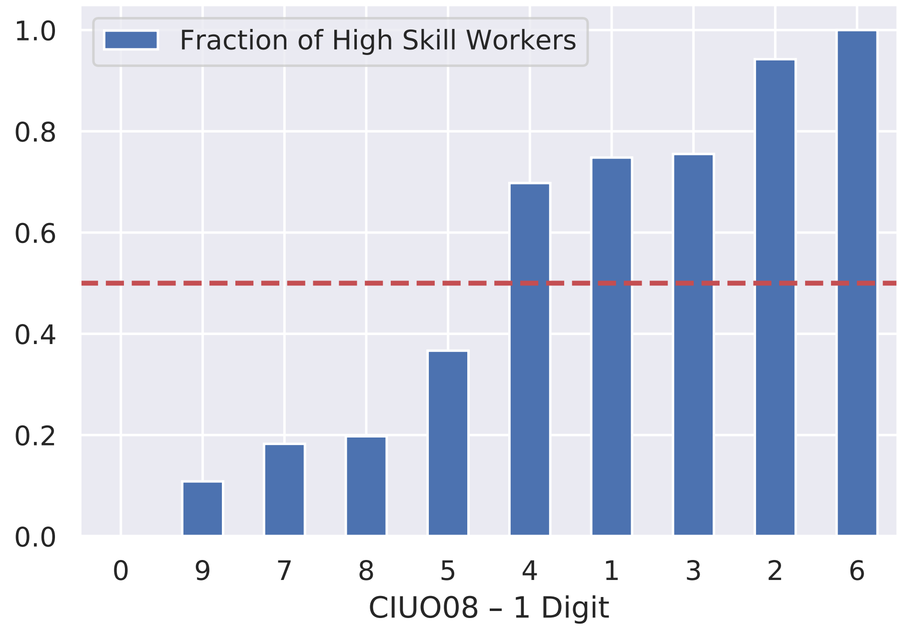


FIGURE 8 High-skill share by occupations at 1 digit ISCO classification in 2011. Calculated from the first-trimester wave of the National Household Survey (*Encuesta Permanente de Hogares*) in 2011.

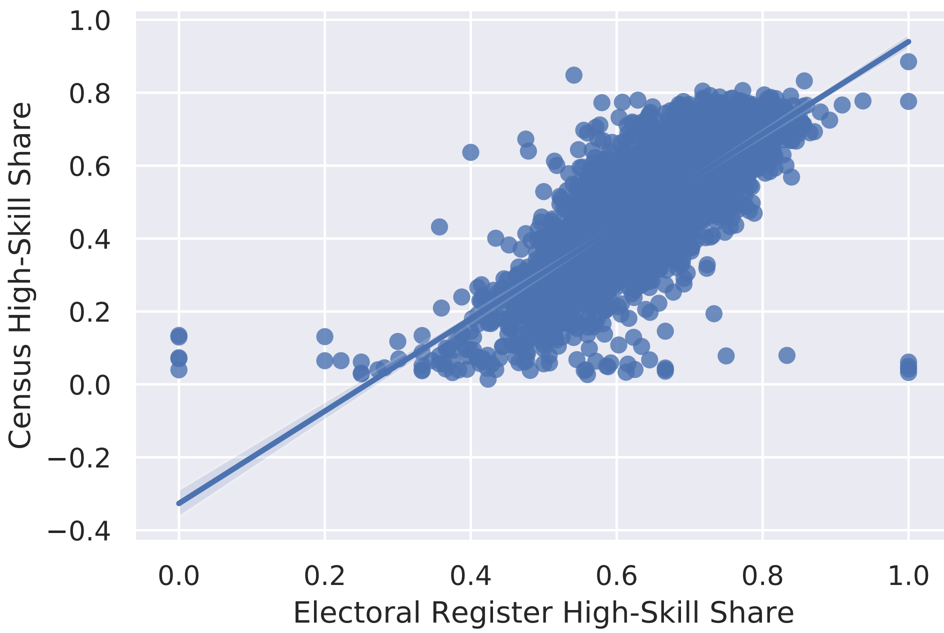


FIGURE 9 High-skilled share in 2010 census versus electoral register data in 2011. Regression at the census tract level of high-skilled share observed in the 2010 census on the high-skilled share measure constructed from the individual level data from the 2011 electoral register. The slope of the regression line is 1.2, and is significant at a 1% level. The R-squared of the regression is 0.6.

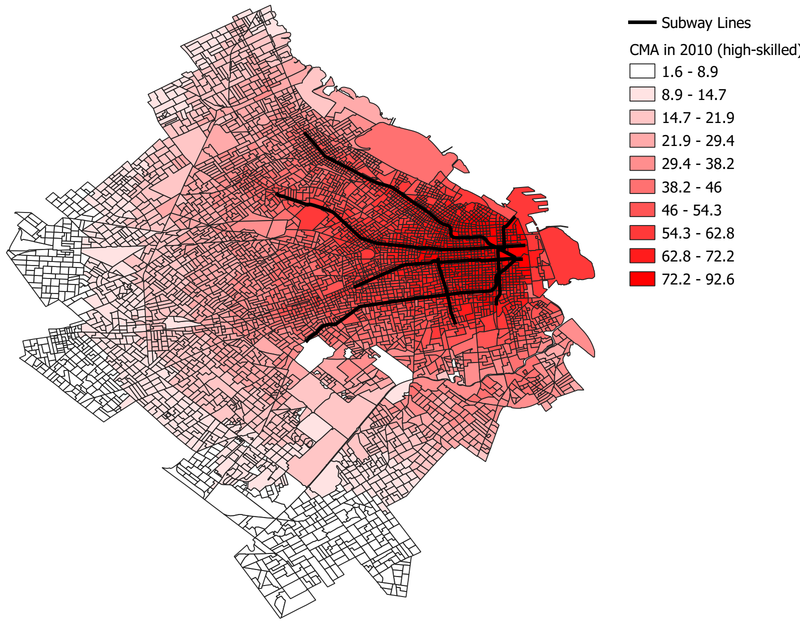


FIGURE 10 Commuter Market Access of high-skilled workers in 2011.

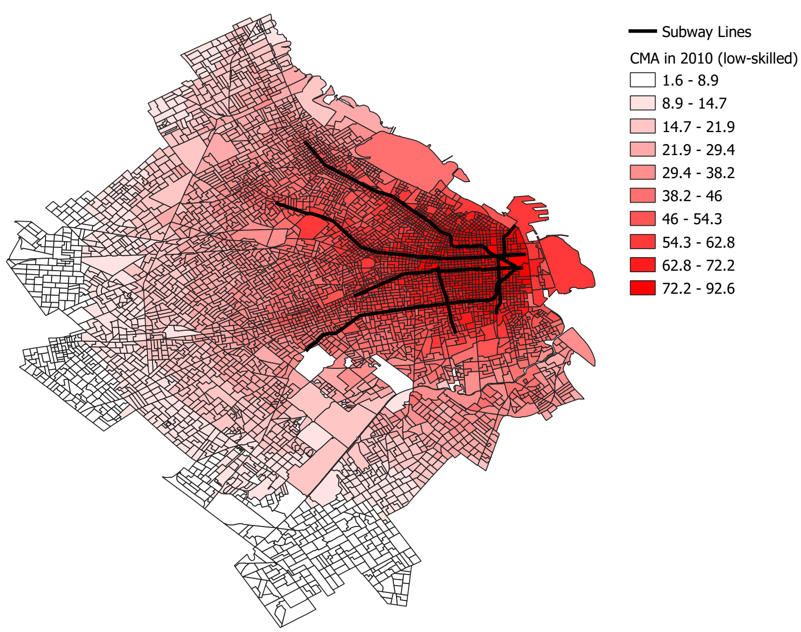


FIGURE 11 Commuter Market Access of low-skilled workers in 2011.

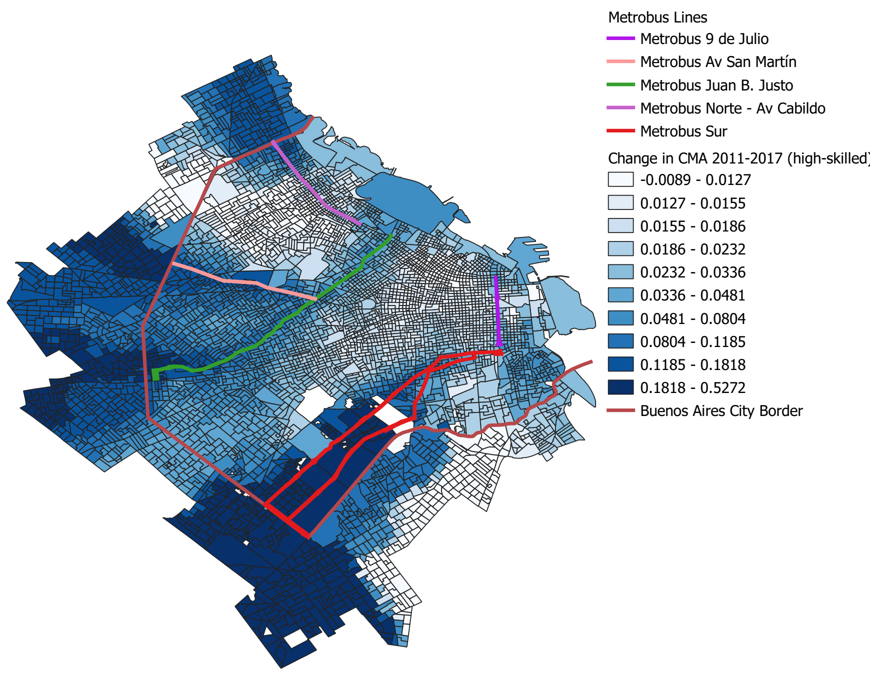


FIGURE 12 Proportional change in Commuter Market Access for high-skilled workers between 2011 and 2017.

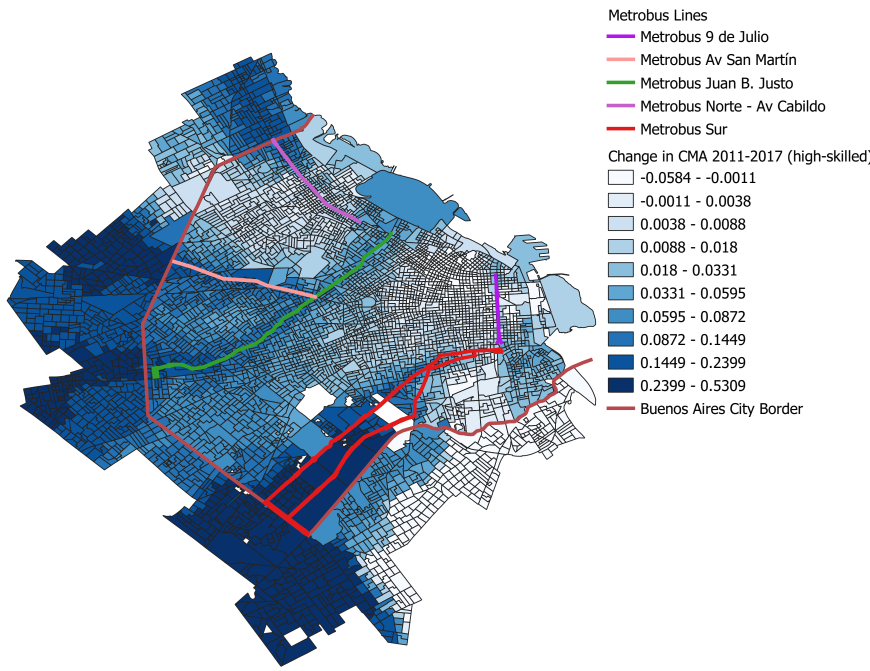


FIGURE 13 Proportional change in Commuter Market Access for low-skilled workers between 2011 and 2017.

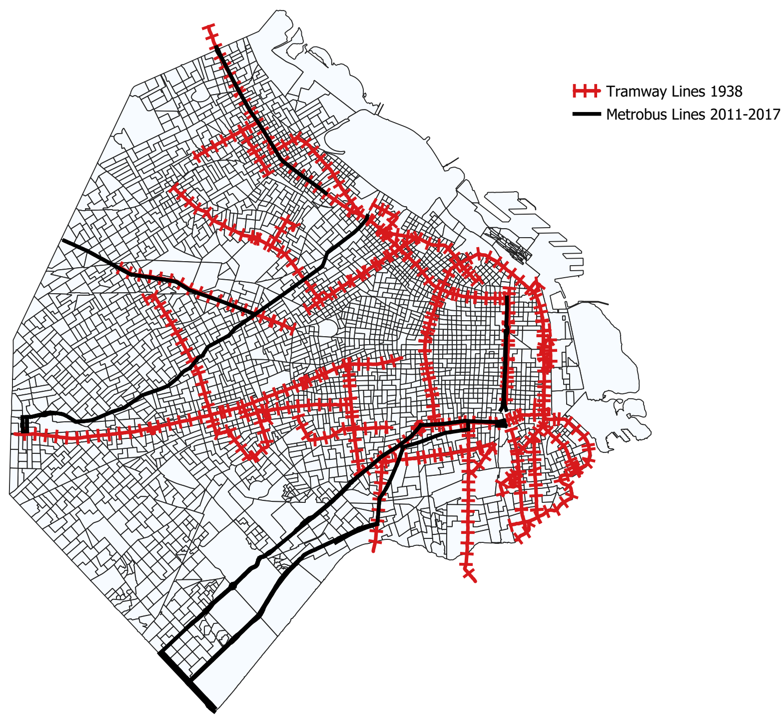


FIGURE 14 Tramway lines running through two-way avenues in 1938, and *Metrobus* BRT lines built up to 2017.

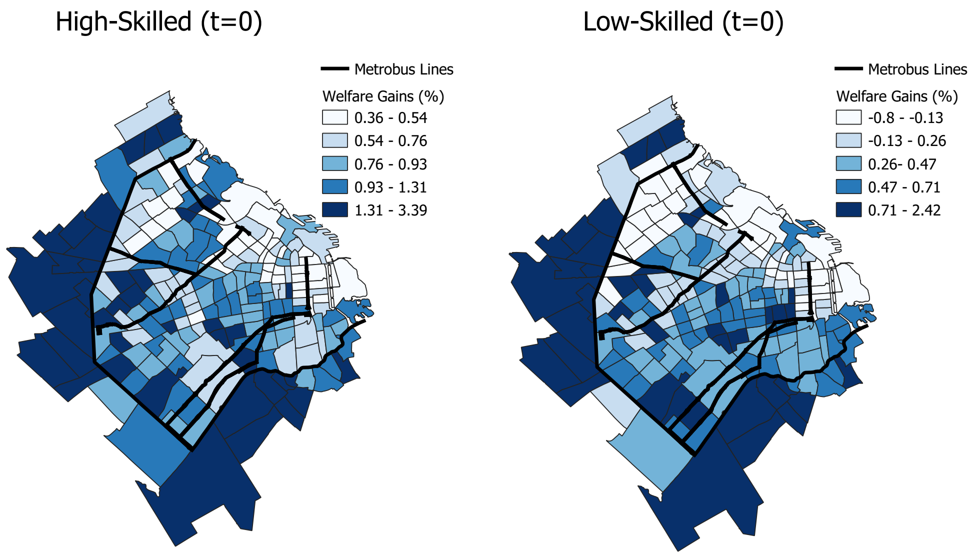


FIGURE 15 Average welfare gains (in percentage terms) for high-skilled and low-skilled incumbents at t = 0 by district.

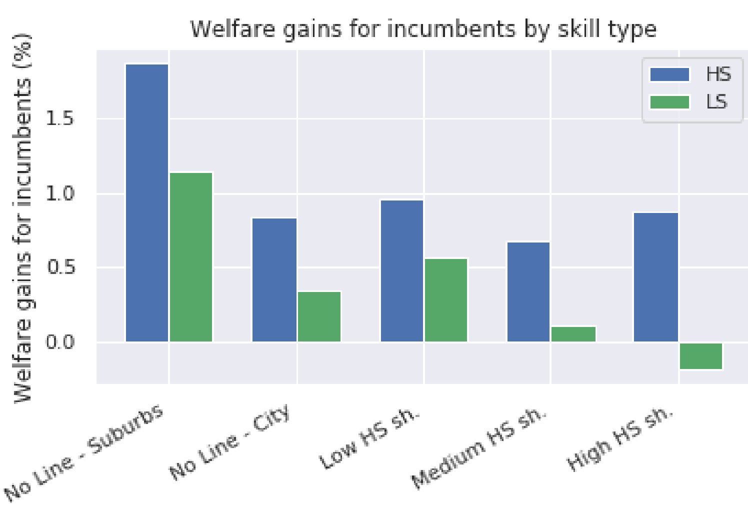


FIGURE 16 Average welfare gains for incumbent residents at t = 0 by line group and skill type.

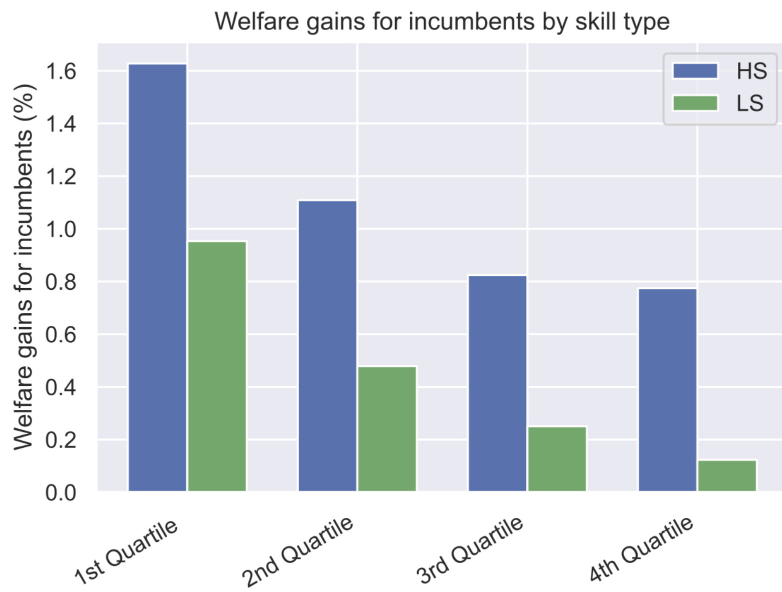


FIGURE 17 Average welfare gains for incumbent residents at $t = 0$ by skill type and quartile of the initial high-skilled share distribution of their residential location.

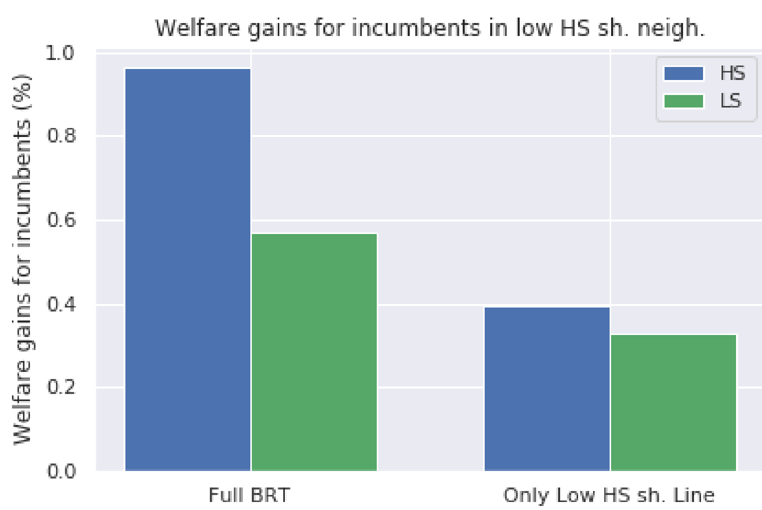


FIGURE 18 Comparison of welfare gains for each skill type between model with full BRT, and model where only the southern line that goes through the districts with the lowest high-skilled share is built.

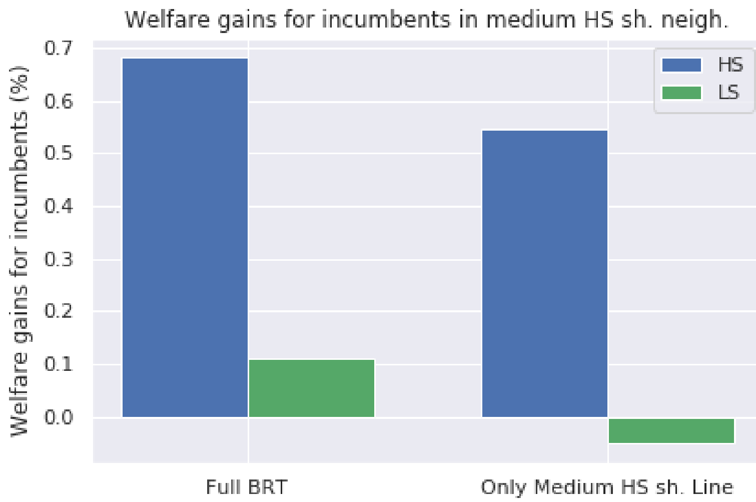


FIGURE 19 Comparison of welfare gains for each skill type between model with full BRT, and model where only the lines that go through the districts with the medium level of high-skilled share are built.

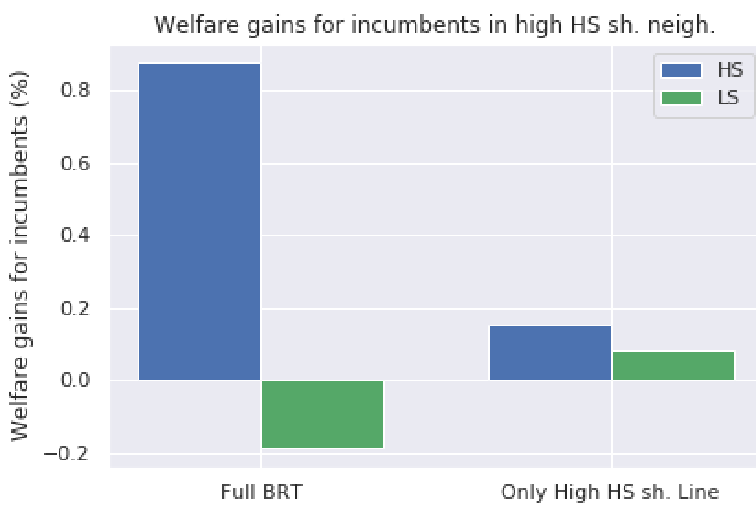


FIGURE 20 Comparison of welfare gains for each skill type between model with full BRT, and model where only the line that goes through the districts with the highest high-skilled share is built.

## 11 | TABLES

Some examples that demonstrate use of syntax and format, follow.

TABLE 1 Percentage of trips made by mode of transport by skill type

| Mode of Transport | Low-Skilled | High-Skilled | Total |
|---|---|---|---|
| Car / motorcycle | 16.0% | 23.1% | 18.3% |
| Taxi | 1.7% | 3.0% | 2.1% |
| Bus | 41.3% | 42.5% | 41.7% |
| Subway | 2.4% | 11.5% | 5.3% |
| Train | 2.9% | 5.3% | 3.6% |
| Walking / bicycle | 35.3% | 14.6% | 28.7% |
| Other | 0.4% | 0.0% | 0.3% |

*Notes:* Percentages calculated from the 2010 mobility survey for the metropolitan region of Buenos Aires (*Encuesta de Movilidad Domiciliaria*, or *EnMoDo*), taking all trips with origin within the City of Buenos Aires and surrounding districts (*partidos* in Spanish) in the Province of Buenos Aires, for people of age 18 or above. Any person with at least some post-secondary education was categorized as high-skilled, and every person with with no more than secondary education was categorized as low-skilled. The *partidos* from the Buenos Aires province that were included are: *Avellaneda, General San Martín, La Matanza, Lanús, Caseros, Vicente López*, and *Lomas de Zamora*.

TABLE 2 International Standard Classification of Occupations (ISCO) one-digit occupation groups

| Code | Occupation Description |
|---|---|
| 1 | Managers |
| 2 | Professional |
| 3 | Technicians and associate professionals |
| 4 | Clerical support workers |
| 5 | Service and sales workers |
| 6 | Skilled agricultural, forestry and fishery workers |
| 7 | Craft and related trades workers |
| 8 | Plant and machine operators, and assemblers |
| 9 | Elementary occupations |
| 0 | Armed forces occupations |

TABLE 3 First stage regressions with historical IV instrument

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | $\Delta\log(CMA)$ | $\Delta\log(CMA)$ | $\Delta\log(CMA) \times$ hs sh. avg.$_0$ | $\Delta\log(CMA)$ | $\Delta\log(CMA) \times$ hs sh. avg.$_0$ |
| $\Delta\log(CMA_{inst})$ | 0.524*** | 0.695*** | 0.0930*** | 0.695*** | 0.0930*** |
| | (0.0213) | (0.0382) | (0.00907) | (0.0382) | (0.00907) |
| hs share$_0$ | 0.109*** | 0.228*** | 0.147*** | 0.228*** | 0.147*** |
| | (0.0100) | (0.0242) | (0.00574) | (0.0242) | (0.00574) |
| $\Delta\log(CMA_{inst}) \times$ hs share$_0$ | | -0.778*** | 0.0288 | -0.778*** | 0.0288 |
| | | (0.144) | (0.0343) | (0.144) | (0.0343) |
| Constant | -0.0196*** | -0.0430*** | -0.0253*** | -0.0430*** | -0.0253*** |
| | (0.00246) | (0.00499) | (0.00118) | (0.00499) | (0.00118) |
| F - stat | 604.67 | 100.04 | 90.90 | 100.04 | 90.90 |
| N | 3731 | 3731 | 3731 | 3731 | 3731 |

**Notes:** Each column shows a regression of one of the endogenous regressors from Tables 4, 5, and 6 on the instrumented changes in market access, where the instrument is constructed assuming the BRT system was built following the 1938 tramway lines, as well as the instrumented market access interacted with the initial share of high-skilled residents. Cragg-Donald Wald F statistics are reported for each regression. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

TABLE 4 Reduced form results: high skill share

| | $\Delta \log(\text{hs share})$ | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| | OLS | IV | OLS | IV |
| $\Delta \log(CMA)$ | -0.119 | 0.270 | -0.571 | -0.508 |
| | (0.059)** | (0.091)*** | (0.096)*** | (0.295)* |
| hs $\text{share}_0$ | 0.703 | 0.633 | 0.413 | 0.098 |
| | (0.017)*** | (0.021)*** | (0.034)*** | (0.208) |
| $\Delta \log(CMA) \times$ hs sh. $\text{avg.}_0$ | | | 2.261 | 4.103 |
| | | | (0.266)*** | (1.517)*** |
| cons | -0.450 | -0.457 | -0.399 | -0.364 |
| | (0.003)*** | (0.004)*** | (0.007)*** | (0.035)*** |
| N | 3,731 | 3,731 | 3,731 | 3,731 |

**Notes:** regressions of change in high skill share between 2011 and 2017 on change in market access between 2011 and 2017, initial average high skill share for contiguous census tracts, and the interaction term. The instrument used for the IV regressions is a measure of change of log CMA where the 2017 CMA is calculated assuming the BRT lines were built following the 1938 tramway lines. Cragg-Donald Wald F statistics from the first stage regression are reported for each IV estimation. Observations are at the census tract level for all census tracts at a distance of 2500 meters of a metrobus line or less. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

TABLE 5 Reduced form results: floorspace prices

| | $\Delta\log$(floorspace price) | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| | OLS | IV | OLS | IV |
| $\Delta\log(CMA)$ | -0.181 | 0.158 | -0.375 | -1.099 |
| | (0.090)** | (0.161) | (0.160)** | (1.899) |
| hs $\text{share}_0$ | 0.810 | 0.749 | 0.685 | -0.665 |
| | (0.039)*** | (0.047)*** | (0.064)*** | (0.815) |
| $\Delta\log(CMA)\times$ hs sh. $\text{avg.}_0$ | | | 0.973 | 7.188 |
| | | | (0.452)** | (11.964) |
| cons | 0.018 | 0.013 | 0.040 | 0.948 |
| | (0.006)*** | (0.006)** | (0.012)*** | (0.249)*** |
| Comune/Partido FE | NO | NO | NO | YES |
| N | 3,731 | 3,731 | 3,731 | 3,731 |

**Notes:** regressions of change in floorspace prices between 2011 and 2017 on change in market access between 2011 and 2017, initial average high skill share for contiguous census tracts, and the interaction term. The instrument used for the IV regressions is a measure of change of log CMA where the 2017 CMA is calculated assuming the BRT lines were built following the 1938 tramway lines. Floorspace prices are calculated as the average sale price per square meter, according to geo-referenced online asking price data for the city of Buenos Aires proper, and from the average land prices from Secretaría de Planeamiento (2012) and Secretaría de Planeamiento (2017) for the suburbs in the Buenos Aires Province. Observations are at the census tract level for all census tracts at a distance of 2500 meters of a metrobus line or less. Cragg-Donald Wald F statistics from the first stage regression are reported for each IV estimation. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

TABLE 6 Reduced form results: population changes

| IV regressions | | | |
|---|---|---|---|
| | $\Delta \log(\text{Tot. pop.})$ | $\Delta \log(\text{HS pop.})$ | $\Delta \log(\text{LS pop.})$ |
| $\Delta \log(CMA)$ | -0.496 | -0.892 | 0.160 |
| | (0.369) | (0.499)* | (0.314) |
| hs $\text{share}_0$ | -0.661 | -1.026 | -0.170 |
| | (0.276)** | (0.380)*** | (0.208) |
| $\Delta \log(CMA) \times$ hs sh. $\text{avg.}_0$ | 4.645 | 6.906 | 1.529 |
| | (2.011)** | (2.760)** | (1.535) |
| cons | -0.251 | -0.159 | -0.408 |
| | (0.046)*** | (0.064)** | (0.036)*** |
| N | 3,731 | 3,731 | 3,717 |

**Notes:** regressions of change in high-skilled population (column 1), low-skilled population (column 2), and total population (column 3) between 2011 and 2017 on change in market access between 2011 and 2017, initial average high skill share for contiguous census tracts, and the interaction term. The instrument used for the IV regressions is a measure of change of log CMA where the 2017 CMA is calculated assuming the BRT lines were built following the 1938 tramway lines. Observations are at the census tract level for all census tracts at a distance of 2500 meters of a metrobus line or less. Cragg-Donald Wald F statistics from the first stage regression are reported for each IV estimation. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

TABLE 7 Bilateral commuting probabilities

| | ln Bilateral Commuting Prob. 2010 | | | |
|---|---|---|---|---|
| | High-skilled | | Low-skilled | |
| trip time | -0.037 | -0.029 | -0.047 | -0.038 |
| | (0.014)*** | (0.012)** | (0.009)*** | (0.009)*** |
| Estimation | OLS | PPML | OLS | PPML |
| Fixed Effects | Yes | Yes | Yes | Yes |
| $R^2$ | 0.48 | - | 0.28 | - |
| N | 415 | 415 | 424 | 424 |

**Notes:**Estimation of bilateral commuting probabilities using equation 8. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

TABLE 8 High-skilled share by BRT line

| Order Built | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Line Name | J. B. Justo | 9 de Julio | Sur | Cabildo | San Martin |
| Median HS share | 0.46 | 0.54 | 0.33 | 0.65 | 0.48 |
| Mean HS share | 0.48 | 0.55 | 0.31 | 0.63 | 0.48 |

**Notes:** Each column shows the median and mean high-skilled share for census tracts that are at most one kilometer from a BRT line. The high-skilled share was calculated as the proportion of population in each census tract with at least some post-secondary education, according to the 2010 national census data.

TABLE 9 Regressions of home ownership rate for high- and low- skilled residents

| | Home Ownership Rate (HS) | Home Ownership Rate (LS) |
|---|---|---|
| HS Share | $-0.108^{***}$ | $-0.070^{***}$ |
| | (2.73) | (2.57) |
| constant | $0.673^{***}$ | $0.759^{***}$ |
| | (32.63) | (59.48) |
| $R^2$ | 0.03 | 0.02 |
| N | 201 | 201 |

**Notes:** Results for regressions of home ownership rate for high- and low-skilled residents by district on the initial high-skilled share. * p<0.10, ** p<0.05, *** p<0.01.

## REFERENCES

AGCBA (2015) Informe final de auditoría con informe ejecutivo. *Tech. rep.*, Buenos Aires, Argentina.

— (2019) Informe final de auditoría con informe ejecutivo. *Tech. rep.*, Buenos Aires, Argentina.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. and Wolf, N. (2015) The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, **83**, 2127–2189. URL: https://www.econometricsociety.org/doi/10.3982/ECTA10876.

Allen, T. and Arkolakis, C. (2019) The Welfare Effects of Transportation Infrastructure Improvements. URL: http://www.nber.org/papers/w25487. Working paper.

Allen, T., Arkolakis, C. and Li, X. (2016) Optimal City Structure. Working paper.

Almagro, M. and Domínguez-Iino, T. (2019) Location Sorting and Endogenous Amenities: Evidence from Short-Term Rentals in Amsterdam. Working paper.

Alves, G., Berniell, L., de la Mata, D., Fernandez, D., Juncosa, C., Rotondo, S. et al. (2018) Distribución espacial del empleo formal en la ciudad autónoma de buenos aires: un diagnóstico a partir de registros administrativos.

Artuç, E., Chaudhuri, S. and McLaren, J. (2010) Trade shocks and labor adjustment: A structural empirical approach. *American Economic Review*, **100**, 1008–1045. URL: http://pubs.aeaweb.org/doi/10.1257/aer.100.3.1008.

Balboni, C. (2019) In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities. Working paper.

Balboni, C., Bryan, G., Morten, M. and Siddiqi, B. (2020) Transportation, Gentrification, and Urban Mobility: The Inequality Effects of Place-Based Policies. Working paper.

Bardaka, E., Delgado, M. S. and Florax, R. J. (2018) Causal identification of transit-induced gentrification and spatial spillover effects: The case of the denver light rail. *Journal of Transport Geography*, **71**, 15–31.

Baum-Snow, N. (2007) Did highways cause suburbanization? *The quarterly journal of economics*, **122**, 775–805.

Baum-Snow, N., Freedman, M. and Pavan, R. (2018) Why has urban inequality increased? *American Economic Journal: Applied Economics*, **10**, 1–42.

Baum-Snow, N. and Turner, M. A. (2017) Transport infrastructure and the decentralization of cities in the people's republic of china. *Asian development review*, **34**, 25–50.

Berg, C. N., Deichmann, U., Liu, Y. and Selod, H. (2017) Transport policies and development. *Journal of Development Studies*, **53**, 465–480.

Bryan, G., Glaeser, E. and Tsivanidis, N. (2020) Cities in the developing world. *Annual Review of Economics*, **12**, 273–297.

Bryan, G. and Morten, M. (2019) The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy*, **127**, 2229–2268.

Caliendo, L., Dvorkin, M. and Parro, F. (2019) Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica*, **87**, 741–835.

Chapple, K. and Loukaitou-Sideris, A. (2019) *Transit-oriented displacement or community dividends?: Understanding the effects of smarter growth on communities*. MIT Press.

Couture, V., Gaubert, C., Handbury, J. and Hurst, E. (2018) Income Growth and the Distributional Effects of Urban Spatial Sorting. Working paper.

David, H., Palmer, C. J. and Pathak, P. A. (2017) Gentrification and the amenity value of crime reductions: Evidence from rent deregulation. *Tech. rep.*, National Bureau of Economic Research.

Dawkins, C. and Moeckel, R. (2016) Transit-induced gentrification: Who will stay, and who will go? *Housing Policy Debate*, **26**, 801–818.

Desmet, K., Kopp, R. E., Kulp, S. A., Nagy, D. K., Oppenheimer, M., Rossi-Hansberg, E. and Strauss, B. H. (forthcoming) Evaluating the economic cost of coastal flooding. *American Economic Journal: Macroeconomics*.

Dingel, J. I., Tintelnot, F. and Griffin, K. C. (2020) Spatial Economics for Granular Settings. Working paper.

Dragan, K., Ellen, I. and Glied, S. A. (2019) Does gentrification displace poor children? new evidence from new york city medicaid data. *Tech. rep.*, National Bureau of Economic Research.

Duranton, G. and Turner, M. A. (2012) Urban growth and transportation. *Review of Economic Studies*, **79**, 1407–1440.

Fajgelbaum, P. and Redding, S. (2018) Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914. Working paper.

Feenstra, R. C., Luck, P., Obstfeld, M. and Russ, K. N. (2018) In search of the armington elasticity. *Review of Economics and Statistics*, **100**, 135–150.

Fujimoto, T. and Krause, U. (1985) Strong ergodicity for strictly increasing nonlinear operators. *Linear Algebra and its Applications*, **71**, 101–112.

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

Glaeser, E. L., Kahn, M. E. and Rappaport, J. (2008) Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, **63**, 1–24.

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

Hannon, E., Nijssen, J. T., Stern, S. and Sumers, B. (2020) The effects of roads on trade and migration: Evidence from a planned capital city. *Tech. rep.*, McKinsey & Company.

Heblich, S., Redding, S. J. and Sturm, D. M. (2020) The making of the modern metropolis: evidence from london. *The Quarterly Journal of Economics*, **135**, 2059–2133.

Kamrowska-Zaluska, D. (2017) Buenos aires–toward comprehensive development and sustainable mobility. In *IOP conference series: materials science and engineering*, no. 6, 062022. IOP Publishing.

Morten, M. and Oliveira, J. (2018) The effects of roads on trade and migration: Evidence from a planned capital city. *NBER Working Paper*, **22158**, 43.

Padeiro, M., Louro, A. and da Costa, N. M. (2019) Transit-oriented development and gentrification: a systematic review. *Transport Reviews*, **39**, 733–754.

Pathak, R., Wyczalkowski, C. K. and Huang, X. (2017) Public transit access and the changing spatial distribution of poverty. *Regional Science and Urban Economics*, **66**, 198–212.

Redding, S. J. and Turner, M. A. (2015) Transportation costs and the spatial organization of economic activity. In *Handbook of regional and urban economics*, vol. 5, 1339–1398. Elsevier.

Roberts, M., Melecky, M., Bougna, T. and Xu, Y. (2020) Transport corridors and their wider economic benefits: A quantitative review of the literature. *Journal of Regional Science*, **60**, 207–248.

Sanderson, E. and Windmeijer, F. (2016) A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of Econometrics*, **190**, 212–221.

Secretaría de Planeamiento (2012) Distribución territorial del precio de oferta de los terrenos en la región metropolitana de buenos aires. *Tech. rep.*

— (2017) Distribución territorial del precio de oferta de los terrenos en la región metropolitana de buenos aires. *Tech. rep.*

Severen, C. (2018) Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification. Working paper.

Singh, D. Z. (2018) The tales of two mobility infrastructures: The street and the underground railway of buenos aires, 1880s–1940s. In *Architectures of Hurry—Mobilities, Cities and Modernity*, 65–82. Routledge.

Stock, J. and Yogo, M. (2005) Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, 109–120.

Su, Y. (2020) The rising value of time and the origin of urban gentrification. *Available at SSRN 3216013*.

Tehrani, S. O., Wu, S. J. and Roberts, J. D. (2019) The color of health: Residential segregation, light rail transit developments, and gentrification in the united states. *International Journal of Environmental Research and Public Health*, **16**, 3683.

Tsivanidis, N. (2019) Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio. Working paper.

UN (2018) 2018 revision of world urbanization prospects.

Windmeijer, F. A. and Santos Silva, J. M. (1997) Endogeneity in count data models: an application to demand for health care. *Journal of applied econometrics*, **12**, 281–294.

# A | DATA APPENDIX

## A.1 | Electoral Register Data

The electoral register (*padrón electoral*) for the city of Buenos Aires contains information on each citizen eligible for voting[54] that resides in the city of Buenos Aires, which amounts to approximately 2.5 million individuals.[55] This register contains the full name, sex of birth, year of birth, national ID number, residential address, and a description of occupation.[56] I use the full name, national ID number, and sex of birth to uniquely identify each citizen across registers.[57] This allows me to merge the 2011, 2013, 2015, and 2017 registers into a panel data set that contains addresses for each individual for every year that they appear in the data set.

### A.1.1 | Geocoding Observations

I assign geographical coordinates (latitude and longitude) to every street address that can be identified within the electoral register for each year using a combination of two geocoding services, the government of the city of Buenos Aires geocoding API[58], and the HERE geocoding API.[59] After extensive string cleaning and normalization, I was able to geocode 139,263 distinct addresses.[60] I discard government addresses and police stations, which are often used to provide an address for people who do not have a fixed residential address.[61] Finally, close to 5% of the individuals in the data set had residential addresses that could not be geocoded. Most of these individuals live in informal settlements, and therefore do not have standardized addresses that could be interpreted by the geocoding services. In the end, I obtain geocoded addresses that give me the street level location of the residential address for 2,382,326 individuals in 2011, 2,438,693 individuals in 2013, 2,446,541 individuals in 2015, and 2,461,564 individuals in 2017.

### A.1.2 | Classifying Occupations

The electoral register for the year 2011 includes a variables that describes the occupation of each person in the register. These descriptions were grouped into 4522 distinct occupations, which were then classified manually into one of the ten one-digit groups from the International Standard Classification of Occupations (ISCO). In order to guarantee the quality of the classification, I only classified occupations when it was clear the they belonged to a group, and when the original description was informative. Slightly more than half of the sample (1,305,217) had an occupation value of "student" (*estudiante* in Spanish). I decided to drop these individuals from my main analysis, since it wasn't clear what occupation or

[54] Every Argentine citizen over 18 years of age was eligible for voting in 2011. Starting in the 2013 elections, eligibility was extended to all citizens over 16 years of age. Since I restrict myself to a balanced panel of individuals that can be tracked between 2011 and 2017, I don't include individuals that are less than 18 years of age in 2013.

[55] The total number of voters registered in the city changes by less than 1% between 2011 and 2017.

[56] This description of occupation was only available in the 2011 register, so I don't have any time-series variation on occupation by individual.

[57] The national ID number and sex of birth are enough to uniquely identify 98% of the observations, for the remaining observations I use the full name to supplement the merging procedure.

[58] Available here: `https://usig.buenosaires.gob.ar/apis/`

[59] Available here: `https://developer.here.com/products/geocoding-and-search`

[60] These addresses do not contain the exact apartment number for multi-family residential buildings.

[61] Keeping these addresses does not change the main results of the paper, but increases the population substantially in certain census tracts in 2011 relative to the population observed in the 2010 census.

education level they might have. For another 399,185 individuals, no occupation was listed. In total, I was able to reliably identify an occupation for 607,775 individuals. I will track these individuals through time in order to construct the migration probabilities that will be used in the estimation of the model.

## A.2 | Construction of Employment Measure by District

The city's Annual Household Survey (*Encuesta Anual de Hogares*) allows me to calculate a measure of employment by skill type by commune. However, there are only 15 communes within the city, and, in order to capture changes in market access, it is convenient to have a smaller measure of employment at a smaller spatial unit.[62] In order to obtain a measure of employment at the district level,[63] I combine the measure of employment at the commune level with the land use map from 2010-2011 by skill type ($L_c^g$). I use the land use map to count the number of plots in each commune that are used for commercial or productive uses ($l_c^g$), I then count the number of plots in each district, within a commune, that are used for commercial or productive uses ($l_{v,c}^g$). With these three measures, I estimate the employment by skill type as

$$L_{v,C} = L_C \times \frac{l_{v,C}}{l_C}.$$

The underlying assumption is that workers of a given skill type within a commune are distributed between districts in proportion to the fraction of the total land used for commercial and productive uses in the commune that is in that district.

## A.3 | Commute Times

In order to calculate the commute times used in the model $\tau_{ij}$, I create a model of the city's transport network using ArcMap's Network Analyst Tool and GIS data on the street grid, the subway lines, the bus lines, the BRT lines, and the above-ground trains. I assign an average speed for subways, trains, buses, and walking within the network based on the average speeds calculated by mode of transport in the 2010 commuting survey (*ENMODO*). With this transport network put in place, I estimate commute times as the times for the minimum-time routes between the centroids of each residential location (defined either as a census tract or a district) to the centroid of each employment location (district), where the minimum-time route is calculated using the shortest path Dijkstra algorithm. I first calculate these commute times in a model of the city's transport network that does not include the BRT lines, and then, I add each lines sequentially according the the time they were inaugurated. Once a line or a set of lines have been added, I assume that buses that run through those lines increase their speed according the the increases in speed measures by the City Government, and re-calculate the minimum commute times.

[62] Alternatively, one could calculate average commute times for different points within a larger spatial unit, and use this average commute time as the commute time from any residential location to an employment location defined at the commune level. Both the reduced form and the quantitative results remain qualitatively similar when employing this alternative method.

[63] There are 167 districts, or an average of 11.13 districts per commune.

# B | SUPPLEMENTARY REDUCED FORM RESULTS

## B.1 | Reduced Form Results Using Low-Skilled CMA

TABLE B.1 Reduced form results using changes in market access for low-skilled residents: high-skilled share

| | $\Delta\log(\text{hs share})$ | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| | OLS | IV | OLS | IV | IV |
| $\Delta\log(CMA_l)$ | -0.024 | 0.424 | -2.634 | -1.088 | -2.450 |
| | (0.076) | (0.088)*** | (0.646)*** | (0.708) | (1.369)* |
| hs $\text{share}_0$ | 0.224 | 0.362 | 0.170 | 0.237 | 0.157 |
| | (0.037)*** | (0.047)*** | (0.085)** | (0.071)*** | (0.114) |
| $\Delta\log(CMA) \times$ hs sh. $\text{avg.}_0$ | | | 4.362 | 2.379 | 3.660 |
| | | | (0.992)*** | (1.117)** | (2.111)* |
| constant | -0.481 | -0.591 | -0.457 | -0.508 | -0.400 |
| | (0.025)*** | (0.033)*** | (0.058)*** | (0.048)*** | (0.078)*** |
| Neigh. FE | NO | NO | NO | NO | YES |
| N | 2,282 | 2,282 | 2,282 | 2,282 | 2,282 |

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

**Notes:** regressions of change in high skill share between 2011 and 2017 on change in market access for low-skilled workers between 2011 and 2017, initial average high skill share for contiguous census tracts, and the interaction term. The instrument used for the IV regressions is a measure of change of log $CMA_l$ where the 2017 $CMA_l$ is calculated assuming the BRT lines were built following the 1938 tramway lines. Cragg-Donald Wald F statistics from the first stage regression are reported for each IV estimation. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

TABLE B.2 Reduced form results using changes in market access for low-skilled residents: floorspace prices

| | $\Delta \log(\text{floorspace price})$ | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| | OLS | IV | OLS | IV | IV |
| $\Delta \log(CMA_l)$ | -0.178 | 0.470 | -4.609 | -1.871 | -5.907 |
| | (0.107)* | (0.078)*** | (1.150)*** | (1.246) | (1.659)*** |
| hs $\text{share}_0$ | 0.125 | 0.325 | -0.202 | 0.132 | -0.528 |
| | (0.057)** | (0.079)*** | (0.070)*** | (0.086) | (0.119)*** |
| $\Delta \log(CMA) \times$ hs sh. $\text{avg.}_0$ | | | 7.375 | 3.684 | 10.021 |
| | | | (1.781)*** | (1.932)* | (2.495)*** |
| constant | 0.071 | -0.088 | 0.277 | 0.041 | 0.392 |
| | (0.037)* | (0.053)* | (0.048)*** | (0.059) | (0.078)*** |
| Neigh. FE | NO | NO | NO | NO | YES |
| N | 2,282 | 2,282 | 2,282 | 2,282 | 2,282 |

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

**Notes:** regressions of change in floorspace prices between 2011 and 2017 on change in market access for low-skilled workers between 2011 and 2017, initial average high skill share for contiguous census tracts, and the interaction term. The instrument used for the IV regressions is a measure of change of log $CMA_l$ where the 2017 $CMA_l$ is calculated assuming the BRT lines were built following the 1938 tramway lines. Floorspace prices are calculated as the average sale price per square meter, according to geo-referenced online asking price data for the city of Buenos Aires. Cragg-Donald Wald F statistics from the first stage regression are reported for each IV estimation. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

# C | THEORY APPENDIX

## C.1 | Derivation of Value Function

In this section I show how to go from equation (5) to equation (6). Consider

$$V^g_{n,t} = E_0\left[\max_{\{i\}}\left\{\frac{w^g_{j,t}\varepsilon_{j,t}}{d_{ij,t}}\right\}\right] + E_0\left[\max_{\{j\}}\left\{\beta V^g_{i,t+1} - \mu^g_{ni} + \eta^g_{i,t}\right\}\right].$$

I will first show that

$$E_0\left[\max_{\{i\}}\left\{\frac{w^g_{j,t}\varepsilon_{j,t}}{d_{ij,t}}\right\}\right] = \tilde{T}\Phi^{\frac{1}{\theta}}_{Rn,t}.$$

Let $y_j = \frac{w_j \varepsilon_j}{d_{nj}}$, where I drop the time subscript for clarity. We will assume $\varepsilon_j$ is distributed extreme value type II, so that $F(\varepsilon_j) = \exp((-T\varepsilon)^{-\theta})$. Then, the CDF of $y_j$ is

$$\begin{aligned}
G_j(y) &= P(y_j \leqslant y) \\
&= P(\frac{w_j \varepsilon_j}{d_{nj}} \leqslant y) \\
&= P\left(\varepsilon_j \leqslant y \frac{d_{nj}}{w_j}\right) \\
&= e^{-\left(Ty\frac{d_{nj}}{w_j}\right)^{-\theta}}.
\end{aligned}$$

Let $\tilde{y} = \max_j\{y_j\}$, then we can define the CDF of $\tilde{y}$ as

$$\begin{aligned}
G(y) &= P(\tilde{y} \leqslant y) \\
&= \prod_{j=1}^{J} P(y_j \leqslant y) \\
&= \prod_{j=1}^{J} \exp\left[\left(\frac{-y}{T^{-1}\frac{w_j}{d_{nj}}}\right)^{-\theta}\right] \\
&= \exp\left[-\sum_{j=1}^{J}\left(\frac{y}{T^{-1}\frac{w_j}{d_{nj}}}\right)^{-\theta}\right] \\
&= \exp\left[-y^{-\theta}T^{-\theta}\sum_{j=1}^{J}\left(\frac{w_j}{d_{nj}}\right)^{\theta}\right] \\
&= \exp\left(-\left(\frac{y}{T^{-1}\left[\sum_{j=1}^{J}\left(\frac{w_j}{d_{nj}}\right)^{\theta}\right]^{\frac{1}{\theta}}}\right)^{-\theta}\right).
\end{aligned}$$

Which implies that $\tilde{y}$ is distributed Frechet with a shape parameter $\theta$ and a scale parameter $T^{-1}\left[\sum_{j=1}^{J}\left(\frac{w_j}{d_{nj}}\right)^{\theta}\right]^{\frac{1}{\theta}}$. This implies that the expected value of $\tilde{y}$ is

$$\begin{aligned}
E\left(\max_{j=1}^{J}\left\{\frac{w_j \varepsilon_j}{d_{nj}}\right\}\right) = E(\tilde{y}) &= T^{-1}\Gamma\left(1 - \frac{1}{\theta}\right)\left[\sum_{j=1}^{J}\left(\frac{w_j}{d_{nj}}\right)^{\theta}\right]^{\frac{1}{\theta}} \\
&= \tilde{T}\Phi_{Rn}^{\frac{1}{\theta}}.
\end{aligned}$$

Where $\tilde{T} = T^{-1}\Gamma\left(1 - \frac{1}{\theta}\right)$, and $\Phi_{Rn} = \sum_{j=1}^{J}\left(\frac{w_j}{d_{nj}}\right)^{\theta}$.

For the second term in (5), recall that $\eta_{i,t}^{g}$ is distributed extreme value type I with parameters $(-\gamma\nu_g, \nu_g)$, where $\gamma$ is the Euler-Mascheroni constant. Let us define

$$\mathcal{O}_{n,t}^{g} = E_0\left[\max_{\{j\}}\left\{\beta V_{i,t+1}^{g} - \mu_{ni}^{g} + \eta_{i,t}^{g}\right\}\right].$$

We can rewrite this expression as

$$\begin{aligned}
\mathcal{O}^g_{n,t} &= E\{\sum_{i=1}^{I} \beta V^g_{i,t+1} - \mu^g_{ni} + \eta^g_{i,t}\} \times \Pr[(\beta V^g_{i,t+1} - \mu^g_{ni} + \eta^g_{i,t}) \geqslant \beta V^g_{m,t+1} - \mu^g_{nm} + \eta^g_{m,t}\ \forall\ m = 1, ..., N] \\
&= \sum_{i=1}^{I} \int (\beta V^g_{i,t+1} - \mu^g_{ni} + \eta^g_{i,t}) f(\eta^g_{i,t}) \times \prod_{m \neq i} F\left(\beta(V^g_{i,t+1} - V^g_{m,t+1}) - (\mu^g_{ni} - \mu^g_{nm}) + \eta^g_{i,t}\right) d\eta^g_{i,t} \\
&= \sum_{i=1}^{I} \int (\beta V^g_{i,t+1} - \mu^g_{ni} + \eta^g_{i,t}) \frac{1}{\nu_g} \exp\left(\frac{\eta^g_{i,t}}{\nu_g} - \gamma\right) \times \\
&\times \exp\left(-\sum_{i=1}^{I} \exp\left(-\frac{\beta(V^g_{i,t+1} - V^g_{m,t+1}) - (\mu^g_{ni} - \mu^g_{nm}) + \eta^g_{i,t}}{\nu_g} - \gamma\right)\right) d\eta^g_{i,t}.
\end{aligned}$$

Where the last step used the fact that the pdf of $\eta^g_{i,t}$ is

$$f(x) = \frac{1}{\nu_g} \exp\left(-\frac{x}{\nu_g} - \gamma - \exp\left(-\frac{x}{\nu_g} - \gamma\right)\right)$$

and the CDF of $\eta^g_{i,t}$ is

$$F(x) = \exp\left(-\exp\left(-\frac{x}{\nu_g} - \gamma\right)\right).$$

Let us define $b^g_{i,t} = \frac{\eta^g_{i,t}}{\nu_g} + \gamma$, $a^g_{im,t} = \frac{\beta(V^g_{i,t+1} - V^g_{m,t+1}) - (\mu^g_{ni} - \mu^g_{nm}) + \eta^g_{i,t}}{\nu_g}$, $\tilde{a}^g_{i,t} = \ln \sum_{m=1}^{I} \exp(a^g_{im,t})$, and $\tilde{b}^g_{i,t} = b^g_{i,t} - \tilde{a}^g_{i,t}$. First consider the change of variables using $b^g_{i,t}$:

$$\begin{aligned}
\mathcal{O}^g_{n,t} &= \sum_{i=1}^{I} \int (\beta V^g_{i,t+1} - \mu^g_{ni} + \nu_g(b^g_{it} - \gamma)) \exp\left(-b^g_{i,t} - \sum_{i=1}^{I} \exp\left(a^g_{im,t} - b^g_{i,t}\right)\right) db^g_{i,t} \\
&= \sum_{i=1}^{I} \int (\beta V^g_{i,t+1} - \mu^g_{ni} + \nu_g(b^g_{it} - \gamma)) \exp\left(-b^g_{i,t} - \exp\left(-(b^g_{i,t} - \tilde{a}^g_{i,t})\right)\right) db^g_{i,t}.
\end{aligned}$$

Doing a second change of variables using $\tilde{b}^g_{i,t}$ we obtain:

$$\mathcal{O}^g_{n,t} = \sum_{i=1}^{I} \exp(\tilde{a}^g_{i,t})(\beta V^g_{i,t+1} - \mu^g_{ni} + \nu_g(\tilde{a}^g_{it} - \gamma)) + \nu_g \int \tilde{b}^g_{i,t} \exp\left(-\tilde{b}^g_{i,t} - \exp\left(-\tilde{b}^g_{i,t}\right)\right) d\tilde{b}^g_{i,t}.$$

Using the fact that $\gamma = \int x \exp(-x - \exp(-x))$, we obtain

$$\mathcal{O}^g_{n,t} = \sum_{i=1}^{I} \exp(\tilde{a}^g_{i,t})(\beta V^g_{i,t+1} - \mu^g_{ni} + \nu_g \tilde{a}^g_{it}).$$

Substituting in for $\tilde{a}^g_{i,t}$, we find that

$$\mathcal{O}^g_{n,t} = \nu_g \ln \sum_{i=1}^{I} \left(\exp(\beta V^g_{i,t+1} - \mu^g_{ni})\right)^{\frac{1}{\nu_g}}.$$

Combining the first and second term, we arrive at equation (6).

### C.2 | Deriving Neighborhood Migration Flows

Let $m^g_{in,t}$ be the fraction of residents of type $g$ that are residing in location $n$ at the beginning of period $t$, and move to location $i$ by the end of period $t$. Since there is a continuum of agents in each residential location at any time $t$, we can apply the strong law of large numbers and express this fraction as the probability that any worker living in location $n$ chooses to move to location $i$ at the end of time $t$, or

$$\begin{aligned} m^g_{in,t} &= P\left[(\beta V^g_{i,t+1} - \mu_{in} + \eta^g_{i,t}) \geqslant (\beta V^g_{m,t+1} - \mu^g_{mn} + \eta^g_{m,t}) \forall m = 1, ..., N\right] \\ &= \int_{-\infty}^{\infty} f(\eta^g_{i,t}) \prod_{m \neq i} F(\beta(V^g_{i,t+1} - V^g_{m,t+1}) - (\mu^g_{in} - \mu^g_{mn}) + \eta^g_{i,t})\, d\eta^g_{i,t}. \end{aligned}$$

Let $z^g_{im,t} = \beta(V^g_{i,t+1} - V^g_{m,t+1}) - (\mu^g_{in} - \mu^g_{mn})$, then

$$\begin{aligned} m^g_{in,t} &= \int_{-\infty}^{\infty} f(\eta^g_{i,t}) \prod_{m \neq i} F(z^g_{im,t} + \eta^g_{i,t})\, d\eta^g_{i,t} \\ &= \int \frac{1}{\nu} \left( \frac{-\eta^g_{i,t}}{\nu_g} - \gamma - \exp\left( \frac{-\eta^g_{i,t}}{\nu_g} - \gamma \right) \right) \prod_{m \neq i} \exp\left( -\exp\left( \frac{-z^g_{im,t}}{\nu_g} - \frac{\eta^g_{i,t}}{\nu_g} - \gamma \right) \right)\, d\eta^g_{i,t} \\ &= \int \frac{1}{\nu_g} \exp\left( \frac{-\eta^g_{i,t}}{\nu_g} - \gamma \right) \prod_{m \neq i} \exp\left( -\exp\left( \frac{-z^g_{im,t}}{\nu_g} - \frac{\eta^g_{i,t}}{\nu_g} - \gamma \right) \right)\, d\eta^g_{i,t} \\ &= \int \frac{1}{\nu_g} \exp\left( \frac{-\eta^g_{i,t}}{\nu_g} - \gamma \right) \exp\left( -\sum_{m=1}^{N} \exp\left( \frac{-z^g_{im,t}}{\nu_g} - \frac{\eta^g_{i,t}}{\nu_g} - \gamma \right) \right)\, d\eta^g_{i,t}. \end{aligned}$$

Define $\lambda^g_t = \log\left(\sum_{m=1}^N \exp\left(\frac{-z^g_{im,t}}{\nu_g}\right)\right)$, $x^g_t = \frac{\eta^g_{i,t}}{\nu_g} + \gamma$, $y^g_t = x^g_t - \lambda^g_t$. Note that

$$\frac{dx^g_t}{d\eta^g_{i,t}} = \frac{1}{\nu_g} \Rightarrow d\eta^g_{i,t} = \nu_g dx^g_t,$$

$$\text{and } \frac{dy^g_t}{d\eta^g_{i,t}} = \frac{dx^g_t}{d\eta^g_{i,t}} \Rightarrow dy^g_t = dx^g_t.$$

Therefore, we can rewrite $m^g_{in,t}$ as

$$
\begin{aligned}
m^g_{in,t} &= \int_{-\infty}^{\infty} \frac{1}{\nu_g} \exp(-x^g_t) \exp(-\exp(\lambda^g_t) \exp(-x^g_t))\nu_g \, dx^g_t \\
&= \int \exp(-y^g_t - \lambda^g_t) \exp(-\exp(\lambda^g_t) \exp(-y^g_t - \lambda^g_t)) dy^g_t \\
&= \exp(\lambda^g_t) \int \exp(-y^g_t - \exp(y^g_t)) dy^g_t \\
&= \exp(-\lambda^g_t) \left[\exp(-\exp(y^g_t))\right] |^{\infty}_{-\infty} \\
&= \exp(-\lambda^g_t) \\
&= \frac{1}{\sum_{m=1}^{n} \exp\left(\frac{-z^g_{in,t}}{\nu_g}\right)} \\
&= \frac{1}{\sum_{m=1}^{n} \exp\left[\frac{1}{\nu_g}(-\beta(V^g_{i,t+1} - V^g_{m,t+1}) + (\mu^g_{in} - \mu^g_{mn})\right]} \\
&= \frac{\left[\exp(\beta V^g_{i,t+1} - \mu^g_{in})\right]^{\frac{1}{\nu_g}}}{\sum_{m=1}^{n} \left[\exp\left((\beta V^g_{m,t+1}) - \mu^g_{mn})\right)\right]^{\frac{1}{\nu_g}}}.
\end{aligned}
$$

### C.3 | Deriving Labor Demand

From the first order conditions of the producers problem, we know that

$$w^h_{j,t} \tilde{L}^h_{Fj,t} = (1 - \rho_j) X_{jt},$$

$$w^l_{j,t} \tilde{L}^l_{Fj,t} = \rho_j X_{jt}.$$

Where $X_{jt}$ is the total expenditure on variety j at time t. From the CES demand for varieties of the consumption good, we know that in equilibrium $X_{jt} = p_{jt}^{1-\sigma} X$, where $X = \sum_{i=1}^{I} \alpha \sum_{g \in \{h,l\}} \bar{y}_{igt}$ is the total expenditure on consumption in the economy, and $\bar{y}_{igt}$ is the mean income of a worker of type g in location i at time t. Perfect competition will imply that the price of each variety is equal to its marginal cost at each time t, $p_{jt} = A_j^{-1} (w^l_{jt})^{\rho_j} (w^h_{jt})^{(1-\rho_j)}$. Combining these results and re-arranging terms we get

$$\tilde{L}^l_{Fj,t} = \frac{\rho_j}{w^l_j} \left(\frac{(w^l_j)^{\rho_j} (w^h_j)^(1 - \rho_j)}{A_j}\right)^{(1-\sigma)} \sum_i \sum_g \alpha \tilde{T}_g \left(\sum_s \left(\frac{w^g_s}{d_{is}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}} L^g_{Ri,t}, \quad \text{(C.1)}$$

$$\tilde{L}^h_{Fj,t} = \frac{1 - \rho_j}{w^h_j} \left(\frac{(w^l_j)^{\rho_j} (w^h_j)^(1 - \rho_j)}{A_j}\right)^{(1-\sigma)} \sum_i \sum_g \alpha \tilde{T}_g \left(\sum_s \left(\frac{w^g_s}{d_{is}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}} L^g_{Ri,t}. \quad \text{(C.2)}$$

Let $\mathbf{w}_t = \{\{w^h_{jt}\}^J_{j=1}, \{w^l_{jt}\}^J_{j=1}\}$, note that the right-hand side of equations C.1 and C.2 depends on $\mathbf{w}_t$, the residential population vector for high- and low-skilled workers at time t (which is assumed to be known to the producer at time t), and known parameters. Therefore, we can define

$$f_{jlt}(\mathbf{w}_t) = \frac{\rho_j}{w_j^l}\left(\frac{(w_j^l)^{\rho_j}(w_j^h)^(1-\rho_j)}{A_j}\right)^{(1-\sigma)}\sum_i\sum_g \alpha\tilde{T}_g\left(\sum_s\left(\frac{w_s^g}{d_{is}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}} L_{Ri,t}^g, \quad \text{(C.3)}$$

$$f_{jht}(\mathbf{w}_t) = \frac{1-\rho_j}{w_j^h}\left(\frac{(w_j^l)^{\rho_j}(w_j^h)^(1-\rho_j)}{A_j}\right)^{(1-\sigma)}\sum_i\sum_g \alpha\tilde{T}_g\left(\sum_s\left(\frac{w_s^g}{d_{is}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}} L_{Ri,t}^g. \quad \text{(C.4)}$$

### C.4 | Deriving Average Match Productivity

Consider

$$\begin{aligned}\bar{\varepsilon}_{jt}^g &= E[\varepsilon|g,t,\ \text{choose j}]\\ &= \sum_{i=1}^{I} E[\varepsilon|g,t,\ \text{choose j if living in i}]Pr(i|j,g,t)\\ &= \sum_{i=1}^{I} T_g\Pi_{j|git}^{-\frac{1}{\theta_g}}\frac{1}{d_{ij,t}}Pr(i|j,g,t).\end{aligned}$$

Where the last equality comes from the properties of $\epsilon$ being a random variable distributed Fréchet, conditional on $\frac{w_{j,t}^g\varepsilon_{j,t}}{d_{ij,t}}$ being the maximal element of $\left\{\frac{w_{s,t}^g\varepsilon_{s,t}}{d_{sj,t}}\right\}_{s=1}^{I}$. Solving for $Pr(i|j,g,t)$, we get

$$\begin{aligned}Pr(i|j,g,t) &= \Pi_{i|gjt}\\ &= \frac{\Pi_{j|git}\Pi_{git}}{\sum_{r=1}^{I}\Pi_{j|grt}\Pi_{grt}}\\ &= \frac{\Pi_{j|git}L_{Ri,t}^g}{\sum_{r=1}^{I}\Pi_{j|grt}L_{Rr,t}^g}.\end{aligned}$$

Which implies that

$$\bar{\varepsilon}_{jt}^g = T_g\sum_i^I \pi_{j|tig}^{-\frac{1}{\theta_g}}\frac{1}{d_{ij,t}}\frac{\pi_{j|tig}L_{Ri,t}^g}{\sum_n^I \pi_{j|tng}L_{Rn,t}^g}. \quad \text{(C.5)}$$

### C.5 | Derivation of Equations in First Difference

#### C.5.1 | Deriving Equation 17

Let us first define $\dot{u}_{n,t+1}^g = \exp(V_{n,t+1}^g - V_{n,t}^g)$. Then, from equation 6 we get

$$\left(\dot{u}^g_{n,t+1}\right)^{\frac{1}{\nu_g}} = \left[\frac{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)}{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t} r^{\alpha-1}_{Rgn,t}\right)}\right]^{\frac{1}{\nu_g}} \left[\frac{\sum_{i=1}^{I}\left(\exp(\beta V^g_{i,t+2} - \mu^g_{in})\right)^{\frac{1}{\nu_g}}}{\sum_{i=1}^{I}\left(\exp(\beta V^g_{i,t+1} - \mu^g_{in})\right)^{\frac{1}{\nu_g}}}\right].$$

Multiplying and dividing each term in the sum $\sum_{i=1}^{I}\left(\exp(\beta V^g_{i,t+2} - \mu^g_{in})\right)^{\frac{1}{\nu_g}}$ by $\left(\exp(\beta V^g_{i,t+1} - \mu^g_{in})\right)^{\frac{1}{\nu_g}}$ we obtain

$$\left(\dot{u}^g_{n,t+1}\right)^{\frac{1}{\nu_g}} = \left[\frac{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)}{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t} r^{\alpha-1}_{Rgn,t}\right)}\right]^{\frac{1}{\nu_g}} \left[\frac{\sum_{k=1}^{I}\left(\exp(\beta V^g_{k,t+2} - \mu^g_{kn})\right)^{\frac{1}{\nu_g}} \frac{\left(\exp(\beta V^g_{k,t+1} - \mu^g_{kn})\right)^{\frac{1}{\nu_g}}}{\left(\exp(\beta V^g_{k,t+1} - \mu^g_{kn})\right)^{\frac{1}{\nu_g}}}}{\sum_{i=1}^{I}\left(\exp(\beta V^g_{i,t+1} - \mu^g_{in})\right)^{\frac{1}{\nu_g}}}\right].$$

Reordering terms, and using the fact that

$$m^g_{kn,t} = \frac{\left[\exp(\beta\beta V^g_{k,t+1} - \mu^g_{kn})\right]^{\frac{1}{\nu_g}}}{\sum_{m=1}^{n}\left[\exp\left((\beta\beta V^g_{m,t+1}) - \mu^g_{mn})\right)\right]^{\frac{1}{\nu_g}}},$$

we obtain

$$\begin{aligned}
\left(\dot{u}^g_{n,t+1}\right)^{\frac{1}{\nu_g}} &= \left[\frac{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)}{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t} r^{\alpha-1}_{Rgn,t}\right)}\right]^{\frac{1}{\nu_g}} \left[\sum_{k=1}^{I} m^g_{kn,t} \frac{\left(\exp(\beta V^g_{k,t+2} - \mu^g_{kn})\right)^{\frac{1}{\nu_g}}}{\left(\exp(\beta V^g_{k,t+1} - \mu^g_{kn})\right)^{\frac{1}{\nu_g}}}\right] \\
&= \left[\frac{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)}{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t} r^{\alpha-1}_{Rgn,t}\right)}\right]^{\frac{1}{\nu_g}} \left[\sum_{k=1}^{I} m^g_{kn,t} \left(\exp(\beta(V^g_{k,t+2} - V^g_{k,t+1}))\right)^{\frac{1}{\nu_g}}\right] \\
&= \left[\frac{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)}{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t} r^{\alpha-1}_{Rgn,t}\right)}\right]^{\frac{1}{\nu_g}} \left[\sum_{k=1}^{I} m^g_{kn,t} \left(\dot{u}^g_{k,t+2}\right)^{\frac{\beta}{\nu_g}}\right].
\end{aligned}$$

Which implies that

$$\dot{u}^g_{n,t+1} = \left[\frac{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)}{\exp\left(\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgn,t} r^{\alpha-1}_{Rgn,t}\right)}\right] \left[\sum_{k=1}^{I} m^g_{kn,t} \left(\dot{u}^g_{k,t+2}\right)^{\frac{\beta}{\nu_g}}\right]^{\nu_g}. \tag{C.6}$$

### C.5.2 | Deriving Equation 18

Let us define $\dot{m}^g_{in,t+1} = \frac{m^g_{in,t+1}}{m^g_{in,t}}$. Then, from equation 7

$$\dot{m}^g_{in,t+1} = \frac{\frac{\left[\exp(\beta V^g_{i,t+2} - \mu^g_{in})\right]^{\frac{1}{\nu_g}}}{\sum_{m=1}^{N}\left[\exp(\beta V^g_{m,t+2} - \mu^g_{mn})\right]^{\frac{1}{\nu_g}}}}{\frac{\left[\exp(\beta V^g_{i,t+1} - \mu^g_{in})\right]^{\frac{1}{\nu_g}}}{\sum_{m=1}^{N}\left[\exp(\beta V^g_{m,t+1} - \mu^g_{mn})\right]^{\frac{1}{\nu_g}}}}$$

$$= \frac{\left(\exp(\beta(V^g_{i,t+2} - V^g_{i,t+1}))\right)^{\frac{1}{\nu_g}}}{\sum_{k=1}^{I} m^g_{kn,t}\left(\exp(\beta(V^g_{k,t+2} - V^g_{k,t+1}))\right)^{\frac{1}{\nu_g}}}$$

$$= \frac{\left(\dot{u}^g_{i,t+2}\right)^{\frac{\beta}{\nu_g}}}{\sum_{k=1}^{I} m^g_{kn,t}\left(\dot{u}^g_{k,t+2}\right)^{\frac{\beta}{\nu_g}}}.$$

### C.5.3 | Deriving Equation 19

Let us define $\dot{r}_{Ri,t+1} = \frac{r_{Ri,t+1}}{r_{Ri,t}}$. From equation 16 we obtain

$$\dot{r}_{Ri,t+1} = \frac{\frac{\sum_{g\in\{h,l\}} L^g_{Ri,t+1}\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgi,t+1}(1-\alpha)}{\bar{H}_{Ri}}}{\frac{\sum_{g\in\{h,l\}} L^g_{Ri,t}\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgi,t}(1-\alpha)}{\bar{H}_{Ri}}}$$

$$= \frac{\sum_{g\in\{h,l\}} L^g_{Ri,t+1}\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgi,t+1}}{\sum_{g\in\{h,l\}} L^g_{Ri,t}\tilde{T}_g \Phi^{\frac{1}{\theta_g}}_{Rgi,t}}.$$

## C.6 | Numerical Solution Algorithm

In this section I provide a more detailed explanation of the numerical solution algorithm implemented in the main quantitative model of this paper.

1. Initiate the algorithm at $t = 0$ with a guess for a path of $\{\{\{\dot{u}^{g0}_{t+1}\}^{I}_{n=1}\}^{\infty}_{t=0}\}_{g\in\{h,l\}}$, such that $\dot{u}^g_{n,T+1} = 1$ for all T large enough, and for all $n$. Take as given $\{L^g_{R0}, L^g_{F0}, m^g_{-1}, \{d_{ij,t}\}^{I,J,\infty}_{i=1,j=1,t=1}, \{A_j\}^{J}_{j=1}, \{\rho_j\}^{J}_{j=1}, \{\bar{H}_{Ri}\}^{I}_{i=1}, \{r_{Ri,0}\}^{I}_{i=1}\}$.
2. For all $t \geqslant 0$, use $\{\{\dot{u}^{g0}_{t+1}\}^{I}_{n=1}\}^{\infty}_{t=0}$ to solve for $\{m^g_t\}^{\infty}_{t=1}$ using equation (18).
3. Use $L^g_{R0}$, $\{m_t\}^{\infty}_{t=1}$, and equation (10) to solve for $\{L^g_{Rt}\}^{\infty}_{t=0}$.
4. Use $\{L^g_{Rt}\}$ and $L^g_{F0}$ to estimate model consistent wages for period zero ($t = 0$), $\{w^g_{j,0}\}^{J}_{j=1}$ using the labor market clearing condition for each skill type, (15).
   This step follows closely the way equilibrium wages are obtained in Ahlfeldt et al. (2015) using employment population and residential population vectors for $t = 0$. The key here is that these vectors are observed in the data, and so we can find the wages at $t = 0$ that are consistent with the distribution of employment and residential population observed in the data, conditional on the estimated parameters.

5. Use $\{L^g_{Rt}\}$, $\{w^g_{j,0}\}^J_{j=1}$, $L^g_{F0}$, and vector equation implied by (15) to solve forward for the model consistent $\{L^g_{Ft}\}^\infty_{t=1}$ and $\{\{w^g_{j,t}\}^J_{j=1}\}^\infty_{t=0}$.
   In this section I implement a numerical algorithm to find model-consistent wages for high- and low-skilled workers by inverting the labor demand functions from equations C.3 and C.4. See Subsection C.6.1 for details on this algorithm.
6. Use $\{L^g_{Rt}\}$, $\{\Phi_{Rgt}\}$, and equation (19) to solve for $\{\dot{r}_{Rn,t+1}\}$. Then use $\{r_{Ri,0}\}^I_{i=1}$ and $\{\dot{r}_{Rn,t+1}\}$ to solve for $\{r_{Rn,t}\}$.
7. For each t, use $\{\{w^g_{j,t+1}\},\{m^g_t\},\{d_{ij,t}\},\{\dot{r}_{Rn,t+1}\}$, and $\{\dot{u}^g_{i,t+2}\}\}$ to calculate backwards $\{\dot{u}^g_{n,t+1}\}$ using equation (17) for each skill type g. This will result in a new sequence $\{\dot{u}^{g1}_{n,t+1}\}^\infty_{t=0}$.
8. Verify if $\{\dot{u}^{g1}_{n,t+1}\}^\infty_{t=0} \approx \{\dot{u}^{g0}_{n,t+1}\}^\infty_{t=0}$, if not, then start the again from step 1, with $\{\dot{u}^{g1}_{n,t+1}\}^\infty_{t=0}$ as your new guess. If indeed $\{\dot{u}^{g1}_{n,t+1}\}^\infty_{t=0} \approx \{\dot{u}^{g0}_{n,t+1}\}^\infty_{t=0}$, then $\{\dot{u}^{g1}_{n,t+1}\}^\infty_{t=0}$ is the solution.

### C.6.1 | Algorithm for Model-Consistent Wages

In order to calculate model-consistent wages at time $t > 0$ using the residential population vector at time t, $\{L^g_{Ri,t}\}$ for residents of both skill types, I use equations C.3 and C.4, as well as the labor market clearing condition (equation 15) to express wages for both skill types as a function of the full vector of wages, the full vector of residential population, and estimated parameters.

Take 15 for each skill type g, and replace $\tilde{L}^g_{Fj,t}$ with $f_{jgt}(\mathbf{w}_t)$ from equations C.3 and C.4. Rearranging terms we get

$$\left(w^l_{j,t}\right)^{\theta_l+1+(\sigma-1)\rho_j} = \frac{\rho_j}{\bar{\varepsilon}^l_{j,t}}\left(\frac{(w^h_{j,t})^(1-\rho_j)}{A_j}\right)^{(1-\sigma)} \sum_i\sum_g \alpha\tilde{T}_g\left(\sum_s\left(\frac{w^g_{s,t}}{d_{is,t}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}}\left[\sum_i \frac{\frac{L^l_{Ri,t}}{(d_{ij,t})^{\theta_l}}}{\Phi_{Ril,t}}\right]^{-1},$$

and

$$\left(w^h_{j,t}\right)^{\theta_h+1+(\sigma-1)(1-\rho_j)} = \frac{(1-\rho_j)}{\bar{\varepsilon}^h_{j,t}}\left(\frac{(w^l_{j,t})^{\rho}_j}{A_j}\right)^{(1-\sigma)} \sum_i\sum_g \alpha\tilde{T}_g\left(\sum_s\left(\frac{w^g_{s,t}}{d_{is,t}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}}\left[\sum_i \frac{\frac{L^h_{Ri,t}}{(d_{ij,t})^{\theta_h}}}{\Phi_{Rih,t}}\right]^{-1}.$$

Which imply that

$$w^l_{j,t} = \left[\frac{\rho_j}{\bar{\varepsilon}^l_{j,t}}\left(\frac{(w^h_{j,t})^(1-\rho_j)}{A_j}\right)^{(1-\sigma)} \sum_i\sum_g \alpha\tilde{T}_g\left(\sum_s\left(\frac{w^g_{s,t}}{d_{is,t}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}}\left[\sum_i \frac{\frac{L^l_{Ri,t}}{(d_{ij,t})^{\theta_l}}}{\Phi_{Ril,t}}\right]^{-1}\right]^{\frac{1}{\theta_l+1+(\sigma-1)\rho_j}}, \tag{C.7}$$

and

$$w^h_{j,t} = \left[\frac{(1-\rho_j)}{\bar{\varepsilon}^h_{j,t}}\left(\frac{(w^l_{j,t})^{\rho}_j}{A_j}\right)^{(1-\sigma)} \sum_i\sum_g \alpha\tilde{T}_g\left(\sum_s\left(\frac{w^g_{s,t}}{d_{is,t}}\right)^{\theta_g}\right)^{\frac{1}{\theta_g}}\left[\sum_i \frac{\frac{L^h_{Ri,t}}{(d_{ij,t})^{\theta_h}}}{\Phi_{Rih,t}}\right]^{-1}\right]^{\frac{1}{\theta_h+1+(\sigma-1)(1-\rho_j)}}. \tag{C.8}$$

Stacking all the wages for high- and low-skilled into one vector of length $2 \times J$, $\mathbf{w}_t = \{\{w^h_{jt}\}^J_{j=1}, \{w^l_{jt}\}^J_{j=1}\}$, we get

$$\mathbf{w}_t = \mathbf{g}(\mathbf{w}_t). \tag{C.9}$$

Where $\mathbf{g} : \mathbb{R}^{2J} \to \mathbb{R}^{2J}$ is a vector function such that

$$\mathbf{g}_{g,t} = \begin{cases} \left[ \left[ \frac{(1-\rho_j)}{\bar{\varepsilon}^h_{j,t}} \left( \frac{(w^l_{j,t})^\rho_j}{A_j} \right)^{(1-\sigma)} \sum_i \sum_g \alpha \tilde{T}_g \left( \sum_s \left( \frac{w^g_{s,t}}{d_{is,t}} \right)^{\theta_g} \right)^{\frac{1}{\theta_g}} \left[ \sum_i \frac{\frac{L^h_{Ri,t}}{(d_{ij,t})^{\theta_h}}}{\Phi_{Rih,t}} \right]^{-1} \right] \right]^{\frac{1}{\theta_h + 1 + (\sigma-1)(1-\rho_j)}} & \text{if } g = h \\ \left[ \left[ \frac{\rho_j}{\bar{\varepsilon}^l_{j,t}} \left( \frac{(w^h_{j,t})^{(1-\rho_j)}}{A_j} \right)^{(1-\sigma)} \sum_i \sum_g \alpha \tilde{T}_g \left( \sum_s \left( \frac{w^g_{s,t}}{d_{is,t}} \right)^{\theta_g} \right)^{\frac{1}{\theta_g}} \left[ \sum_i \frac{\frac{L^l_{Ri,t}}{(d_{ij,t})^{\theta_l}}}{\Phi_{Ril,t}} \right]^{-1} \right] \right]^{\frac{1}{\theta_l + 1 + (\sigma-1)\rho_j}} & \text{if } g = l. \end{cases}$$

In this way we have defined the full vector of wages as a fixed point that solves the vector equation $\mathbf{w}_t = \mathbf{g}(\mathbf{w}_t)$. In order to obtain a unique (to-scale) $\mathbf{w}_t$, we can start with any initial vector, and apply iteratively $\mathbf{g}(\mathbf{w}_t)$ until the difference between one iteration and the previous is small enough, where I define the difference in terms of the $L^2$ distance between vectors. This algorithm will result in a unique (to-scale) vector that solved the vector equation $\mathbf{w}_t = \mathbf{g}(\mathbf{w}_t)$. This follows from the fact that $\mathbf{g}(\mathbf{x})$ is strictly increasing in $\mathbf{x}$, and weakly homogeneous, which implies that $\mathbf{g}(\lambda \mathbf{x}) = f(\lambda)\mathbf{g}(\mathbf{x})$, for $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\frac{f(\lambda)}{\lambda}$ is strictly increasing and $f(0) = 0$. With these properties, Fujimoto and Krause (1985) show that there must exist a unique up-to-scale solution to $\mathbf{w}_t = \mathbf{g}(\mathbf{w}_t)$.

## C.7 | Deriving Expected Utility Expression

From equation (7), and the assumption that the mobility costs are zero for non-movers ($\mu_{nn} = 0$) we know that

$$m^g_{nn,t} = \frac{\left(\exp(\beta V^g_{n,t+1})\right)^{\frac{1}{\nu}}}{\sum_{m=1}^I \left(\exp(\beta V^g_{m,t+1} - \mu^g_{mn})\right)^{\frac{1}{\nu_g}}}.$$

Taking logs at both sides we get

$$\ln(m^g_{nn,t}) = \frac{1}{\nu_g} \left(\beta V^g_{n,t+1}\right) - \ln \left( \sum_{m=1}^I \left(\exp(\beta V^g_{m,t+1} - \mu^g_{mn})\right)^{\frac{1}{\nu_g}} \right).$$

.

From (6) we know that

$$V^g_{n,t} = \tilde{T}_g \Phi^{\frac{1}{\theta}}_{Rgn,t} r^{\alpha-1}_{Rn,t} + \nu_g \ln \left( \sum_{m=1}^I \left(\exp(\beta V^g_{m,t+1} - \mu^g_{mn})\right)^{\frac{1}{\nu_g}} \right)$$

which, along with the previous equation implies that

$$V^g_{n,t} = \tilde{T}_g \Phi^{\frac{1}{\theta}}_{Rgn,t} r^{\alpha-1}_{Rn,t} + \beta V^g_{n,t+1} - \nu_g \ln(m^g_{nn,t}).$$

Iterating this equation forward and substituting we obtain

$$V^g_{n,t} = \sum_{s=t}^{\infty} \beta^{s-t} \left[ \tilde{T}_g \Phi_{Rgn,s}^{\frac{1}{\theta_g}} r_{Rn,t}^{\alpha-1} - \nu_g \ln(m^g_{nn,s}) \right]. \tag{C.10}$$

## C.8 | Calibrating Additional Parameters

I fix $T_l = 1$, and I calibrate $T_h$ in order to equalize the average wage premium in the model to the average average wage premium in the city according to the 2011 first quarter National Household Survey (*Encuesta Permanente de Hogares*). Let $\hat{WP} = 1.69$ be the observed average wage premium, I calculate $T_h$ so that

$$\hat{WP} = \frac{T_h \sum_i \Phi_{Rih0}^{1/\theta_h} \lambda_{ih}}{\sum_i \Phi_{Ril0}^{1/\theta_l} \lambda_{il}}, \tag{C.11}$$

where

$$\lambda_{ig} = \frac{L^g_{Ri}}{L^h_{Ri} + L^l_{Ri}}.$$

This results in $T_h = 1.146$.

In order to calculate $\{\rho_j\}_{j=1}^J$, and $\{A_j\}_{j=1}^J$, I first estimate $\{w^g_{j0}\}_{j=1,g\in\{h,l\}}^J$ using equation 15; the residential population vector for $t = 0$, $\{L^g_{Ri,t}\}$, by skill type; and the employment population vector for $t = 0$, $\{L^g_{Fj,t}\}$, both of which are observed data. I calculate the model-consistent wages following the algorithm used in Ahlfeldt et al. (2015). With this wage vector, and the employment population at $t = 0$, I estimate the $\{\rho_j\}_{j=1}^J$ from the following equation:

$$\frac{\tilde{L}^h_{Fj0} w^h_{j0}}{\tilde{L}^l_{Fj0} w^l_{j0}} = \frac{1-\rho_j}{\rho_j},$$

and I estimate the vector of total factor productivities, $\{A_j\}_{j=1}^J$, from the following equation:

$$A_j = \left[ \frac{\tilde{L}^l_{Fj0} w^l_{j0}}{\rho_j} [(w^l_{j0})^{\rho_j} (w^h_{j0})^{(1-\rho_j)}]^{\sigma-1} \left[ \sum_i \sum_{g\in\{h,l\}} \alpha \tilde{T}_g \left( \sum_s \left( \frac{w^g_{s0}}{d_{is0}} \right)^{\theta_g} \right)^{\frac{1}{\theta_g}} L^g_{Ri0} \right]^{-1} \right]^{\frac{1}{\sigma-1}}.$$

I will do this by first calculating

$$X = \sum_i \sum_{g\in\{h,l\}} \alpha \tilde{T}_g \left( \sum_j \left( \frac{w^g_{j0}}{d_{ij0}} \right)^{\theta_g} \right)^{\frac{1}{\theta_g}} L^g_{Ri0},$$

and then calculating

$$A_j = \left[ \frac{\tilde{L}^l_{Fj0} w^l_{j0}}{\rho_j} [(w^l_{j0})^{\rho_j} (w^h_{j0})^{(1-\rho_j)}]^{\sigma-1} [X]^{-1} \right]^{\frac{1}{\sigma-1}}$$

for each $j = 1, ..., J$

## C.9 | Derivation of Estimating Equation for Migration Elasticities

From equation 7 we know that

$$\ln(m^g_{nn,t}) = \frac{1}{\nu_g} \left( \beta V^g_{n,t+1} \right) - \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+1} - \mu^g_{mn}) \right)^{\frac{1}{\nu_g}} \right),$$

and

$$\ln(m^g_{in,t}) = \frac{1}{\nu_g} \left( \beta V^g_{n,t+1} - \mu_{in} \right) - \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+1} - \mu^g_{mn}) \right)^{\frac{1}{\nu_g}} \right).$$

Which implies that

$$\ln \left( \frac{m^g_{in,t}}{m^g_{nn,t}} \right) = \frac{1}{\nu_g} \left( \beta \left( V^g_{i,t+1} - V^g_{n,t+1} \right) - \mu_{in} \right),$$

so that

$$\ln \left( \frac{m^g_{in,t}}{m^g_{nn,t}} \right) = \frac{\beta}{\nu_g} \left( \tilde{T}_g \left( \Phi^{\frac{1}{\theta}}_{Rgi,t+1} r^{\alpha-1}_{Ri,t+1} - \Phi^{\frac{1}{\theta}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1} \right) + \right.$$
$$\left. + \nu_g \left( \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+2} - \mu^g_{mi}) \right)^{\frac{1}{\nu_g}} \right) - \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+2} - \mu^g_{mn}) \right)^{\frac{1}{\nu_g}} \right) \right) - \frac{\mu_{in}}{\beta} \right).$$

Now consider that

$$\ln \left( \frac{m^g_{in,t+1}}{m^g_{ii,t+1}} \right) = -\frac{\mu_{in}}{\nu_g} + \left( \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+2} - \mu^g_{mi}) \right)^{\frac{1}{\nu_g}} \right) - \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+2} - \mu^g_{mn}) \right)^{\frac{1}{\nu_g}} \right) \right),$$

which implies that

$$\nu_g \ln \left( \frac{m^g_{in,t+1}}{m^g_{ii,t+1}} \right) + \mu_{in} = \nu_g \left( \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+2} - \mu^g_{mi}) \right)^{\frac{1}{\nu_g}} \right) - \ln \left( \sum_{m=1}^{I} \left( \exp(\beta V^g_{m,t+2} - \mu^g_{mn}) \right)^{\frac{1}{\nu_g}} \right) \right).$$

Combining this result with the expression for $\ln \left( \frac{m^g_{in,t}}{m^g_{nn,t}} \right)$ derived above, we obtain

$$\ln \left( \frac{m^g_{in,t}}{m^g_{nn,t}} \right) = \frac{\beta}{\nu_g} \left( \tilde{T}_g \left( \Phi^{\frac{1}{\theta}}_{Rgi,t+1} r^{\alpha-1}_{Ri,t+1} - \Phi^{\frac{1}{\theta}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1} \right) + \beta \ln \left( \frac{m^g_{in,t+1}}{m^g_{ii,t+1}} \right) \right) - \mu_i n \frac{1-\beta}{\nu_g}. \tag{C.12}$$

Taking expectation at time t of equation C.12 we get

$$E_t\left[\ln\left(\frac{m^g_{in,t}}{m^g_{nn,t}}\right)\right] = E_t\left[\frac{\beta}{\nu_g}\left(\tilde{T}_g\left(\Phi^{\frac{1}{\theta}}_{Rgi,t+1} r^{\alpha-1}_{Ri,t+1} - \Phi^{\frac{1}{\theta}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right) + \beta \ln\left(\frac{m^g_{in,t+1}}{m^g_{ii,t+1}}\right)\right) - \mu_i n \frac{1-\beta}{\nu_g}\right].$$

Which can be interpreted as the linear regression equation:

$$\log\left(\frac{m^g_{in,t}}{m^g_{nn,t}}\right) = -\mu^g_{in}\frac{1-\beta}{\nu_g} + \frac{\beta}{\nu_g}\left[\tilde{T}_g\left(\Phi^{\frac{1}{\theta_g}}_{Rgi,t+1} r^{\alpha-1}_{Ri,t+1} - \Phi^{\frac{1}{\theta_g}}_{Rgn,t+1} r^{\alpha-1}_{Rn,t+1}\right)\right] + \beta \log\left(\frac{m^g_{in,t+1}}{m^g_{ii,t+1}}\right) + \omega^g_{t+1}. \tag{C.13}$$

## D | SUPPLEMENTARY QUANTITATIVE RESULTS

### D.1 | Welfare Results Under Different Home Ownership Assumptions

First, note that in the main specification of the model, I assume that residents receive a lump-sum housing income $\psi^g_{i,t}$, which represents the fraction of all housing rents that are payed to this resident, and is calculated as

$$\psi^g_{i,t} = o_{ig}\frac{E_{i,t}}{L_{Ri,t}},$$

where $o_{ig}$ is the share of homeowners of type $g$ living in residential location $i$, $E_{i,t}$ is the total expenditure on housing in $i$ at time $t$, and $L_{Ri,t}$ is the total population in $i$ at time $t$.

Under the alternative assumption that all residents, regardless of skill type, receive the same lump-sum housing income $\psi_{i,t}$, we must simply modify the interpretation of $o_{ih} = o_{il} = o_i$ to be the share of home owners living in residential location $i$ in 2010 regardless of the skill type. All the remaining equations and derivations follow exactly in the same way.

Using the same logic, if we instead assumed that all residents were renters and all rent was payed to an absentee landlord that does not consume in the city, this would be equivalent to assuming that $o_{ig} = 0$ for both $g \in \{h, l\}$. In this way, all the remaining equations and derivations do not change, and this assumption can be considered a special case of the more general model derived in the paper.

Table D.1 summarizes the aggregate welfare gains under each home ownership assumption. As we can see from this table, the aggregate welfare results do not change significantly with the different assumptions regarding home ownership. The greatest differences with respect to the assumption in the main specification of the model come, not surprisingly, when we assume that all rents are collected by an absentee landlord.

From figure D.1 we can see that the welfare gains for incumbent residents under the assumption that the lump-sum rent payment does not depend on the skill type are distributed very similarly across residential locations with different level of initial high-skill share, relative to the welfare gains under the assumption that home ownership varies by skill type within residential locations. Given that there is a high degree of correlation between home ownership rates for high- and low-skilled residents, this results is not so surprising.

We can see from figure D.2 that the assumption of an absentee landlord does change the spatial distribution of welfare gains. Although the main pattern remains the same, residential locations with lower levels of initial high-skill share saw higher gains, under this alternative assumption, the difference in welfare gains between quartiles of the initial high-skill distribution is significantly more pronounced, especially for low-skill residents. In fact, under this assumption, low-skill residents suffer (on average) welfare losses in residential locations at the fourth quartile of the initial high-skill share distribution. Implying that the welfare loss produced by the increase in rents more than erases the gains from the increase in market access for low-skill workers living in these locations.

TABLE D.1 Aggregate welfare gains for high- and low-skilled incumbents under three different assumptions regarding home ownership.

| | Aggregate Welfare Gains | |
|---|---|---|
| | High-skilled | Low-Skilled |
| Home ownership by skill type | 1.00% | 0.63% |
| Average home ownership | 1.00% | 0.64% |
| Absentee landlord | 1.05% | 0.50% |

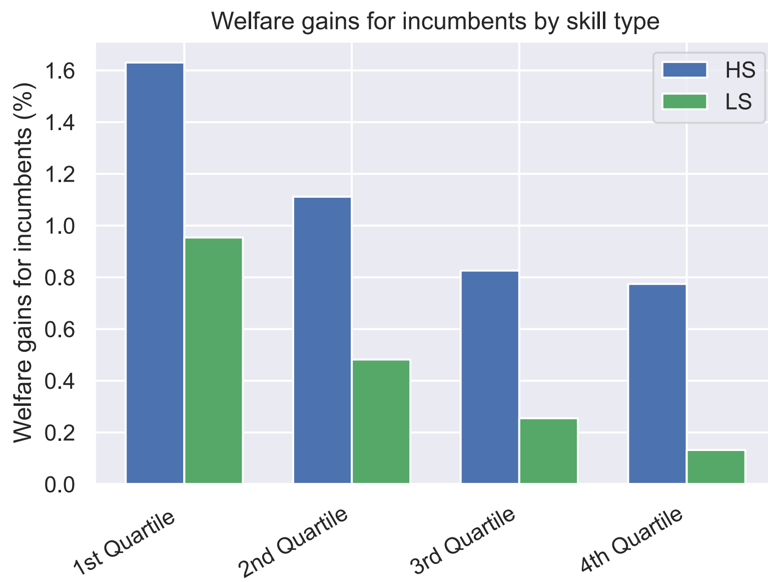


FIGURE D.1 Average welfare gains for incumbent residents at $t = 0$ by skill type and quartile of the initial high-skilled share distribution of their residential location. These welfare calculations were made under the assumption that the lump-sum rent payments were proportional to the proportion of home owners in each residential location in 2010, regardless of skill type.

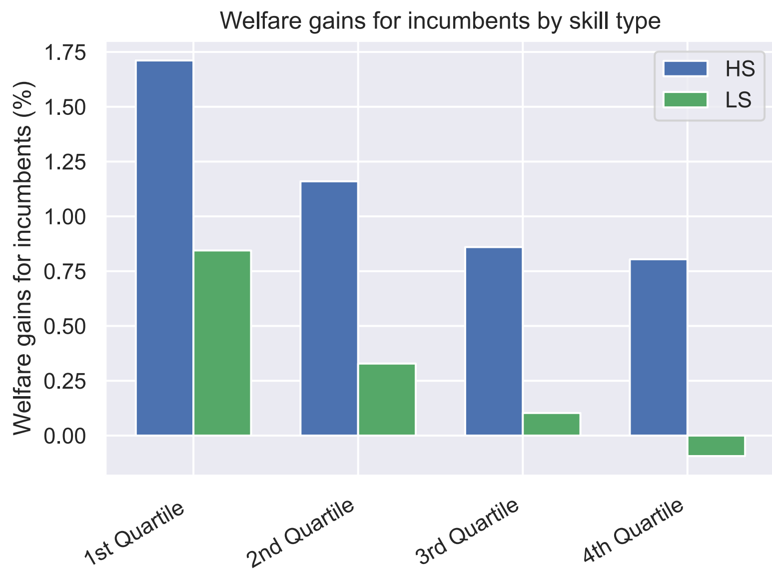


FIGURE D.2 Average welfare gains for incumbent residents at $t = 0$ by skill type and quartile of the initial high-skilled share distribution of their residential location. These welfare calculations were made under the assumption that all residents are renters and all rents are collected by an absentee landlord.

## D.2 | Model Results at the Census Tract Level

TABLE D.2 Reduced form results versus model estimation at census tract level

| Reduced Form (IV) | | Model | |
|---|---|---|---|
| $\Delta \log(\text{hs share})$ | | $\Delta \log(\text{hs share}_{BRT}) - \Delta \log(\text{hs share}_{CF})$ | |
| $\Delta \log(CMA)$ | -2.308 | $\Delta \log(CMA_{BRT}) - \Delta \log(CMA_{CF})$ | -2.79 |
| | (0.522)*** | | (0.34)*** |
| hs $\text{share}_0$ | 0.113 | hs $\text{share}_0$ | -0.05 |
| | (0.052)** | | (0.02) |
| $\Delta \log(CMA) \times$ hs sh. $\text{avg.}_0$ | 4.143 | $\Delta \log(CMA_{BRT}) - \Delta \log(CMA_{CF}) \times \text{hs share}_0$ | 5.66 |
| | (0.847)*** | | (0.65)*** |
| cons | -0.419 | cons | -0.01 |
| | (0.036)*** | | (0.01) |
| N | 2,282 | N | 2,282 |

**Notes:** The left-hand side shows the results from column 4 of table 4, while the right-hand side shows the results from the difference in difference estimation within the model at the census tract level, comparing changes between $t = 0$, and $t = 8$, and between the model with the BRT system, and the counterfactual without a BRT system put in place. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

## D.3 | Calculations for GDP Gains Net of Construction Costs

Let $X_t$ be the aggregate income in the city at time $t$,

$$\sum_i \sum_{g \in \{h,l\}} \alpha \tilde{T}_g \left( \sum_j \left( \frac{w_{j,t}^g}{d_{ij,t}} \right)^{\theta_g} \right)^{\frac{1}{\theta_g}} L_{Ri,t}^g.$$

We can calculate the net present value of the benefits of the *Metrobus*, net of construction costs by calculating the discounted sum of the excess growth in aggregate income in the city and subtracting the discounted sum of the construction costs:

$$\text{NPV}_{\text{GDP gains}} = \left[ \sum_{t=0}^{\infty} \beta^t \left( \frac{X_t}{X_0} - \frac{\hat{X}_t}{X_0} \right) \right] \text{GDP}_0 - \sum_{t=0}^{4} C_t.$$

Where $\hat{X}_t$ is the aggregate income at time $t$ under the counterfactual scenario where no BRT was put in place, $\text{GDP}_0$ is the two year GDP for 2010 and 2011 for the city of Buenos Aires, and $C_t$ is the estimated construction cost measured in 2010 dollars for the bi-annual period that corresponds to each period $t$ in the model. These construction costs are estimated by taking the average construction cost per kilometer, obtained from AGCBA (2015), and AGCBA (2019), and multiplying this number by the total number of kilometers built in each bi-annual period. Table D.3 shows the resulting estimated values. Once a net present value is obtained (measures in 2010 dollars), I calculate the fifty-year annuity payment that would lead to the same net present value, at an annual discount rate of $\beta^{-\frac{1}{2}}$, and express this annuity payment in terms of the city's 2010 GDP.

TABLE D.3 Net present value of GDP gains from the *Metrobus*

| | |
|---|---|
| NPV of benefits of BRT (mm US dollars 2010) | 13387 |
| NPV of Construction Costs (mm US dollars 2010) | 121 |
| NPV of net gains of BRT (mm US dollars 2010) | 13266 |
| NPV gains as 2010 GDP constant growth equivalent (%) | 0.4 |

**Notes:** Net present value of GDP gains from the *Metrobus* lines built between 2011 and 2017, net of construction costs.