The Aggregate and Distributional Effects of Urban Transit Networks: Evidence from Bogotá's TransMilenio*

Nick Tsivanidis†
University of Chicago Booth School of Business

October 14, 2017

PRELIMINARY AND INCOMPLETE‡

Abstract

How large are the benefits to improving transit in cities, and how are the gains shared between low- and high-skilled workers? This paper uses detailed tract-level data to analyze the construction of the world's largest Bus Rapid Transit (BRT) system–TransMilenio–in Bogotá, Colombia. First, I build a quantitative general equilibrium model of a city where low and high-skill workers sort over where to live, where to work, and whether or not to own a car. Second, I develop a new reduced form methodology derived from general equilibrium theory to evaluate the effects of the system based on a notion of "commuter market access". Third, I structurally estimate the model to quantify the effects of the system on output and welfare. I find that while the system caused large increases in welfare and output that were more than worth the costs, the gains accrued disproportionately to high-skilled workers. This suggests improving public transit is not a precise tool to target welfare improvements for the poor. Allowing for increased building densities in treated locations would have led to 50% higher welfare gains, underscoring the benefits to cities pursuing a unified transit and land use policy.

*I am extremely grateful to my committee chair Chang-Tai Hsieh and members Marianne Bertrand, Erik Hurst and Robert E. Lucas Jr. for their continuous guidance and support. I also thank Rodrigo Adao, Jonathan Dingel, Rick Hornbeck, Ralph Ossa, Felix Tintelnot and Nancy Stokey for helpful discussion and comments, as well as participants of numerous working groups at Chicago, Chicago Fed, IGC Cities and Universidad de Javeriana. I would like to express a particular thanks to Andres Villaveces who was immensely helpful in providing data for this project, as well as to Marco Antonio Llinas Vargas, Helmut Menjura and those at DANE who granted access to microdata. I gratefully acknowledge financial support from the Chicago Booth Social Enterprise Initiative, the University of Chicago Urban Network, the Energy Policy Institute at Chicago and the North American Regional Science Council. The results in this paper use confidential microdata from DANE and have been reviewed to ensure no confidential information has been disclosed. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of DANE. All errors are mine.

†Ph.D. candidate, Booth School of Business. Email: ntsivanidis@uchicago.edu.

‡PLEASE DO NOT CITE OR CIRCULATE

1 Introduction

How large are the economic gains to improving public transit systems within cities and how are they distributed between the rich and poor? With 2.5 billion people predicted to move into cities by 2050, mostly in developing countries, governments will spend vast sums on commuting infrastructure to reduce congestion associated with this rapid urban growth.[1] The reliance of the poor on public transit suggests they may benefit the most. Yet measuring the benefits of these systems is challenging: while individuals save time on any particular commute, their decisions of where to live and work will likely change as new alternatives become more attractive and land and labor markets adjust. The lack of detailed intra-city data in less developed countries coinciding with the construction of large transit systems makes the task of evaluating their causal impact even more daunting.

This paper exploits uniquely detailed spatial data I construct before and after the opening of the world's largest Bus Rapid Transit (BRT) system–TransMilenio–in Bogotá, Colombia to make three contributions to our understanding of the aggregate and distributional effects of urban transit systems. First, I build a quantitative general equilibrium model of a city where low and high-skill workers sort over where to live, where to work, and whether or not to own a car. Second, I develop a new reduced form methodology derived from general equilibrium theory to evaluate the effects of changes in commuting networks in cities. I show that a wide class of models (nesting a special case of my own) contain a log-linear relationship between outcomes and the transit network through "commuter market access". For individuals this reflects access to jobs while for firms it reflects access to workers; both are easily computed using data on employment and residence across the city. I use the implied regression framework to empirically evaluate the effect of TransMilenio on outcomes such as population, employment and house prices. Third, I instrument for changes in market access to identify the model's structural elasticities, and use the estimated model to quantify the effects of the system and counterfactual policies.

I have three main results. First, changes in commuter market access perform better than traditional distance-based approaches in explaining the heterogeneous adjustment of population, employment and house prices to TransMilenio. This suggests the framework can be applied elsewhere to improve predictions about the effects of transit on the spatial organization of cities. Second, I find the system provided large aggregate gains for the city, increasing average welfare by 3.8% and output by 3.08% (net of construction and operating costs) at my most conservative estimates. However, these gains would have been around one third larger had the government implemented a complementary change in zoning policy to allow housing supply to respond where it was most needed. Third, I find that high-skilled workers benefitted slightly *more* than the low-skilled, suggesting (perhaps surprisingly) that improving public transit is not a precise tool to target welfare improvements for the poor.

To build intuition, I find certain key channels explain the incidence of public transit across worker groups. The first is mode choice: the group that relies on public transit benefits more. This operates in favor of the low-skilled who are poorer and less willing to pay for cars. The second is the elasticity

[1] For example, figures in McKinsey (2016) suggest a need for $40 trillion of spending to close the transport infrastructure gap.

of commuting decisions to commute costs, which determines how willing individuals are to bear high commute costs to work in a particular destination. In the model, this is determined by the heterogeneity of workers' match-productivities with firms in different locations. For example, a high-skilled IT worker may be more willing to incur a long commute to an especially well-paid position. A low-skilled cleaner who receives similar wages wherever they work may instead substitute towards less costly alternatives. In the data I find high-skilled workers are less sensitive to commute costs, and thus bear a greater incidence along this channel.[2] Lastly, there are geographic factors specific to the transit network and the city in question: where house prices appreciate following the system's construction, whether it connects locations of dense residence and employment, and how these characteristics differ where each group lives and works. In Bogotá these forces worked in favor of the high-skilled. That the net effect of these forces benefits the high-skilled underlines the importance of using a general equilibrium model to fully account for the channels through which investments in transit affect welfare.

Opened in 2000, TransMilenio is the world's most used BRT system with a daily volume of over 2.2 million trips. The system operates more like a subway than the informal bus system that preceded it: buses run in single-use lanes with express and local services, passengers pay at station entrances using smart cards, and buses are boarded at stations rather than at roadside. BRT provides an attractive alternative to subways in rapidly growing developing country cities since they are able to deliver similar reductions in commuting times at a fraction of the cost, and are much faster to build.[3] I collect new sources of data covering 2,800 census tracts on residence, employment, commuting patterns, and land markets spanning the system's construction.

Prior to TransMilenio's opening, low-skilled workers commuted using a network of informal buses which were on average 30% slower than cars. To understand the implications of improving public transit on worker welfare, I develop a quantitative general equilibrium model of a city where workers choose where to live, where to work, and how to commute. Non-homothetic preferences mean that the high-skilled live in high amenity neighborhoods and are more likely to own cars. Individuals work in different locations due in part to differential demand for skills from firms across the city. For example, retail and manufacturing establishments demand more low-skilled workers while real estate and financial brokerage businesses rely on the high-skilled. Differences in residential locations, commuting elasticities and the relative demand for worker skills turn out to be crucial in determining the distributional effects from improving transit.

A large literature estimates average treatment effects of transit based on proximity to stations. In contrast, I show that for a wide class of models featuring a gravity equation for commute flows the full direct and indirect effects of the entire transit network on firms and workers can be summarized by a

[2]This does not imply one should observe the high-skilled taking longer commutes. It says that from any location of residence, the low-skilled will be less willing to commute to locations with high commute costs ceteris paribus. Average commute times and distances are greater for the low-skilled in Bogotá (as in many other cities) since they live further from high employment densities in central areas.

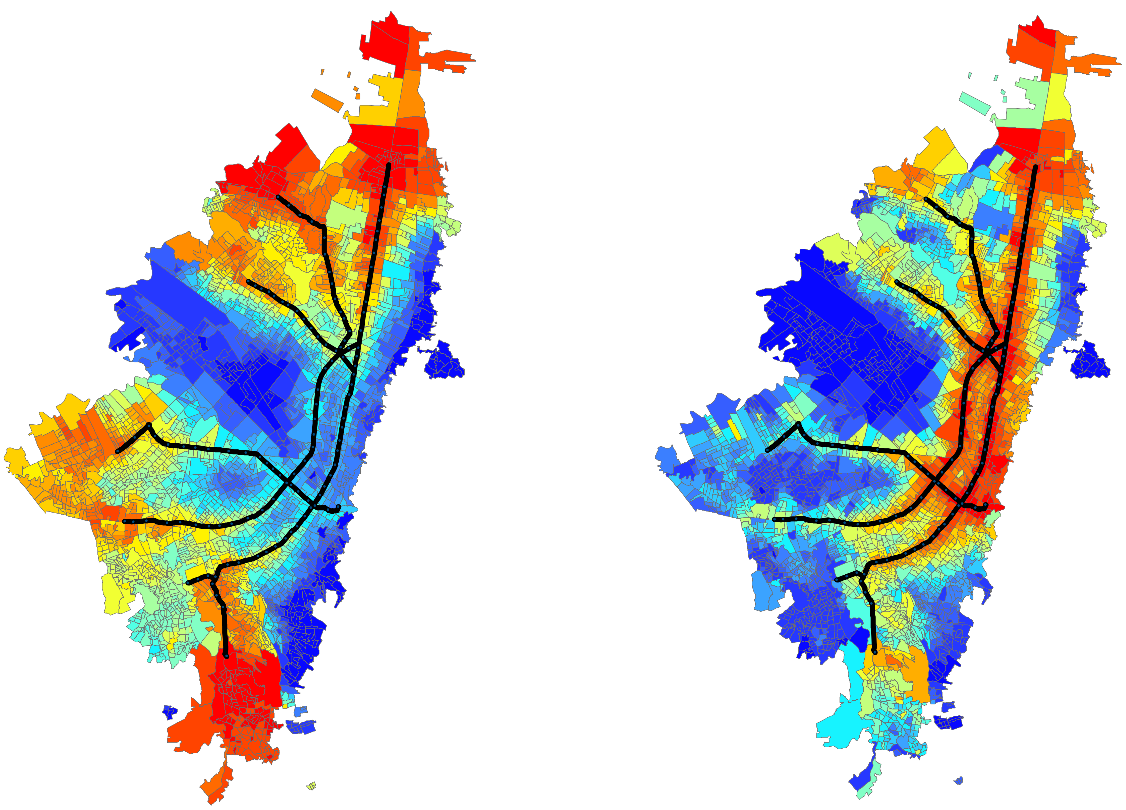
[3]For example, the per mile cost of construction of the subway in Colombia's second largest city, Medellín, was 10 times that of TransMilenio, all the while maintaining similar system speeds. Moreover, TransMilenio took less than 18 months to construct and open, compared to the 12 years taken by Metro Medellín.

single variable: commuter market access. Importantly, these terms are easily computed using data on residence and employment in the city, as well as a measure of commute costs. Figure 1 plots the change in commuter market access as a result of TransMilenio. For residents, this captures access to high paying jobs through the commuting network. Tracts towards the edge of the city far from the high density of jobs in the center experienced a much larger improvement in market access. For firms, it reflects access to workers. Central locations benefit most from increased access to workers supplied along all spokes of the network.[4] Changes in market access capture a wide heterogeneity in treatment effects from TransMilenio (seperately for workers and firms) that would be missed by looking at distance to the system alone.

Figure 1: Change in Commuter Market Access from TransMilenio

(a) Residential CMA

(b) Firm CMA



Note: Plot shows the baseline instrument for the change in CMA induced by holding population and employment fixed at their initial level and changing only commute costs. Tracts are grouped into 20 even groups based on the their change in CMA, and shaded accordingly with hotter colors indicating a larger increase in CMA. Black line shows the TransMilenio routes as of 2006. See section X for full discussion.

In a special case of my model, the equilibrium has a reduced form in which outcomes such as population, employment and house prices can be written as log-linear functions of commuter market access. Moreover, I show any model with log-linear demand for residents and workers across the city has a similar representation. The framework is therefore isomorphic to a number of alternative assumptions over production technologies, housing supply and worker preferences. I use the implied regression specifications

[4] Note that firm commuter access also increases away from the center the North of the city. This is due to the high density of (low-skill) workers in the South, as discussed in the next section.

to empirically evaluate the impact of TransMilenio through improvements in market access. I address non-random route placement by predicting the location of TransMilenio in two ways. First, I use a historical tram system that was last placed in 1921. Second, using engineering estimates for the cost of building BRT on different types of land use I solve for least cost construction paths connecting terminals at the end of the system with the central business district as was the intent of the government. These routes are then used to instrument for changes in market access.[5] I run falsification tests exploiting the timing of station openings as well as using residual variation in market access conditional on distance to stations and instruments to provide additional evidence that the effects are causal.

I find that changes in commuter market access perform well in predicting the heterogeneous response of population, employment and land markets in response to TransMilenio. Improvements in residential commuter market access also led to growth in both commute distances and wages, supporting the intuition that it measures access to jobs. Interestingly, the system caused a re-sorting of workers by skill group, with the high-skilled moving into high amenity, expensive neighborhoods in the North while the low-skilled moved into the poorer neighborhoods of the South. This suggests that transit has the potential to increase residential segregation across worker groups within cities.[6]

In the final part of the paper, I structurally estimate the full (non-linear) model. Some parameters, such as spillovers in productivities and amenities, are challenging to estimate in cross-sectional data. For example, a location's productivity may be a cause or consequence of the number of workers employed there. Since the supply of workers and residents in the model is a log-linear function of market access, my instruments provide exogenous variation in the number of individuals living and working across the city. This allows me to identify these key elasticities through a GMM procedure.

I estimate an agglomeration elasticity roughly three times the size of median estimates in the US but close to other studies using experimental approaches. I provide one of the first estimates in a less developed country city, suggesting that these forces can be particularly strong in poorer countries. I also find a substantial elasticity of amenities to the college share of residents, reflecting the endogeneity of neighborhood characteristics such as crime. I recover the remaining structural parameters that vary by skill group, allowing me to quantify the distributional effects of TransMilenio.

The model performs well in matching a number of non-targeted moments such as income, employment and commute flows by skill group. The amenities and productivities recovered from the model correlate well with observable proxies like local homicide rates. I also check the robustness of my quantitative results to alternative parameter values, and incorporate home ownership and the employment of domestic servants in model extensions. Lastly, I use the extreme assumptions of either zero or infinite mobility costs between Bogotá and the rest of Colombia to bound the impact of TransMilenio on welfare, population, land rents and output.[7]

The system led to large aggregate gains in worker welfare and output. Productive locations were able

[5]Formally, my instrument measures the change in residential and firm market access that would have occured had TransMilenio been placed along the routes of each instrument. I exclude the targeted neighborhoods surrounding portals and the CBD from the analysis.

[6]Heilmann (2016) notes a similar finding following the opening of the light rail system in Dallas.

[7]These are known as the open and closed city assumptions in the urban economics literature.

to "import" more workers through the commuting network. This suggests better transit can improve the spatial allocation of labor within cities. The increase in output greatly exceeded construction and operating costs, supporting the notion that BRT can be highly a worthwhile investment. Population decentralized, as improved access to jobs made distant neighborhoods more attractive.

My finding that the high-skilled benefit slightly more than the low-skilled underscores the need to evaluate the full range of determinants of worker welfare rather than considering mode choice alone. However, the effect of different parts of the network is heterogeneous: lines serving poor neighborhoods in the South of the city, as well as a cable car connecting hillside slums with a TransMilenio slated to open in 2018, disproportionately benefit the low-skilled. This suggests there is no universal answer to the distribution of gains from improving transit. The conclusion that the low-skilled benefit less than what is implied by mode choice alone remains generalizable, though, since existing evidence suggests that the key elasticities that vary across groups have similar relative magnitudes in other countries and Bogotá is by no means unusual in its spatial configuration.[8] Moreover, the methodology developed in this paper can be applied to the specific geography and transit systems of any city where data on residence and employment are available to predict the effects of new infrastructure.

The estimated model allows to me assess the impact of counterfactual policies. In the first exercise, I simulate the removal of the feeder bus network that transports individuals in the outskirts of the city to terminals at the end of lines using existing road infrastructure at no additional fare. This part of TransMilenio accounts for just under 40% of the total welfare gains from the system. This underlines the potential for large benefits to providing cheap, complementary services that reach residents in outlying but dense residential areas, thereby reducing the "last mile problem" of travelling between stations and final destinations.

In the second exercise, I evaluate the welfare impacts of a "Land Value Capture" scheme under which development rights to increase building densities near stations are sold by the government to developers, and the extent to which the revenues could have financed the system's construction. Similar schemes have seen great success in Asian cities such as Hong Kong and Tokyo.[9] In contrast, one of the main criticisms of TransMilenio was that the city experienced such a large change in transit without any adjustment of zoning laws to allow housing supply to respond where it was needed. I compare the effects of two alternative policies. The first increases permitted densities within a certain distance of stations, while the second allocates the same number of permits based on predicted growth in commuter market access. The market access based policy increases welfare gains from TrasnMilenio by around 28%, while government revenues cover 127% of the construction costs. Under the distance-based scheme welfare and government revenues only increase by 15% and 53% respectively. These policies disproportionately benefit low skilled workers by dampening house price appreciation towards the edge of the city where they live. My findings suggest large returns to the pursuit of an integrated transit and land use policy, and highlight the

[8]In the model, the elasticity of commuting decisions to commute costs is governed by the heterogeneity of worker match productivity with firms. While I estimate these parameters using data on commute flows, Lee (2015) provides estimates for this Frechet shape parameter across worker education categories in the US, Brazil, India and Mexico. She reports similar relative differences across the college and non-college educated.

[9]See Suzuki et. al. (2015) for an excellent review of these policies.

applicability of commuter market access as an instrument to guide government policy.

The rest of the paper proceeds as follows. Section 2 discusses the paper's contribution to the literature. Second 3 presents the context of Bogotá and TransMilenio. Section 4 develops the model and section 5 outlines the reduced form framework it delivers. Section 6 describes the data. Section 7 presents the reduced form estimation results while section 8 structurally estimates the non-linear model which section 9 uses to quantify the effects of TransMilenio. Section 10 simulates the effects of counterfactual policies. Section 11 concludes.

2 Relation to Previous Literature

This paper contributes to the literatures on urban economics and economic geography.

Within a large body of work that documents the association between transit and urban structure, a smaller strand exploits the opening of new systems to establish a causal relationship. These papers typically measure changes in population and property prices as a function of distance to the central business district (Baum-Snow 2007; Baum-Snow et. al. 2017; Turner and Gonzalez-Navarro 2016) or distance to stations (Gibbons and Machin 2005; Glaeser et. al. 2008; Billings 2011). However, when spatial units are interlinked (as is likely the case within cities), spillovers across treatment and control locations confound causal inference from these comparisons.[10] If these linkages lead to heterogeneous responses, then average treatment effects estimated in one context will no longer be externally valid in another. My approach confronts these challenges by explicitly measuring the full direct and indirect effects of changes in the transit network between connected locations, allowing for a causal identification of transit connections that captures heterogeneous reponses as a function of city geography.

Dating back to at least Harris (1954), a long literature has examined the link between access to goods markets and economic development across regions within countries (see Redding 2010 for a review). Recent work has combined data at the regional level with natural experiments that impact the cost of trading goods across space on local outcomes such as factor prices and population through market access (Redding and Sturm 2008; Donaldson forthcoming; Donaldson and Hornbeck 2015). Yet these models contain no notion of commuting within cities, and therefore are silent on the effects of infrastructure that reduces the cost of moving people rather than goods across space. I show that reduced form relationships between a different notion of access to workers and jobs can be derived in an urban setting wherever data on residential population and employment are available.[11] The structural part of this paper uses the

[10]More precisely, the identification of causal impacts of transit infrastructure from relative comparisons requires that the true treatment effect on one location is independent of whether any other is treated. In other words, locations must not interact. This is known as the stable unit treatment value assumption (Rubin 1980). While this can be overcome if spatial units are sufficiently aggregated until the point where interaction across locations is minimal, this is unlikely to hold for highly local geographies within cities. See Donaldson (2015) for an excellent discussion.

[11]Within the urban planning literature, the notion of "accessibility" as a determinant of land use and prices within cities dates back as far as Hansen (1959). However, the idea has received little attention within urban economics which instead has focused on models where distance from the central business district is a summary statistic for the spatial configuration of the city (Alonso 1964; Mills 1967; Muth 1969, Lucas and Rossi-Hansberg 2002). A small handful of recent papers measure the correlation between residential land prices and an accessibility index given by the distance weighted sum of employment

similar reduced form moments to guide estimation of a non-linear model, allowing me to quantify the distributional effects of transit across worker skill groups.

This paper contributes to the growing body of quantitative work featuring gravity equations for commute flows (Ahlfeldt et. al. 2015; Allen et. al. 2015; Monte et. al. 2017; Owens et. al. 2017). I show that these models share a common measure that summarizes the effect of transit on the supply of residents and workers across the city. These measures are easily computed using data and population and employment, and can be used to guide reduced form analysis of changes in the network. This paper is also the first to structurally estimate a gravity model by combining a large-scale construction of commuting infrastructure with moments derived directly from log-linear relationships between market access and model outcomes.[12] The approach is particularly useful, since I can explicitly control for the direct effects of the instruments and rely only on residual variation for identification.[13]

I extend the approach of these papers to incorporate multiple types of workers, firms and transit modes, which is necessary to assess the distributional effects of urban policies.[14] Importantly, I show that by incorporating multiple types of firms with different demand for worker groups, one can invert the model to solve for unobserved group-specific wages that rationalize the data. Allowing for differences in wages between skill groups across the city turns out to be quantitatively important for assessing the distributional effects of transit.[15]

A large literature has studied the relationship between population density and outcomes such as wages and productivity (see Rosenthal and Strange (2004) for a review). A smaller strand of work uses potentially exogenous sources of variation in the density of economic activity to estimate these spillovers (Greenstone et. al. (2010), Kline and Moretti (2014), Ahlfeldt et. al. (2015)). Other papers examine how amenities depend on the composition of local residents (Bayer, Ferreira and McMillan (2007), Guerrieri, Hartley and Hurst (2013), Diamond (2015)). To my knowledge, this paper provides the first intra-city estimates of productivity and amenity spillovers within a developing country city, using changes in the transit network as labor and resident supply shocks as sources of identifying variation.

(Osland and Thorsen 2008; Ahlfeldt 2011; McArthur et. al. 2012). In contrast, I show similar indices reflecting market access can be derived from general equilibrium theory (seperately for firms and workers), and use a change in the transit network to estimate its causal impact on a range of urban outcomes.

[12]Severen (2016) estimates a model similar to Ahlfeldt et. al. (2015) using the expansion of the Los Angeles subway, but does not use the market access approach developed in this paper in his methology.

[13]Suppose one is estimating the effect of transit on land prices and instruments for the location of the line in question using a historical route. One requirement for validity of the instrument is that the historical system has no direct effect on the outcome. However, this has to be addressed indirectly through other variables since distance itself is perfectly collinear with the instrument. By constructing an instrument for the change in market access had TransMilenio been constructed along the tram route, I am able to directly control for distance to the tram and exploit only residual variation in my instrument for identification.

[14]To my knowledge, Redding and Sturm (2016) is the only other paper in the literature to incorporate multiple types of workers. However, they test only qualitative predictions of their model. I also incorporate non-homothetic demand for cars and amenities, rather than generating sorting through differences in preferences alone.

[15]Without this additional structure, in order to solve for wages one would either need to assume wages are identical for each worker group in every location (i.e. skill groups are perfect substitutes in production) or observe both residence and employment by skill group. In most contexts, comprehensive data on employment by skill group at small spatial scale is prohibitively difficult to obtain.

3 Background

Bogotá is the political and economic center of Colombia, accounting for 16% and 25% of the country's population and GDP respectively. Its population of 8 million inhabitants makes it the world's 9th densest, yet there remains a stark divide between rich and poor.[16] In this section, I provide some background on the structure of the city and its transit system.

3.1 Structure of Bogotá

Residence and Employment Bogotá is characterized by a high degree of residential segregation between the rich and poor. Defining high skill or college workers as individuals who have completed some post-secondary education, panel (a) in Figure 2 plots the share of college residents within a census tract in 1993.[17] The high-skilled are much more likely to live in the North, with low-skilled workers located primarily in the city's South and periphery. Panel (b) shows that these poorer neighborhoods have a much higher population density, reflecting the concentration of smaller housing units that are crowded in.

High- and low-skilled residents work in different kinds of jobs and neighborhoods. Table 1 shows the share of workers employed in each 1-digit industry with post-secondary education. Workers in domestic services, hotels and restaurants, manufacturing and retail are relatively unskilled, while those in real estate, education and financial brokerage tend to be high-skilled. These jobs are located in different parts of the city. Defining high-skill intensive industries as those with college employment shares above the median, Figure 3 shows that while overall employment is concentrated along two bands to the west and north of the city center, high-skill intensive industries are located more towards the North of the city.

Taken together, this shows substantial differences in where the high- and low-skilled live and work in the city.[18]

Commuting Prior to TransMilenio In 1995 the average trip to work in Bogotá took 55 minutes, more than double the average commute in US cities. The vast majority of these commutes was taken by bus (73%), followed by car (17%) and walking (9%).[19] Despite its importance, public transportation in the city was highly inefficient due in large part to its industrial organization. The government allocated the

[16]Colombia is the 11th most unequal country in the world according the ranking of Gini coefficients from the World Bank for most recently available year. The income distribution in Bogotá has a slightly higher Gini than the country as a whole (author's calculation using GEIH data from DANE in 2014). Other figures from DANE.

[17]All datasets are described in detail in Section 6. In this section, population data comes from the 1993 census, employment location data comes from the 1990 economic census, other employment data is from DANE's GEIH and ECH labor surveys and mobility data is from DANE's mobility surveys.

[18]Given the differential spatial distribution of jobs by skill-intensity, it may be that the different residential patterns of skill groups simply reflect better access to industries with greater demand for them (and therefore offer a higher wage). In the appendix, I show that the this channel only explains a small portion of the overall variation in the college share across the city, suggesting that a portion of the observed residential sorting is likely driven by neighborhood attributes other than access to jobs.

[19]The average commute in US cities has increased from 21 minutes in 1980 to 26 minutes in 2015. For an analysis using US census data, see Ingraham, Christopher. 2016 "The astonishing human potential wasted on commutes." Washington Post, Feb. 26. Data from Bogotá comes from the 1995 Mobility Survey. Bicycles and motos account for the remaining 1% of commutes.

administration of routes to companies called "afiliadoras" which acted as intermediaries between the government and bus companies. Afiliadoras sold sell slots to run their routes to bus operators, but since their profits depended only on the number of buses the result was a huge over-supply of vehicles. Moreover, low enforcement meant that up to half of the city's bus fleet operated illegally (Cracknell (2003)).[20] Disregard of the use of bus stops promoted boarding and alighting along curbs, further reducing traffic flows.

The result was that while the crowding of Bogotá's streets slowed traffic overall, buses were much slower than cars. Table 2 uses commuting microdata from the Mobility Survey to explore how speeds compared between buses and cars in 1995. Column (1) shows that commutes by car were around 33% faster than by bus, a result robust to controlling for differences in trip composition by including trip origin-destination fixed effects in column (2). However, the burden of slow public transit fell disproportionately on the city's low-skilled population. Column (3) of the shows that Bogotanos without post-secondary education were about 32% and 24% more likely to use buses as opposed to cars. These differences are robust to controlling for differences in the composition of commutes in column (4).[21]

This dependence on slow public transportation meant that the poor faced a very different distribution of commute times than the rich.

3.2 TransMilenio: The World's Most Used BRT System

Background At the start of his first term as Mayor of Bogotá, Enrique Peñalosa wasted no time in transforming the city's transit infrastructure. TransMilenio was approved in March 1998, its first phase opening a mere 21 months later adding 42 km along Avenida Caracas and Calle 80, two arteries of the city.[22] Phases 2 and 3 added an additional 70km in 2006 and 2011, creating a network spanning the majority of the city (Figure 4). Today the system is recognized as the "gold standard" of BRT systems and with more than 2.2mn riders a day using its 147 stations it is the most heavily patronized system of its kind in the world (Cervero (2013)).[23] Its average operational speed of 26.2kmh reported during phase 1 is on par with that of the New York subway (Cracknell (2003), Johnson (2010)), and provided a pronounced improvement on reported bus speeds of 10kmh on the incumbent bus network (Wright and Hook (2007)).

The system involves exclusive dual bus lanes running along the median of arterial roads in the city segregated from other traffic.[24] In contrast to the informal network that preceded it, buses stop only at stations which are entered using a smart card so that fares are paid before arriving at platforms. Dual lanes allow for both express and local services, as well as passing at stations. Accessibility for poorer

[20]The Department of Mobility estimated to be more than double the amount actually required by the city (e.g. http://www.dinero.com/edicion-impresa/negocios/articulo/buses-nuevo-modelo-para-bogota/17889). A typical practice through which bus companies avoided government controls was through duplication of license plates and vehicle documentation.

[21]The results are the same if workers are compared by income rather than education, as shown in the appendix.

[22]In many cases anticipation of a system may predate its inauguration. However, TransMilenio went from a "general idea" to implementation in only 35 months (Hidalgo and Graftieux (2005)).

[23]For comparison, the London tube carries 5 million passengers per day over a network of 402km, giving it a daily ridership per km of 12,000 compared to TransMilenio's 20,000.

[24]As shown in Figure 4, concrete barriers separate TransMilenio from other lanes and have helped the city to achieve essentially complete compliance.

citizens in the urban periphery is increased through a network of feeder buses that use existing roads to bring passengers to "portals" at the end of trunk lines at no additional cost. Free transfers and a fixed fare further enhance the subsidization of the poor while the government sets fares close to those offered by existing buses.[25]

There are two main reasons why BRT provides an attractive alternative to subways in rapidly growing cities. First, it delivers similar reductions in commuting times at a fraction of the cost: the average per kilometer construction cost is one-tenth of rail (Menckhoff (2005)). Second, BRT is much faster to construct. An illustrative comparison is that of TransMilenio and Metro de Medellín, the subway system in Colombia's second city. While both achieve similar system speeds, Medellín's metro cost 11 times as much as TransMilenio and took 12 years from announcement to opening.[26] BRT has the additional advantage that it can be put down quickly and cheaply to address congestion in rapidly growing cities, but in the years after can be easily repurposed into roadways if subways are built or can remain as part of a multimodal public transit network.[27]

While BRT is not without drawbacks,[28] these features have led to systems being built in more than 200 cities, the vast majority constructed over the past 15 years in Latin America and Asia (BRT Data (2017)).

Route Selection and System Rollout The corridors built during the first phase of the system were consistently mentioned in 30 years of transportation studies as first-priority for mass transit (World Bank (2013)). The city conducted a planning study to reconfirm these suggested routes and identify new ones based on (i) current and future demand level and (ii) expected capital costs. The result was a plan that aimed to connect the city center with dense residential areas in the north, northwest and south of the city (Hidalgo and Graftieux (2005)). Since the cost component was an important determinant of route selection, final lines were placed along wide arterial roads that were cheaper to convert. The number of car lanes was left unchanged either because existing busways were converted or due to road widening.[29]

[25]For example, in 2011 (the only year where fare information is reported in the Mobility Survey), the average bus fare is 1400 COP compared to the 1700 COP fare on TransMilenio. While the fare difference of 21.4% is non-trivial, this does not reflect the free transfers across trunk and feeder lines not offered by the existing bus network leading the relative cost of TransMilenio trips to fall for long commutes.

[26]The difference in times from planning to opening between BRT and subways is not specific to Bogotá. While a comprehensive source on construction times is hard to find, stories of such instances are not. In India planning for subways in Delhi and Bangalore started 18 and 8 years before inauguration respectively, while the BRT in Ahmedabad took only 4 years. New York's second avenue subway line opened on January 1st 2017 having been originally proposed in 1919. In Bogotá, there were a total of 10 attempts to introduce heavy rail between 1947 and 1997, thwarted by high capital costs and vested interests of the public transportation sector (Lleras (2003)).

[27]Many cities currently operate both subways and BRT system, for example Mexico City, Medellín and Guangzhou. Thus, the debate does not need to center around choosing between BRT or a subway, since the two can be complementary both over time and over space within cities.

[28]In recent years complaints of overcrowding on TransMilenio buses and stations have increased (World Bank (2010)), and ultimately BRT becomes congested more quickly than subways as riderships rise.

[29]See Hidalgo and Graftieux (2005) for a discussion of existing busways on phase 1 corridors, and Wright and Hook (2007) who report road widening during phase 2. Inspection of satellite images confirms that the number of road lanes for other traffic was unchanged (see appendix for some examples). That certain routes already contained median busways did not mean that there was efficient bus transit available along them (e.g. Avenida Caracas). While these lanes shared many similar features to TransMilenio, including dual bus lanes and bus stops, within a few years of the opening in 1990 the "the scheme became anarchic as, for example, (i) buses competed for passengers and this, together with little effective stop regulations, resulted in

Two features of the choice process merit emphasis. First, having identified neighborhoods towards the city's periphery to be connected with the center, final routes were chosen to a large extent by the desire to minimize construction cost. Second, lines were far cheaper to construct along the widest arteries of the city, whose availability was limited and determined in large part by the city's historical evolution. I leverage both in constructing instruments for the system's layout.

A notable feature of TransMilenio was that it was rolled out so quickly, primarily to complete a portion of the system within mayor Peñalosa's term that ran between 1998 and 2001.[30] The unanticipated nature of the system's construction, combined with the staggered opening of lines across three phases, provide additional sources of time series variation I will use in my analysis.

Finally, one central criticism of TransMilenio was its singular focus on improving urban mobility without coordinated changes in land use regulation (Bocajero et. al. (2013)). As a result, I show in the appendix that housing supply did not respond to the system's construction. An integrated land use and transit policy tailored towards increasing housing densities near stations promotes a more efficient urban structure where many residents can take advantage of improved commuting infrastructure. Cities such as Hong Kong and Tokyo have had great success in implementing Land Value Capture (LVC) schemes which increase permitted densities around new stations but charge developers for the right to build there. These policies achieved the dual aim of increasing housing supply and raising revenue to finance the construction of the system.[31] In counterfactuals, I quantitatively assess the welfare gains from TransMilenio had Bogotá pursued a similar policy.

Trip Characteristics and Effects on Congestion In the appendix, I provide additional details on the way in which TransMilenio is used and its effects on other modes which I briefly summarize here. First, TransMilenio is a quantitatively important mode of transit that is more likely to be for longer trips compared with other modes.[32] Second, TransMilenio is disproportionately likely to be used for commutes to work rather than leisure trips when compared with other modes, motivating the focus on access to jobs in this paper. Third, TransMilenio use appears to have come primarily from substitution away from buses for commuting. Fourth, conditional on car ownership the rich and poor are equally likely to use TransMilenio, consistent with the similar fares charged compared to traditional buses.

In general, BRT is likely to have complex and ambiguous effects on the speeds of other modes. Holding commute choices constant, higher passenger capacity of BRT should decrease the total number of vehicles needed to transport the same volume of commuters. At the same time, roadspace may be taken

bus stop congestion and hazardous operating conditions, (ii) buses without a license to operate on Av. Caracas were attracted to the busway seeking passengers and (iii) the bus volume was so big and bus standards were so poor that air pollution became a serious issue" (Cracknell (2003)).

[30] Peñalosa's upheaval of the status quo faced entrenched opposition both from the incumbent bus industry and car owners, ultimately leading him to be voted out of office in 2001.

[31] By increasing the response of housing quantities rather than prices, these policies also shift some of the incidence of the infrastructure from land owners to city residents. For a comprehensive review of Land Value Capture schemes, see Suzuki et. al. (2015).

[32] This is commensurate with the fixed time costs of entering and exiting stations, and time spent walking between stations and trip origins and destinations.

away from cars which might increase congestion (though this was not the case with TransMilenio). There will then be a web of equilibrium effects in which commuters substitute between modes in response to the change in times and vehicle volumes along each part of network adjust to any change in least cost routes between locations. Both data limitations and the challenge of incorporating these forces within a general equilibrium model put a full analysis of these forces beyond the scope of this paper. However, in the appendix I use commuting microdata to compare changes in speeds for cars and buses along routes where TransMilenio was built compared to other control trips and find no significant differences. This suggests my abstraction from the effects of TransMilenio on other mode speeds appears to be a reasonable approximation to reality.[33]

4 A Quantitative Model of a City with Heterogeneous Skills

This section presents a general equilibrium model of a city. High and low skill workers decide where to live, where to work, and how to commute between a large number of discrete locations. Individuals are attracted to neighborhoods with nice amenities, good access to jobs and low house prices. Public transit is available to everyone to commute between home and work, but only those willing to pay to own a car have the option to drive. Firms from multiple industries are located across the city and produce using labor and commercial floorspace. Some locations are more productive than others. Each industry differs in its demand for workers: for example, hotels and restaurants demand more low-skilled workers while financial brokerage firms require more high-skilled. Since industries may be located in different places, wages for low- and high-skill workers will differ across the city. Each location has a fixed amount of floorspace supply which landowners allocate to either residential or commercial use. In equilibrium, the price of floorspace, the share alloacated to each use and wages adjust to clear land and labor markets.

The setup differs from recent quantitative urban models (e.g. Ahlfeldt et. al. (2016)) along two key dimensions. First, I add in multiple skill groups of workers which allows me to assess the distributional effects of the system. Second, I incorporate non-homothetic demand for cars and residential amenities to match the sorting patterns documented in the data.

Despite the interactions between labor and land markets across thousands of locations that occur through the city's commuting network, there will be a single measure that summarizes the effect of the entire network on outcomes in any location given by its "commuter market access". In a simplified case of the model, equilibrium outcomes such as population, employment and floorspace prices can be written as log-linear functions of commuter market access. I use the implied regression framework to empirically assess the effect of TransMilenio in Section XX. In Section XX, I structurally estimate the full non-linear

[33]While the fact that the number of car lanes was left unaffected by TransMilenio may well have been a factor, this finding is consistent with two recent results in the transportation literature. First, Akbar and Duranton (2017) estimate congestion in Bogotá and find that during times used for commuting the elasticity of speed with respect to the number of travelers is a mere 0.06, suggesting imperceptible in trip times in response to changes in traffic. Their interpretation is that during rush hours the city's streets are already so crowded that variation in vehicle numbers has little effect on speeds. Second, Duranton and Turner (2011) find that for the US vehicle-kilometers travelled (VKT) increase one for one with roadway lane kilometers, and as a further implication, find no evidence that the provision of public transportation affects VKT.

model and use it to quantify the aggregate and distribiutional effects of the system.

4.1 Model Setup

The city is comprised of a discrete set of locations $i \in I$. Locations differ by their total amount of floorspace H_i, which can be used for either residential or commercial purposes, productivities, amenities as well as their access to the transit network which determines the time it takes to reach any other location in the city.[34]

The city is populated by different worker skill groups indexed by $g \in G = \{L, H\}$, each of which has a fixed population $\bar{L}_g$.[35] Each worker has an idiosyncratic preference for each combination of where to live and whether or not to own a car, as well as a match productivity with firms in each location, and chooses the combination that maximizes their utility. I assume timing is such that workers first choose where to live and whether or not to own a car, and then choose where to work.[36] Firms in different industries $s \in S$ produce in a location using labor and commercial floorspace under perfect competition. Absentee land lords own floorspace which they to residential and commercial use to maximize profits. In equilibrium, wages and the price and use of floorspace adjust to clear land and labor markets.

4.2 Workers

A worker ω in group g chooses a location i in which to live, a location j in which to work, and whether or not to own a car denoted by $a \in \{0, 1\}$. Individuals derive utility from consumption of a freely traded numeraire good ($C_{i\omega}$); consumption of residential floorspace ($H_{Ri\omega}$); an amenity reflecting common components of how members of that group enjoy living in i under car ownership a (u_{iag}); and have a disutility from commuting that reduces their productivity at work ($d_{ija} \geq 1$). Workers are heterogeneous in their match-productivity with firms where they work ($\epsilon_{j\omega}$) and their preference for each residence-car ownership pair ($\nu_{ia\omega}$).

Commute costs differ by car ownership because car owners can choose between commuting by car or public transit (such as walking, bus or TransMilenio), whereas workers without cars can only choose between public modes. In addition, cars provide an amenity benefit capturing the ways in which cars are used for leisure activities, but come at a fixed cost of ownership $p_a > 0$.[37]

[34]The choice to keep the total supply of floorspace fixed is motivated by the result that this mostly unaffected by TransMilenio as documented in the appendix. In Section XX I explore the impact of allowing floorspace supply to respond to the system.

[35]This is the "closed city" assumption. In quantitative exercises, I also consider the alternative extreme of perfect mobility between the city and the rest of the country ("open city" assumption).

[36]If the timing is reversed, or if all decisions are made simultaneously, the model no longer admits closed form expressions for commute flows and residential populations. In Monte Carlo exercises I simulate workers' choices where timing is reversed, and find the distribution of residence and employment are qualitatively unchanged. This is intuitive since in the case without non-linearities (when $p_a = \bar{h} = 0$), one can show analytically that the expressions for residence and employment are isomorphic under alternative timing assumptions.

[37]In the appendix, I outline a third stage mode choice problem in which individuals decide how to commute between home and work conditional on their decision on car ownership. Car owners can choose between cars and public modes (walk, bus, TransMilenio) while non-car owners may only use public transportation. The result is that car owners face different average commute times for each trip; thse are what I report in this section.

Assuming that individuals have Stone-Geary preferences in which they need a minimum amount of floorspace $\bar{h}$ in which to live, utility of a worker who has made choice (i, j, a) is

$$U_{ijag\omega} = \max_{C_{i\omega}, H_{Ri\omega}} u_{iag} C_{i\omega}^{\beta} (H_{Ri\omega} - \bar{h})^{1-\beta} \nu_{ia\omega}$$
$$\text{subject to } C_{i\omega} + r_{Ri} H_{Ri} + p_a a = \frac{w_{jg} \epsilon_{j\omega}}{d_{ija}}$$

Solving for the optimal demand for housing and consumption good yields the following expression for indirect utility

$$U_{ijag\omega} = u_{iag} \left(\frac{w_{jg} \epsilon_{j\omega}}{d_{ija}} - p_a a - r_{Ri} \bar{h} \right) r_{Ri}^{\beta-1} \nu_{ia\omega} \tag{1}$$

where the iceberg commute cost $d_{ija} = \exp(\kappa t_{ija})$ increases with the time t_{ija} it takes to commute between i and j under car ownership a. The parameter $\kappa > 0$ controls the size of these commute costs.[38]

In contrast to other work with homothetic preferences (e.g. Ahlfeldt et. al. 2016), the fixed nature of expenditures on cars and housing allows the model to match the Engel curves I document for car ownership and housing expenditure,[39] and drives sorting of workers over car ownership and residential neighborhoods by income. When cars are quicker than public modes of transit, the rich are more willing to pay to the fixed cost since their value of time is higher. Similarly, the fixed expenditure on subsistence housing means that the poor spend a greater share of income on housing and find neighborhoods with cheaper housing more attrractive. Since housing is expensive in high amenity locations in equilibrium, the poor (rich) to sort into low (high) amenity neighborhoods.

Workers first choose where to live and whether or not to own a car, and then choose where to work. I now solve their problem by backward induction.

4.2.1 Employment Decisions

Having chosen where to live i and whether or not to own a car a, individuals draw a vector of match-productivities ϵ_ω with firms in locations across the city.[40] I assume this is drawn from a multivariate Frechet distribution

$$F_g(\epsilon_1, \ldots, \epsilon_J) = \exp\left(- \left[\sum_j \tilde{T}_g \epsilon_j^{-\frac{\tilde{\theta}_g}{1-\rho_g}} \right]^{1-\rho_g} \right).$$

The parameter $\tilde{\theta}_g$ measures the dispersion of productivities for type-g workers (comparative advantage), with a higher $\tilde{\theta}_g$ corresponding to a smaller dispersion, while the parameter ρ_g determines the correlation of

[38]The consensus within the literature is that a semi-log gravity equation best fits the commuting data within cities, which will come from this specification of commute costs d_{ija} (e.g. Fortheringham and O'Kelly (1989)). I assume the commute cost affects productivity at work (i.e. by reducing effective labor supply) rather than overall utility, since this simplifies the gravity equation for commute flows. In Monte Carlo exercises I show that in a simulated city, the effect of improving commuting infrastructure is quantitatively similar if commute costs reduce utility directly.

[39]See the appendix for both figures and explanations of their construction.

[40]In the appendix, I show this can be microfounded by a process of undirected job search where workers and firms meet according to a poisson process with match productivity learned after each meeting.

an individual's talent across locations (absolute advantage). If $\rho_g = 1$ then draws are perfectly correlated within individuals while if $\rho_g = 0$ then they are perfectly uncorrelated. The scalar $\tilde{T}_g$ controls the overall level of productivities for workers in a particular group.

With these draws in hand, linearity of (1) means that workers simply choose to work in the location that offers the highest income net of commute costs $\max_j\{w_{jg}\epsilon_{j\omega}/d_{ija}\}$. Properties of the Frechet distribution imply that the probability a worker of type g who has made choice (i, a) decides to work in j is given by

$$\pi_{j|iag} = \frac{(w_{jg}/d_{ija})^{\theta_g}}{\sum_s (w_{sg}/d_{isa})^{\theta_g}} \equiv \frac{(w_{jg}/d_{ija})^{\theta_g}}{\Phi_{Riag}} \tag{2}$$

where $\theta_g \equiv \tilde{\theta}_g/(1-\rho_g)$ reflects the relative strength comparative advantage for members of skill group g.

Individuals are more likely to commute to a location when it pays a high wage net of commute costs (the numerator) relative to those in all other locations (the denominator). The sensitivity of employment decisions to commute costs is governed by the dispersion of productivity. When workers have similar matches with firms in different locations (high θ_g), then commuting decisions are more sensitive to commute costs. Differences in this productivity heterogeneity across skill groups will be an important factor in the incidence of commute costs, since it determines the extent to which individuals are willing to bear high commute costs to work in a location.

Residential Commuter Market Access Expected income prior to drawing the vector of match productivities is directly related to the denominator in (2) through

$$\bar{y}_{iag} = T_g \Phi_{Riag}^{1/\theta_g}, \tag{3}$$

where T_g is a transformation of the location parameter of the Frechet distribution.[41].

I define the term Φ_{Riag} as *Residential Commuter Market Access* (RCMA). We will see that it summarizes the effect of the entire commuting network on the supply of residents to a location: it rises when a location is close (in terms of commute costs) to well-paid jobs. I will return to the content and measurement of market access in the next section.

4.2.2 Residential Location and Car Ownership Decisions

In the first stage, individuals choose where to live and whether or not to own a car in order to maximize their expected indirect utility

$$U_{iag\omega} = u_{iag}\left(\bar{y}_{iag} - p_a a - r_{Ri}\bar{h}\right) r_{Ri}^{\beta-1} \nu_{ia\omega}.$$

[41] In particular, $T_g \equiv \gamma_{\theta,g}\tilde{T}_g^{1/\theta_g}$ and $\gamma_{\theta,g} = \Gamma\left(1 - \frac{1}{\theta_g(1-\rho_g)}\right)$ where $\Gamma(\cdot)$ is the gamma function.

I assume that the idiosyncratic preferences $\nu_{ia\omega}$ are drawn from a Frechet distribution with shape parameter $\eta_g > 1$. The supply of type-g individuals to location i and car ownership a is then

$$L_{Riag} = \lambda_U \left(u_{iag} \left(\bar{y}_{iag} - p_a a - r_{Ri} \bar{h} \right) r_{Ri}^{\beta - 1} \right)^{\eta_g}. \tag{4}$$

where λ_U is an equilibrium constant.[42] Intuitively, workers are more attracted to locations with high amenities, expected incomes and low house prices, with an elasticity determined by the dispersion of their idiosyncratic preferences η_g. The entire transit network only matters for individuals' residential choices in so far as it affects residential commuter market access, which determines workers expected incomes through (3).[43]

4.2.3 Aggregation

Firm Commuter Market Access and Labor Supply Using the commuting probabilities (2), the supply of workers to any location by summing over the number of residents who commute there $L_{Fjg} = \sum_{i,a} \pi_{j|iag} L_{Riag}$. This implies

$$L_{Fjg} = w_{jg}^{\theta_g} \Phi_{Fjg} \tag{5}$$
$$\text{where } \Phi_{Fjg} = \sum_{i,a} d_{ija}^{-\theta_g} \frac{L_{Riag}}{\Phi_{Riag}}$$

Labor supply in the model takes a log-linear form that depends on two forces. First, more workers commute to destinations paying higher wages. Second, firms attract more workers when they have better access to them through the commuting network, captured through the term Φ_{Fjg}. This is because individuals care about wages net of commute costs.

I define the term Φ_{Fjg} as a location's *Firm Commuter Market Access* (FCMA). It summarizes the effect of the entire commuting network for firms in a location through its effect on labor supply.

The effective units of labor supplied to a location will depend both on the number of individuals who choose to work there as well as their average productivity. Properties of the Frechet distribution imply the average productivity of workers who commute to j from (i, a) is inversely related to the share who choose to do so through $T_g \pi_{j|iag}^{-1/\theta_g}$. This reflects selection: if the net wage per unit of labor in a location is low, few individuals will choose to work there but those that do have exceptionally high idiosyncratic productivities. Total effective labor supply is given by $\tilde{L}_{Fjg} = \bar{\epsilon}_{jg} L_{Fjg}$, where $\bar{\epsilon}_{jg}$ is the average productivity of type-g workers who decide to work in j.[44]

---

[42]In particular, $\lambda_U = \bar{L}_g (\gamma_{\eta,g} / \bar{U}_g)^{\eta_g}$ where $\gamma_{\eta,g} = \Gamma \left(1 - \frac{1}{\eta_g} \right)$ and $\bar{U}_g$ is the overall level of utility for group-g individuals.

[43]Locations will be populated by members of group g only if they are desirable ($\bar{u}_{iag} > 0$) and affordable ($\bar{y}_{iag} - p_a a - r_{Ri} \bar{h} > 0$). Thus, the expression for residential populations in (4) applies only for active locations $\mathcal{A}_{Rg} = \left\{ (i, a) : \bar{u}_{iag} > 0,\ r_{Ri} < (\Phi_{Riag}^{1/\theta} - p_a a) / \bar{h} \right\}$ that are both desirable and affordable for members of group-g, and is zero otherwise. For clarity, I omit this additional notation in the text.

[44]In particular, $\bar{\epsilon}_{jg} = T_g \sum_{i,a} \frac{\pi_{j|iag}^{-1/\theta_g}}{d_{ija}} \frac{\pi_{j|iag} L_{Riag}}{\sum_{r,o} \pi_{j|rog} L_{Rrog}}$.

Worker Welfare Finally, I equate the overall welfare of group-g residents with the expected utility prior to drawing their idiosyncratic preferences in the first stage. This is given by

$$\bar{U}_g = \gamma_{\eta,g} \left[\sum_{i,a} \left(u_{iag} \left(\bar{y}_{iag} - p_a a - r_{Ri} \bar{h} \right) r_{Ri}^{\beta - 1} \right)^{\eta_g} \right]^{1/\eta_g} \tag{6}$$

where $\gamma_{\eta,g} = \Gamma\left(1 - \frac{1}{\eta_g}\right)$ is a constant.

4.3 Firms

Technology There are $s \in \{1, \ldots, S\}$ industries which produce varieties differentiated by location in the city under perfect competition. Their output is freely traded, and consumers have CES preferences over each variety with elasticity of substitution $\sigma_D > 1$.[45] Firms produce using a Cobb-Douglas technology over labor and commercial floorspace

$$Y_{js} = A_{js} N_{js}^{\tilde{\alpha}_s} H_{Fjs}^{1-\tilde{\alpha}_s}$$
$$\text{where } N_{js} = \left(\sum_g \alpha_{sg} \tilde{L}_{Fjgs}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

where the labor input is a CES aggregate over the effective labor across skill groups with elasticity of substitution σ, $\tilde{\alpha}_s = \sum_g \alpha_{sg}$ is the total labor share and A_{js} is the productivity of location j for firms in industry s which they take as given. Industries differ in the intensity in which they use different types of workers α_{sg}. Some, such as real estate and financial services, require all else equal a higher share of high-skill workers while others, such as hotels and restaurants, rely on the low-skilled.

Factor Demand Perfect competition implies that the price of each variety is equal to its marginal cost $p_{js} = W_{js}^{\tilde{\alpha}_s} r_{Fj}^{1-\tilde{\alpha}_s} / A_{js}$, where r_{Fj} is the price of commercial floorspace in j and

$$W_{js} = \left(\sum_g \alpha_{sg}^{\sigma} w_{jg}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

is the cost of labor for firms of industry s in location j. Intuitively, labor costs differ by industries due to their differential skill requirements.

[45] The CES aggregate is the numeraire good introduced in the consumer's problem. While the assumption of representative firms with a fixed location seems restrictive, in the next section I show this production technology is isomorphic to more realistic setups.

Solving the firm's cost minimization problem, the demand for labor and commercial floorspace is

$$\tilde{L}_{Fjgs} = \left(\frac{w_{jg}}{\alpha_{sg} W_{js}}\right)^{-\sigma} N_{js} \tag{7}$$

$$H_{Fjs} = (1 - \tilde{\alpha}_s) \frac{X_{js}}{r_{Fj}} \tag{8}$$

where X_{js} is firm sales.[46]

4.4 Floorspace

Market Clearing There is a fixed amount of floorspace H_i in any location, a fraction ϑ_i of which is allocated to residential use and $1 - \vartheta_i$ to commercial use. For any allocation, market clearing for residential floorspace requires that the supply of residential floorspace $H_{Ri} = \vartheta_i H_i$ equals demand:

$$r_{Ri} = (1 - \beta) \frac{E_i}{H_{Ri} - \beta \bar{h} L_{Ri}} \tag{9}$$

where $L_{Ri} = \sum_{g,a} L_{Riag}$ is the total number of residents in i.

Likewise, the supply of commercial floorspace $H_{Fj} = (1 - \vartheta_i) H_j$ must equal that demanded by firms:

$$r_{Fj} = \frac{\sum_s (1 - \tilde{\alpha}_s) \left(W_{js}^{\tilde{\alpha}_s} r_{Fj}^{1-\tilde{\alpha}_s} / A_{js}\right)^{1-\varsigma} X}{H_{Fj}}. \tag{10}$$

Floorspace Use Allocation Landowners allocate the fraction ϑ_i of floorspace allocated to residential use to maximize profits. They receive r_{Ri} per unit of floorspace allocated to residential use, but land use regulations limit the return to each unit allocated to commercial use to $(1 - \tau_i) r_{Fi}$. Landowners allocate floorspace to its most profitable use so that

$$\begin{aligned} \vartheta_i = 1 \ \text{ if } r_{Ri} > (1 - \tau_i) r_{Fi} \\ (1 - \tau_i) r_{Fi} = r_{Ri} \ \ \forall \{i : \vartheta_i \in (0, 1)\} \\ \vartheta_i = 0 \ \text{ if } (1 - \tau_i) r_{Fi} > r_{Ri} \end{aligned} \tag{11}$$

4.5 Externalities

Productivities A long literature points to the importance of productivity spillovers in cities.[47] I allow a location's productivity to depend on an exogenous component that reflects features independent of the

[46] From CES demand sales are given by $p_{js}^{1-\sigma_D} X$ where $X = \sum_i \beta (E_i - \bar{h} r_{Ri} L_{Ri})$ is total spending on goods in the city and $E_i = \sum_{g,a} \bar{y}_{iag} - p_a a$ is total spending on goods and housing from residents in i.

[47] This idea dates back at least to Adam Smith (1776), and was articulated more fully in Marshal (1890). Two prominent examples establishing this relationship are Ciccone and Hall (1996) using regional data and Ahlfeldt et. al. (2016) using intra-city data. See Rosenthal and Strange (2004) for a review.

density of economic activity (e.g. access to roads, electricity, crime) as well as a production externality that depends on the density of employment in that location

$$A_{js} = \bar{A}_{js} \left(\frac{\tilde{L}_{Fj}}{T_j} \right)^{\mu_A},$$

where $\tilde{L}_{Fj} = \sum_s \tilde{L}_{Fjs}$ is the total effective labor supplied to that location and T_j is the total units of land. The strength of agglomeration externalities is governed by the parameter μ_A.[48]

Amenities Similarly, I allow amenities in a neighborhood to depend on an exogenous component which also varies by car ownership (e.g. leafy streets, close to getaways surrounding the city) and a residential externality that depends on its college share of residents

$$u_{iag} = \bar{u}_{iag} \left(\frac{L_{RiH}}{L_{Ri}} \right)^{\mu_{U,g}}.$$

In contrast to existing urban models (e.g. Ahlfeldt et. al. 2015), endogenous amenities depend on the composition of residents across skill groups rather than the total density of residents. This seems especially applicable to poorer cities that lack strong public goods provision; in Bogotá, where crime is a significant problem, the rich often pay for private security around their buildings which increase the sense of safety in those areas.[49] This externality provides an additional force towards residential segregation, since the high-skilled are more willing to pay to live in high amenity neighborhoods and by doing so increase the amenities even more. While this sorting force could be driven by the subsistence housing requirement alone, I allow the strength of residential externalities $\mu_{U,g}$ to potentially differ across groups so that some groups may prefer to live near the high-skilled all else equal. I let the data speak to the relative strength of these forces towards residential segregation in estimation.

4.6 Equilibrium

I now define general equilibrium in the city.

Definition. Given vectors of exogenous location characteristics $\{H_i, \bar{u}_{iag}, \bar{A}_{js}, t_{ija}, \tau_i\}$, city group-wise populations $\{\bar{L}_g\}$ and model parameters $\{\bar{h}, \beta, \alpha, p_a, \kappa, \theta_g, \rho_g, T_g, \eta_g, \alpha_{sg}, \sigma_D, \sigma, \mu_A, \mu_U\}$, an equilibrium is defined as a vector of endogenous objects $\{L_{Riag}, L_{Fjg}, w_{jg}, r_{Ri}, r_{Fi}, \vartheta_i, \bar{U}_g\}$ such that

1. **Labor Market Clearing** The supply of labor by individuals (5) is consistent with demand for labor by firms (7),

[48]Unlike Ahlfeldt et. al. (2016) I do not allow for spillovers across locations. The authors find very local spillovers across space which go to zero within 15 minutes of walk time. This is echoed by Rossi-Hansberg et. al. (2010) who find spillovers from revitalized houses fall approximately one half every 1,000 feet. Given that I estimate my model at the census tract, including spillovers across spatial units led to such a large estimate for spatial decay that essentially shut down this channel across spatial units in my setting. I show later that the regression approach can be extended to include such spillovers, though.

[49]Evidence for the US discussed in section 2 suggests amenities in cities do depend on the composition of residents.

2. **Floorspace Market Clearing** The market for residential floorspace clears (9) and its price is consistent with residential populations (4), the market for commercial floorspace clears (10) and floorspace shares are consistent with land owner optimality (11)

3. **Closed City** Populations add up to the city total, i.e. $\bar{L}_g = \sum_{i,a} L_{Riag} \ \forall g$.

With this definition in hand, I now characterize existence and uniqueness of equilibria in this economy.

Proposition 1. *An equilibrium exists in this city. Moreover, in a special case of the model with one group of workers, firms and commute modes and no non-homotheticities ($\bar{h} = p_a = 0$) and a fixed allocation of floorspace, a sufficient condition for the equilibrium to be unique is that*

$$\mu_A \leq 1 - \alpha + \frac{\sigma + \theta - 1}{(\sigma - 1)(\theta - 1)}$$

$$\mu_U \leq \frac{1 + \eta(1 - \beta)}{\eta} - \frac{\beta}{\theta - 1}$$

$$\beta(\sigma - 1)\mu_A + \sigma\mu_U \leq \frac{\sigma}{\eta} + \sigma(1 - \beta) + \beta(1 + (\sigma - 1)(1 - \alpha))$$

Relative to existing papers (Ahlfeldt et. al. 2015, Allen et. al. 2016), the model adds multiple groups of workers, industries and transit modes, along with two fixed expenditures that potentially influence the set of locations inhabited by each group. Despite these additional features, the first part of the proposition ensures an equilibrium still exists in the city.[50] The second part of the proposition shows that in a special case of the model, the equilibrium is unique only if spillovers are sufficiently weak. In the presence of strong spillovers, fundamental productivities and amenities becomes less important and different urban configurations can be supported as equilibria. While multiplicity does not pose a problem for my estimation strategy (which only requires that two equilibria be observed), I address it through my selection mechanism when solving for counterfactual equilibria.

4.7 Intuition for Welfare Effects

To build intuition for the channels through which changes in the transit network affect welfare, I totally differentiate the expression for average utility (6). The change in utility in response to a small change in commute costs is given by

$$d \ln \bar{U}_g = \sum_{i,a} \lambda_{iag} \left(\underbrace{-\mu^Y_{iag} \sum_j \pi_{j|iag} d \ln d_{ija}}_{\text{Partial Equilibrium}} + \underbrace{d \ln u_{iag} + \mu^Y_{iag} \sum_j \pi_{j|iag} d \ln w_{jg} - \left[(1 - \beta) + \mu^R_{iag}\right] d \ln r_{Ri}}_{\text{General Equilibrium}} \right)$$

[50] Intuitively, land owner optimality ensures that enough residential floorspace is provided so that each group of workers can afford to live in the city. Were this not the case, then if low-skill workers could afford to live in the city, the CES production technology means that commercial floorspace prices and all wages would equal zero and the city would in effect "shut down", providing zero payments to landowners. In the proof, I also consider the open city assumption.

where $\lambda_{iag} = L_{Riag}/\bar{L}_g$ is the share of type-g individuals living in i under car ownership a, $\pi_{j|iag}$ is the conditional commuting probability, $\mu^Y_{iag} = \frac{\bar{y}_{iag}}{\bar{y}_{iag}-p_a a-r_{Ri}\bar{h}}$ is the ratio of gross labor income to labor income net of fixed expenditures, and $\mu^R_{iag} = \frac{r_{Ri}\bar{h}}{\bar{y}_{iag}-p_a a-r_{Ri}\bar{h}}$ is the ratio of the subsistence housing expenditure to net income.

There are both partial and general equilibrium effects of reductions in commute costs on worker welfare. The partial equilibrium effect is greater if costs are reduced between locations where many people live and work (reflected through λ_{iag} and $\pi_{j|iag}$ respectively). General equilibrium effects depend on the response of wages and amenities (which raise welfare) and residential floorspace prices (which reduce welfare). Standard approaches to measure the benefits of improvements in commuting infrastructure measure only the partial effects through the value of time savings (e.g. Small and Verhoef 2005). I assess the importance of accounting for general equilibrium forces in quantitative exercises.

The effects of improvements in public transit differ across skill groups in a way that is ex ante ambiguous. First, within any residential location the low-skilled rely more on public transport and so put more weight on the improvement. Second, worker groups can differ in the elasticity of commute flows to commute costs (controlled by θ_g). The group with less sensitive commute decisions are more willing to tolerate high commute costs in the initial equilibrium, and thus put more weight on costly commutes (reflected through $\pi_{j|iag}$). If the percentage drop in commute costs as a result of the improvement is greater for longer commutes, then welfare gains will be greater for this group. Third, low-skill workers are hurt more by house price apperciation since spend a greater share of their income on housing (μ^R_{iag} is greater) but also experience a greater percentage change in wages given the fixed expenditures on housing and cars (μ^Y_{iag} is greater). Fourth, low- and high-skill individuals live and work in different locations (reflected both through λ_{iag} and $\pi_{j|iag}$). The group for whom the improvement reduces commute costs the most between locations of residence and employment will benefit the most.

The direction of these forces depends on parameter estimates as well as the geography of the city and its transit improvements. While this decomposition helps develop intuition, the estimated model allows me to quantify the net effect on worker welfare.

5 Using The Model To Guide Empirical Work

In this section, I show that in a special case the model's equilibrium has a reduced form representation in which outcomes such as population, employment and floorspace prices can be written as log-linear functions of commuter market access. This provides a regression framework I use to empirically evaluate the effect of TransMilenio. In fact, I show this framework applies to a wide class of models featuring a gravity equation for commute flows and is robust to a number of alternative modelling assumptions. While this simplified framework is unable to assess the distributional effects of the system, I use the intuition developed in this section to guide structural estimation of the full model later in the paper.

5.1 A Special Case of the Model

Commuter Market Access: Measurement and Intuition Consider a simplification of the model with one group of workers, firms and transit modes and a fixed allocation of residential and commercial floorspace. Using the labor supply curve (5) to substitute for wages in (5), I find that given data on population, employment and commute costs as well as a value for the commuting elasticity, commuter market access is the solution to the following system of equations

$$\Phi_{Ri} = \sum_j d_{ij}^{-\theta} \frac{L_{Fj}}{\Phi_{Fj}}$$
$$\Phi_{Fj} = \sum_i d_{ij}^{-\theta} \frac{L_{Ri}}{\Phi_{Ri}}$$

Residential commuter market access is greater when a location is close (in terms of commute costs) to other locations with high employment, particularly so when these other locations have low access to workers. Firm commuter market access is greater when a location is close to other locations with high residence, particularly so when these other locations have low access to jobs.[51] I show below that the solution to this system of equations exists and is unique, so that the market access measures are easily computed using data on population, employment and commute costs as well as a value for the commuting elasticity.

Regression Framework In this simplified model, the equilibrium reduces to the following system

$$L_{Ri} = \lambda_U \left(u_i \Phi_{Ri}^{1/\theta} r_{Ri}^{\beta-1} \right)^{\eta}$$
$$r_{Ri} = \frac{1-\beta}{H_{Ri}} \Phi_{Ri}^{1/\theta} L_{Ri}$$
$$\tilde{L}_{Fi} = w_j^{\theta-1} \tilde{\Phi}_{Fj}$$
$$\tilde{L}_{Fi} = \frac{1}{\alpha} w_i^{\alpha(1-\sigma)-1} A_i^{\sigma-1} r_{Fi}^{(1-\sigma)(1-\alpha)} P^{\sigma-1} E$$
$$r_{Fi} = \left(\frac{A_i^{\sigma-1} w_i^{-\alpha(\sigma-1)} P^{\sigma-1} E}{(1-\alpha) H_{Fi}} \right)^{\frac{1}{1+(\sigma-1)(1-\alpha)}}$$

where $\tilde{\Phi}_{Fi} = \sum_i d_{ij}^{-\theta} \frac{L_{Ri}}{\Phi_{Ri}} \Phi_{Ri}^{\frac{1}{\theta}}$ is an adjusted firm commuter access term that accounts for access to effective units of labor supplied, , $u_i = \bar{u}_i L_{Ri}^{\mu_U}$ and $A_i = \bar{A}_i \tilde{L}_{Fi}$.[52]

The first equation determines the supply of residents to any location given residential floorspace prices.

[51] These expressions are closely related to commute-distance weighted sums of employment and residence respectively, reminiscent of the commuting analogue of "market potential" discussed by Harris (1954) and alluded to in the discussion of accessibility in Hansen (1959). I include these alternative measures in the empirical results, to show the results are robust to measuring commuter market access with these model-independent measures.

[52] For simplicity, I model depence of productivity on total employment rather than employment density. It doesn't affect the results since the units of land in the denominator are absorbed into the structural error.

The second equation is a market clearing condition for residential floorspace, which provides an inverse demand equation for residents. Together, these supply and demand curves determine equilibrium in the market for residents. The third and fourth equations are labor supply and demand schedules that determine equilibrium in the labor market. The fifth equation is a market clearing condition that determines equilibrium in the market for commercial floorspace.[53]

Taking logs, stacking the equations and considering long-differences between two time periods, the change in endogenous variables can be written as the following system

$$A\Delta \ln Y_i = B_R \Delta \ln \Phi_{Ri} + B_F \Delta \ln \tilde{\Phi}_{Fi} + e_i$$

where $\Delta \ln Y = \begin{bmatrix} \Delta \ln L_{Ri} & \Delta \ln r_{Ri} & \Delta \ln r_{Fi} & \Delta \ln \tilde{L}_{Fi} \end{bmatrix}'$ is a vector of log changes in endogenous variables, A is a matrix of coefficients reflecting the interdepence between endogenous variables, B_R and B_F are vectors of coefficients controlling the direct effects of changes in market access on outcomes, and e is a vector of structural residuals containing changes in fundamentals $\bar{u}_i, \bar{A}_i, H_{Ri}, H_{Fi}$.[54] Premultiplying by the coefficient matrix A yields the reduced form

$$\Delta \ln Y_i = A^{-1} B_R \Delta \ln \Phi_{Ri} + A^{-1} B_F \Delta \ln \tilde{\Phi}_{Fi} + A^{-1} e_i \tag{12}$$

where the reduced form coefficients are given by

$$A^{-1}B_R = \begin{bmatrix} \frac{\beta\eta}{\theta(1+\eta(1-\beta-\mu_U))} \\ \frac{1+\eta(1-\mu_U)}{\theta(1-\eta(1-\beta-\mu_U))} \\ 0 \\ 0 \end{bmatrix}, \; A^{-1}B_F = \begin{bmatrix} 0 \\ 0 \\ \frac{\alpha+(\alpha(\sigma-1)+1)\mu_A}{\theta\sigma/(\sigma-1)+(\theta-1)((1-\alpha)(\sigma-1)\mu_A-\alpha)} \\ \frac{\sigma/(\sigma-1)}{\theta\sigma/(\sigma-1)+(\theta-1)((1-\alpha)(\sigma-1)\mu_A-\alpha)} \end{bmatrix}$$

The regression specification (12) reflects the total change of outcomes in response to changes in market access. This reflects both the direct effect (in the B_R and B_F coefficient vectors) and the indirect effect (in A^{-1}) as the response to improved commuter market access filters through labor and land markets. The block structure of reduced form coefficients mean that residential (commercial) outcomes depend only on changes in residential (commercial) commuter market access, so that the specification for each outcome has a simple univariate specification.

5.2 Isomorphisms in Gravity Commuting Models

The previous section shows that the model is well-suited to guide empirical analysis since it delivers a regression framework connecting changes in outcomes with changes in commuter market access. I now show that commuter market access can be easily measured using data on population and employment, and

[53]This can be substituted into the expression for labor demand to eliminate commercial floorspace prices from the system. I retain this since I explore the response of these prices to firm CMA in the empirics. The gravity generalization reduces this and isomporphic models into equilibrium systems in only population, employment, and market access.

[54]See the appendix for a full definition of each term and derivation of the reduced form.

a similar reduced form representation exists for a wide class of models so that the regression framework is robust to a number of alternative modelling assumptions.

Proposition 2. *Consider a model where commute flows are of the "gravity" form*

$$L_{ij} = \gamma_i \delta_j \kappa_{ij} \tag{13}$$

where $\gamma_i, \delta_j > 0$ are endogenous and $\kappa_{ij} > 0$ is exogenous. Then

(i) **Measuring CMA** *The supply of residents and workers to locations are given by $L_{Ri} = \gamma_i \Phi_{Ri}$ and $L_{Fi} = \delta_i \Phi_{Fi}$. Given data $\{L_{Ri}, L_{Fi}\}$ and parameters $\{\kappa_{ij}\}$, the commuter market access terms Φ_{Ri}, Φ_{Fi} are the unique solution to the system*

$$\Phi_{Ri} = \sum_j \frac{L_{Fj}}{\Phi_{Fj}} \kappa_{ij} \tag{14}$$

$$\Phi_{Fi} = \sum_j \frac{L_{Rj}}{\Phi_{Rj}} \kappa_{ji} \tag{15}$$

(ii) **General Gravity Model** *When there is log-linear demand for labor and residents of the form $\tilde{L}_{Fj} = A_j \delta_j^{\alpha}$ and $L_{Ri} = B_i \gamma_i^{\beta} \Phi_{Ri}^{\gamma}$ where $A_i, B_i > 0$ are exogenous constants and the supply of labor (potentially different from the number of workers) is given by $\tilde{L}_{Fj} = \delta_j^{\delta} \tilde{\Phi}_{Fj}$, where $\tilde{\Phi}_{Fj} = \sum_i \frac{L_{Ri}}{\Phi_{Ri}} \kappa_{ij} \Phi_{Ri}^{\epsilon}$ then*

1. *An equilibrium always exists and is unique whenever $|\epsilon(\beta - 1) - \gamma| \leq |\beta - 1||\alpha - 1|$*
2. *The economy has a reduced form representation where residence and employment can be written as*

$$\Delta \ln Y_i = B \Delta \ln \Phi_i + e_i$$

where $\Delta \ln Y_i = \begin{bmatrix} \Delta \ln L_{Ri} \\ \Delta \ln \tilde{L}_{Fi} \end{bmatrix}, \Delta \ln \Phi_i = \begin{bmatrix} \Delta \ln \Phi_{Ri} \\ \Delta \ln \tilde{\Phi}_{Fi} \end{bmatrix}, B = \begin{bmatrix} \frac{\beta - \gamma}{\beta - 1} & 0 \\ 0 & \frac{\alpha}{\alpha - \delta} \end{bmatrix}$ and e_i is a structural error term containing changes in the exogenous constants.

The first part of the proposition shows that unique values of market access can be computed using data on the location of residence and employment, as well as a parameterization of commute costs. This uniqueness property is independent of whether there are multiple model equilibria, since these measures can always be inverted from an observed equilibrium. These measures are solved using only the supply side structure of these models, from the gravity equation for commute flows.

The second part of the proposition shows that when additional structure on the demand for residents and workers across the city, population and employment can be written as log-linear functions of commuter market access. In addition to the log-linear functional forms, this structure requires knowing values for the parameters $\alpha, \beta, \gamma, \delta$ and ϵ.[55] By observation, the result in part (i) implies that knowledge of these

[55] Most models impose additional restrictions between these 5 parameters, which reduces the number of parameters one needs to know (see the appendix for examples).

parameters as well as data on residence, employment and commute costs allows one to solve for the endogenous objects $\{\delta_i, \gamma_i\}$ and location characteristics $\{A_i, B_i\}$ that rationalize the data as an equilibrium of the model. While supply is upward sloping in the shifters $\{\delta_i, \gamma_i\}$, multiple equilibria may occur when demand is upward sloping (determined by the constants α, β). This will be the case in the presence of strong spillovers as seen above.

The gravity equation for commuting that determines the supply side of the model enjoys wide empirical support and is used in vast majority of recent quantitative urban models.[56] While the demand side of the model is expressed solely in terms of employment and residence, this representation can accommodate additional factors (such as residential and commercial housing) so long as the supply and demand equations are log-linear. In the appendix, I show this framework accommodates iso-elastic housing supply, alternative production technologies (e.g. Eaton and Kortum 2002 and individual entrepeneurs who sort over where to produce across the city) and worker preferences (such as utility over leisure).[57] Thus, the regression framework I take to the data is robust to a host of alternative modelling assumptions.

5.3 Relation to the Market Access Literature in Economic Geography

While the commuter market access approach shares similarities with parts of the recent economic geography literature, some key differences distinguish my framework from a simple relabelling of goods as commuters in those models.[58]

First, these models impose additional assumptions of balanced trade and symmetric trade costs to construct their empirical measures. By imposing neither, my approach requires less structure to recover measures of market access from the data. Indeed, one can show that it is precisely the absence of these two conditions that deliver seperate notions of firm and residents commuter market access in this paper.[59] This distinction is important given that changes in firm and residential commuter market access capture very different sources of variation (see Figure 1).

Second, commuter access measures can be computed using data on quantities (of population and employment) alone, compared to the measures in economic geography which require data on both population and income. Comprehensive data on income at small spatial disaggregation within cities is typically pro-

[56]See McDonald and McMillen (2010) for a review of the evidence in support of gravity in commute flows. All the quantitative models in section 2 feature gravity equations for commute flows. Finally, note that my full model only exhibit gravity in commute flows conditional on location of residence. I show in the next section that there exist unique measures of commuter market access in my model given the observed data. In fact, part (i) of proposition 2 can be easily extended to include models with gravity in commute flows conditional on location of residence.

[57]I also show the framework can be used when workers have preference rather than productivity shocks over employment locations, so that there is no difference between effective labor and the number of workers. In this case, a small approximation is used to write the equilibrium in the exact log-linear form.

[58]See Redding and Sturm (2008), Bartelme (2015) and Donaldson and Hornbeck (2015) for examples. My gravity formulation builds on Allen et. al. (2017) were the first to provide general insight into gravity models of trade, and Bartelme (2015) who nested a market approach within this gravity trade framework.

[59]Imposed balanced trade in commuters would require the number of workers in a location to equal the number of residents, which is clearly counterfactual. Since my procedure to calculate commute times results in symmetric commute times, assymetries in the location of residents and employment are the sole driver of the assymmetries in changes in residential and firm commuter market access in Figure 1.

hibitively difficult to obtain, especially in developing countries, whereas population and employment are widely available.

6 Data

In this section I provide an overview of the primary datasets used in the analysis. Additional details are provided in the data appendix.

The primary geographic unit used in the analysis is the census tract ("sección"). Bogotá is partitioned into 2,799 tracts, with an average size of 133,303 square meters and a mean population of 2,429 in 2005.[60] These are contained within larger spatial units including 19 localities and 113 planning zones (UPZs).

My primary source of population data is the Department of Statistics' (DANE) General Census of 1993 and 2005. This provides the residential population of each block by education-level. I define college-educated individuals as those with some post-secondary education defined by their last complete year of study. In 2015, DANE provides population totals at the UPZ. I combine this with the share of college-educated workers in each UPZ in the GEIH survey (described below) to construct population by skill group. This allows me to compute seperate growth rates of college and non-college residents between 2005 and 2015 within each UPZ. I then calculate 2015 census tract population by skill group by inflating the 2005 totals by these growth rates. Details are provided in the appendix.

I use two sources of data on employment. The first is a census covering the universe of establishments from DANE's 2005 General Census and 1990 Economic Census which report the location, industry and employment of each unit. The second is a database of establishments registered with the city's Chamber of Commerce (CCB) in 2000 and 2015. In 2015 this contains the location, industry and employment of each establishment, but in 2000 establishment size is not provided. While I tend to use the census and CCB datasets seperately, a concern is that the spatial distribution of registered employment may be different from that of total employment. In the appendix, I show that the employment and establishment densities in both years of the CCB data are highly correlated with that from the 2005 census. Importantly, coverage is even across different types of neighborhoods, suggesting both that the CCB data is representative of overall employment. Since I rely on establishment counts to measure employment in 2000, I also show that establishment count and employment densities are highly correlated.

Housing market data between 2000 and 2013 comes from Bogotá's Cadastre. Its mission is to keep the city's geographical information up to date; all parcels, formal or informal, are included with the result that the dataset covers 98.6% of the city's more than 2 million properties (Ruiz and Vallejo (2015)).[61] It reports the use, floorspace and land area, value per square meter of land and floorspace, as well as a number of property characteristics. Values in the cadastre are important for the government since they determine property taxes which comprise a substantial portion of city revenue.[62] In developed countries,

[60]Number reported is for tracts with positive population. Almost all tracts (2,768) have positive population in 2005. For comparison, tracts in Bogotá are about 60% smaller than those in New York City which had an average of 4,067 residents in the 2010 census.

[61]I confirmed this high coverage by overlaying the shapefile for available properties over satellite images of the city.

[62]For example, in 2008 property taxes accounted for 19.8% of Bogotá's tax revenues (Uribe Sanchez (2015)).

these valuations are typically determined using information on market transactions. However, Bogotá, like most developing cities, lacks comprehensive records of such data and those available may be subject to systematic under-reporting. As described in the appendix, the city circumvents this through an innovative approach involving sending officials to pose as potential buyers in order to negotiate to a sales price under the premise of a cash payment (Anselin and Lozano-Gracia (2012)). Professional assessors are also sent to value at least one property in one of each of the city's more than 16,000 "homogenous zones" (Ruiz and Vallejo (2015)).[63] As a result, I show the average price per square meter of floorspace in the cadastre is highly correlated with the average price per room of purchases reported in a DANE worker survey. Importantly, the relationship is constant across rich and poor neighborhoods which would not be the case were the cadastre over- or under-valuing expensive properties.

Four waves of microdata on commuting behavior come from the city's Mobility Survey administered by the Department of Mobility and overseen by DANE. Conducted in 1995, 2005, 2011 and 2015, these are representative household surveys in which each member was asked to complete a travel diary for the previous day. The survey reports the demographic information of each traveller and household, including age, education, gender, industry of occupation, car ownership and in some years income. For each trip, the data report the departure time, arrival time, purpose of the trip, mode, as well as origin and destination UPZ.

Employment data at the worker level come from DANE's Continuing Household Survey (ECH) between 2000 and 2005, and its extension into the Integrated Household Survey (GEIH) for the 2008-2015. These are monthly, repeated cross-sectional labor market surveys covering approximately 10,000 households in Bogotá each year. They report individual and household characteristics, as well details on employment such as income, hours worked and industry of occupation across primary and secondary jobs. I was able to access versions of these datasets with the block of each household reported.

Commute times between more than 7.8 million pairs of census tracts by each mode are computed in ArcMap. I obtain the shape of each mode's network by combining spatial datasets provided by the city.[64] To construct the time to traverse each edge of the network, I assign speeds in order to match both reported values in the literature as well as the distribution of commute times observed in the Mobility Surveys.[65]

Finally, I measure the distance of tracts to various spatial features provided by the city. I also use a land use map of the city in 1980 provided by the US Defence Mapping Agency and a Tramway map from Morrison (2007).

[63] Surveyors are sent out to update the characteristics of each property every couple of years. Since the primary data informative about prices is not necessarily updated each year, I focus on long-differences in my analysis.

[64] For example, the TransMilenio network is the union of pedestrian paths, trunk lines and feeder routes; the latter two can only be entered at stations. As described in detail in the appendix, one mode may have different speeds depending on the part of the network. For example, cars have different speeds on primary, secondary and tertiary roads.

[65] While I provide evidence speeds were not changing on routes affected by TransMilenio relative to other locations, the appendix shows that aggregate speeds fell between 1995 and 2005 (a period of city expansion) but remained relatively constant thereafter. I assign speeds to seperately match the commute data for each period, and use the average computed time in the main analysis. In robustness exercises, I run specifications with alternate comute times to ensure my results are not sensitive to this choice.

7 Empirical Analysis

In this section, I use the log-linear relationships between endogenous outcomes and commuter market access derived in Section 5 to empirically assess the effect of TransMilenio on land and labor markets.

7.1 Approach and Identification

Taking logs of the expression for residential outcomes in the first two entries of the reduced form system (12) and delivers my baseline specification

$$\Delta \ln Y_{Rit} = \beta \Delta \ln \Phi_{Rit} + \alpha_\ell + \gamma' X_{it} + \epsilon_{Rit} \tag{16}$$

That is, I regress changes in (log) residential outcome Y_{Rit} in census tract i in year t on changes in (log) residential commuter market access Φ_{Rit}, as well as a set of controls that contain census tract characteristics X_{it} as well as locality fixed effects α_ℓ. An equivalent specification holds for commercial outcomes, which instead depend on firm commuter market access. These commuter market access terms are defined implicitly by the system of equations in population, employment and commute costs $d_{ij}^{-\theta}$. The elasticity β is identified from variation in commuter market access within census tracts over time, comparing tracts within a locality with similar observable chacteristics which experienced different changes in market access.[66] Typically this regression will be estimated in long-differences over a pre- and post-period.

Figure 1 plots the distribution of changes in commuter access across the city induced by the construction of the first two phases of the system.[67] The system increases access to jobs much more for tracts in the outskirts of the city, which were far from the high employment densities towards the center. Firms' access to workers rose more in the center of city, since these locations were best to positioned to take advantage of increased labor supply along all spokes of the network.

Challenges to Identification There are 2 key challenges to estimating the specification (16).

First, changes in market access contain population and employment in both periods. Local productivity and amenity shocks (contained in the error term) that drive movements in residence and employment will therefore be mechanically correlated with the error. I address this by instrumenting for changes in commuter market access holding population and employment fixed at their initial values.[68] This isolates the portion of this variation due to the change in commute costs.

[66] Of course, the reduced form elasticities are outcome-specific but I omit this additional notation here. Note also that for firm outcomes I use the unadjusted firm commuter access instead of the adjusted term that reflects units of effective labor supplied for clarity. The results are qualitatively unchanged if I use the adjusted term (the measures have a XX correlation).

[67] In order to compute the market access terms, I require values for $d_{ij}^{-\theta} = \exp(-\theta\kappa t_{ij})$. The estimation of θ_g, κ is outlined in the next section; I measure θ and t_{ij} by averaging over skill group and car ownership values respectively weighting by population shares of each category. The figure plots the change in CMA induced by holding population and employment fixed at their initial level in 1993 and 1990 respectively (from the population and economic census) and changing only commute costs. This isolates graphically the portion of the change due only to TransMilenio.

[68] I also exclude the location itself from the construction of this instrument to remove the most obvious form of mechanical bias.

Second, the variation in commuter market access due to the change in commute costs may be correlated with the error if the government aimed to support neighborhoods that were growing or to stimulate those that were lagging.[69] I therefore construct two instruments for TransMilenio's placement (full details are in the appendix).

The first instrument takes as given the government's overall strategy of connecting portals in the edge of the city with the CBD as given, excludes those areas from the analysis, and constructs the routes that would have been built if the sole aim had been to minimize costs. I construct these routes by first digitizing a land use map of Bogotá in 1980 to measure the different type of land use on small pixels across the city (e.g. arterial roads, vacant, slope etc). Using engineering estimates for the cost to built BRT on different types of land use, this provides a construction cost raster for the city based on the share of land use in each pixel. This allows me to solve for the least cost paths connecting portals with the CBD in ArcMap. I then instrument for changes in market access by supposing TransMilenio had been located along these least cost routes. This will be a valid instrument under the reasonable assumption that these routes should be uncorrelated with trends in amenities and productivities(conditional on controls).[70]

The second instrument exploits the location of a tram system opened in 1884, had its last extension in 1921 and stopped operating in 1951. The tram was built along wide arterial roads in the city, which should predict the location of TransMilenio since these are cheaper to convert to BRT than narrow ones. Moreover routes should be uncorrelated with changes in productivites and amenities between 2000 and 2012 to the extent that these were unanticipated by city planners in 1921.[71]

My identification assumption is that the instruments have only an indirect effect on outcome growth through market access. An attractive feature of my approach is that I can control for distance to the features (distance to the tram, distance to main roads) and use only residual variation in the instrument. This directly controls for direct effects of the instruments. I include additional historical variables to control for other direct effects of the historical instrument.[72] To provide further evidence in support of my identification assumption, I check the stability of IV point estimates as controls are added and test that both instruments yield similar coefficients. I also run falsification tests in which I test for an absence of effects from changes in market access due to lines before they open, as well as a number of robustness checks, to provide additional evidence of the causal effect of improved commuter market access on urban outcomes.

[69] The discussion in section 3 suggests the former is more likely.

[70] Such construction cost based instruments have been used in regional settings to predict the location of highways been cities (Faber (2015), Alder (2017)).

[71] Much of the trade and economic geography literature have used similar types of historical instruments to predict the location of present day infrastructure (e.g. Baum Snow (2007), Duranton and Turner (2011)).

[72] These incude 1933 population and distance to main road in 1933. To further reduce concerns any direct effect of the tram on outcomes, I extend the tram lines to the edge of the city (which also greatly improves its predictive power over TransMilenio placement).

7.2 Results: Main Outcomes

In this section I present results from the linear specification (16).[73]

Main Outcomes Table 3 presents the main results. In all specifications, only tracts further than 500m from a portal and the CBD are included in order to keep a constant sample across specifications. Columns (1) and (2) report the results from the OLS regressions where the change commuter market access is measured using both the change in commute costs as well as the change in population and employment. In most cases, the point estimates are slightly lower in column (2) (my preferred specification) due to the positive correlation of changes in market access with initial land market and demographic characteristics that caused treated locations to grow faster over the period.

Columns (3) and (4) run the baseline IV specification, which instrument for the total change in market access holding employment and population fixed at their initial levels. The point estimates tend to fall slightly, reflecting the positive correlation between local productivity and amenity shocks that increase outcomes as well as market access as measured in the post-period.[74]

Columns (5) and (6) instrument for the change in market access both by holding initial employment and population constant and computing the change in commute times were TransMilenio to have been constructed using the least cost path instrument. For residential outcomes, the point estimates are larger than columns (3) and (4) suggesting a negative bias between TransMilenio placement and growth in amenities.[75] This seems plausible, given that the system was build to serve areas of the city that had been growing and may therefore have slowed down over the period as they became congested. Commercial outcomes are more noisy, but the overall pattern is that the IV estimates are slightly higher than the previous estimates. That the estimates remain constant as additional controls are added provides additional evidence in support of the exclusion restriction holding.

Finally, columns (7) and (8) use both the tram and LCP instruments. The coefficients remain stable compared to using the LCP instrument alone, and in all but one case I fail to reject validity of the overidentification restrictions.

Heterogeneous Effects of Transit Figure 5 plots the non-parametric relationship between (residual) growth in outcomes and (residual) changes in commuter market access. The relationship appears ap-

[73]One limitation of my data is that variables do not line up over time periods and each specification may therefore rely on use changes over different periods. However, I will always use changes in market access constructed between the two waves in question and measure CMA using the values for population and employment in each period. For example, population regressions using differences between 1993 and 2005 measure changes in market access induced by phase 1 (opened between 2000 and 2003). Land market and employment regressions using differences between 2000-2012 and 2000-2015 respectively measure changes in market access induced by phase 1 and 2 (opened between 2005-2006). I explore whether future station openings predict prior growth in outcomes in falsification regressions below. My main employment and population regressions are weighted by initial establishment counts and population respectively to increase precision, but in robustness checks I show the results also hold in unweighted regressions. I also restrict the sample to tracts within 3km of stations for main specifications to ensure the results are not being driven by changes in commuter market access in very distant tracts, but include all tracts in robustness checks.

[74]F-Stats are not reported for clarity; they are extremely high for in columns (3) and (4).

[75]Part of this could also be due to the instrument correcting for measurement error in market access.

proximately log-linear for each outcome variable, supporting the functional form implied by the model. Importantly, this suggests the model performs well in fitting the heterogeneous effects observed in the data: tracts that experience large improvements in market access report large changes in outcomes.

Robustness In the appendix, I report a number of additional results which I summarize here.

First, I use less model-dependent measures of resident and firm commuter market access. These are commute-time weighted sums of employment and residence respectively, and recall the "market potential" discussed by Harris (1954) and alluded to in the discussion of accessibility in Hansen (1959).[76] The results are robust to using this alternative measure, suggesting my findings are not sensitive to using the structure implied by gravity models. That the coefficient from this measure differs statistically from the coefficient on commuter market access suggests that the adjustments in the gravity-based definition capture meaningful variation. Second, I run falsification tests to check that changes in market induced by particular lines are not associated with growth in outcomes before they open. Third, I condition on distance to stations to show that the effects are driven by changes in market access rather than characteristics of stations (such as changes in crime and complementary infrastructure. Fourth, I assess the response of variables to changes in market access to distant locations more than 1.5km away. Both of these empirical approaches are not possible with a distance-based empirical approach. Fifth, I use alternative speeds to compute the commute times for each mode. Sixth, I vary the commute elasticity θ to 1.5 and 0.5 times its estimated value. Seventh, I include all census tracts in the analysis, rather than those within 3km of a station. Eighth, I run unweighted regressions for employment and population regressions which are weighted in the main specifications. Finally, I use Conley (1999) HAC standard errors (compared to the baseline estimates which cluster by census tract) to allow for arbitary spatial correlation of errors across tracts within 1km of each other. That my results are robust to these alternative specifications provides additional evidence in support of the causal effect of TransMilenio on urban outcomes through improvements in commuter market access.

Comparison with Distance Band-Based Predictions In the appendix, I compare the predictions for residential house price growth in the commuter market access based model with those from a distance-based regression of the change in floorspace prices on two dummies for being closer than 750m from a station and between 750-1500m from a station (relative to the omitted tracts between 1.5-3km away). I find the dissimilarity index for the predicted changes in 0.631, with appreciation over- (under-)predicted in the center (outskirts).[77] Given the log-linear observed relationship between changes in commuter market access and house prices, this suggests that a government relying on the distance-based approach would make incorrect predictions over house price growth in response to new transit.

[76] In particular, I define $RMP_i = \sum_j t_{ij}^{-1} L_{Fj}$ and $FMP_i = \sum_j t_{ij}^{-1} L_{Rj}$ as resident and firm market potential respectively.
[77] This varies between 0 and 1 with 0 indicating identical predictions in each location

7.3 Results: Additional Outcomes

I now examine the impact of commuter market access on other outcomes outside the reduced form log-linear specification but that the full model in Section 4 suggest should be connected.

Wages and Labor Market Outcomes [TO ADD FROM DANE SERVERS]

Commute Distance Table 4 examines whether TransMilenio led to changes in commuting distances. Column (1) of panel A shows that changes in market access caused by TransMilenio indeed were associated with greater probability of using the system in 2015, providing reassurance that the measure is capturing changes in commuting opportunities through the system. The remaining columns run difference-in-difference specifications similar to (16) to how changes in market access affected commute distances within residential locations (UPZs) between 1995 and 2015. Throughout the OLS and IV specifications, improvements in commuter market access led to increases in commute distances, suggesting the system made employment in more distant locations more attractive.

Panel B tests for whether this effect is heterogeneous across high- and low-skill workers. The effect on commute distances is greater, but only mildly so, for low-skill workers. This likely reflects both their greater reliance on public transit as well as a greater sensitively of commute flows to commute costs as shown in the next section.

College Share A key question surrounding the debate on the effects of public transit is whether it leads to a re-sorting of worker (skill) groups. In the US, investments in transit have typically been followed by reductions in the share of rich residents (e.g. Glaeser et. al. 2008) although there is evidence this effect varies across different types of neighborhoods (Heilman 2017). However, the evidence in developing countries is far sparser.

Table 5 explores how the share of college residents in a census tract responds changes in residential commuter market access. Column (1) shows that, on average, there was no significant effect on demographic composition. However, this may mask heterogeneous responses across types of census tracts. For example, the model predicts that the high-skilled are more willing to pay for improved access to jobs in neighborhoods with high amenities.[78] In columns (2) to (5), I test whether the response differed by tracts according to the college share of the surrounding neighborhood.[79] The results show that the college share

[78]Assuming one mode of transit for the moment and the same wage for skill groups (i.e. $\sigma_L \to \infty$), log-linearizing the expression for residential populations (4), the change in residential population can be written as

$$\Delta \ln L_{Rig} \approx \eta_g \frac{\mu_{ig}^Y}{\theta_g} \Delta \ln \Phi_{Ri} - \eta_g (1 - \beta + \mu_{ig}^R) \Delta \ln r_{Ri} + \eta_g \Delta \ln u_{ig}$$

where , $\mu_{ig}^Y = \frac{T_g \Phi_{Ri}^{1/\theta_g}}{T_g \Phi_{Ri}^{1/\theta_g} - r_{Ri} \bar{h}}$ and $\mu_{ig}^R = \frac{r_{Ri} \bar{h}}{T_g \Phi_{Ri}^{1/\theta_g} - r_{Ri} \bar{h}}$. While μ_{ig}^Y is decreasing in income, μ_{ig}^R is decreasing in income since the poor spend a greater share of income on housing. Since μ_{ig}^R is greater in initially expensive (high amenity) neighborhoods, the low-skilled workers are more sensitive to house price appreciation there and thus should be less willing-to-pay for improved commuter market access than the rich.

[79]I measure a tract's surrounding college share using the share of college residents within a 1km disk around each tract

of a tract's residence did increase in response to an increase in market access, but only in neighborhoods with an initially high college share. In other words, the high-skilled were only willing to pay for improved transit access in nicer neighborhoods, and would not trade off these benefits for the lower amenities in poorer locations in the South. In contrast, the low-skilled were more likely to move into neighborhoods with a lower initial college share. Overall, this suggests that TransMilenio increased residential segregation between the low- and high-skilled.

8 Structural Estimation

Having empirically established the causal effect of TransMilenio on land and labor market outcomes through improved commuter market access, in this section I structurally estimate the full model from section 4. Since this model contains multiple groups, it allows to quantitatively assess the distributional effects of TransMilenio.

The section proceeds as follows. I first describe how the model can be inverted to obtain the unobservable wages, amenities and productivities that rationalize the observed data as an equilibrium of the model. I then outline the procedure to estimate the model's parameters. Finally, I present the estimation results and model diagnostics.

8.1 Model Inversion

The model contains location charactersitics, such as productivities, amenities and land use wedges, that are unobserved but needed to solve for counterfactual equilibria. While the presence of agglomeration forces allows for the possibility of multiple equilibria, a key advantage of my approach is that I am able to recover unique values of composite productivities and amenities that rationalize the observed data as an equilibrium.

There is a key difference in the process to solve for unobservables between this paper and recent quantitative urban models (e.g. Ahlfeldt et. al. 2016). In those models, there is one group of workers. Given data on where individuals live and work, it is straightforward for the vector of wages rationalize this distribution of residence and employment given observed commute costs. With wages in hand, recovering the remaining unobservables from the model's equilibrium conditions is straightforward.

In a model with different skill groups of workers, one would need data both on where skill groups live and work for this procedure to work. While data on residence by skill group are typically available from population census', I am not aware of similar datasets that provide employment by skill group within small spatial units within cities. This is where the incorporation of multiple industries becomes useful. In the data I observe employment by industry. Intuitively, given the different skill use intensities across industries, the relative employment by industries in a location should be informative about the relative

centroid in 1993 (excluding the tract itself). I then define a high college dummy equal to one for tracts in the top two terciles of its distribution. The results are robust to using own-tract college share, but this is subject to a mechanical bias from including the lagged value of the dependent variable. The results are also robust to proxying for college share using a dummy for localities in the North of the city, since this is where most of Bogotá's college educated live (Figure 2).

employment across skill groups. The following proposition formalizes this intuition, and shows that a unique vector of group-specific wages can be recovered using data on residence by skill and employment by industry. Recovering the remaining unobservables is simple.

Proposition 3. (Model Inversion) ***(i) Wages*** *Given data on residence by skill group L_{Rg}, employment by industries L_{Fs}, commute costs d_{ija} and car ownership shares $\pi_{a|ig}$ in addition to model parameters, there exists a unique vector of wages that rationalize the observed data as an equilibrium of the model.*

(ii) Productivities and Amenities *Given model parameters, wages and data $\{L_{Rig}, \pi_{a|iag}, L_{Fjs}, H_i, \vartheta_i, r_{Ri}, r_{Fi}\}$ there exists a unique vector of composite amenities and productivities $\{u_{ig}, A_{js}\}$ which rationalize the observed data as an equilibrium of the model.*

With this result in hand, I now describe my procedure to structurally estimate the parameters of the model.

8.2 Parameter Estimation

8.2.1 Parameters Calibrated to Exogenous Values

I calibrate $\{\sigma, \sigma_D, \tilde{\alpha}_s\}$ to existing values from the literature. I set the elasticity of substitution between labor skill groups to $\sigma = 1.3$ based on the review in Card (2009). I set the cost share of commercial floorspace to the estimates from Greenwood, Hercowitz, and Krusell (1997) who measure the share of labor, structures and equipment in value added for the US to be 70 percent, 13 percent, and 17 respectively. A floorspace share of $1 - \tilde{\alpha}_s = 0.156$ corresponds to their estimates renormalized to exclude equipment which is absent from my model. I set this share to be equal across all industries. I set the elasticity of substitution of demand to $\sigma_D = 6$ close to median estimates from Feenstra et. al. (2014). I vary both elasticities of substitution in robustness checks.

I now discuss how I estimate the parameters $\{\beta, \alpha_{sg}, t_{ija}, \kappa, \theta_g, \rho_g, T_g\}$ using relationships from the model.

8.2.2 Parameters Estimated without Solving the Model

Share Parameters I estimate $\beta = 0.24$ to match the long-run housing expenditure share in Bogotá.[80] I estimate the labor shares α_{sg} by industry using the average share of the wage bill paid to college and non-college educated workers in Colombia between 2000 and 2014 in all cities other than Bogotá. Since I assume that firms outside Bogotá aggregate labor using Cobb-Douglas technology, these labor cost shares identify α_{sg}.

Commute Costs The appendix outlines how commute times for car and non-car owners are constructed using averages of the time on each available model implied by a discrete choice model. The commute

[80] See the appendix for the Engel curve for housing and details on its estimation.

time between each pair of census tracts for each mode is computed in ArcMap. This allows me to estimate κ by estimating the mode choice model to the commute microdata by maximum likelihood.

Table 7 reports the results. I obtain an estimate of $\kappa = 0.012$, the semi-elasticity of commute costs to commute times, remarkably close to the estimate of 0.01 in Ahlfeldt et. al. (2015). This is despite me using a different source of variation for identification through the sensitivity of mode choices to commute times within origin-destination pairs. The last entry reports a value of 0.14 for the parameter λ governing the correlation of preference shocks within the public nest. This suggests that preferences are far more heterogeneous between public transit and cars than within the different modes available within public nests (i.e. walk, bus, TransMilenio). Given this order of magnitude difference, my baseline specifications assume that users take the quickest public mode of transportation available but imperfectly substitute across cars and public transit. I explore the sensitivity of my quantitative results to alternative aggregation methods. The remaining parameters reflect average preferences for each mode relative to walking (conditional on commute time). Intuitively, cars are most attractive followed by buses and TransMilenio. That TransMilenio is least desirable likely reflects the high crowds using the system as well as the inconvenience of having to walk between stations and final origins and destinations.

Skill Distribution The gravity equation for commute flows in (2) combined with the specification of commute costs $d_{ija} = \exp(\kappa t_{ija})$ implies a semi-log gravity equation for (conditional) commute flows

$$\ln \pi_{j|iag} = \gamma_{iag} + \gamma_{jg} - \theta_g \kappa t_{ija} + \varepsilon_{ijag}$$

where γ_{iag} and γ_{jg} are fixed effects and ε_{ijag} is classical measurement error. In order to leave sufficient residual variation after conditional on fixed effects, I aggregate to the locality level.

The results are reported in Table 8. Column (1) runs the specification above, including zeros by adding the minimum observed number of commuters to empty observations. Column (2) excludes empty observations. Column (3) weights by the number of trip observations in each origin-destination-skill-car ownership combination to place more weight on more precise estimates. Columns (4) and (5) estimate the model using poisson pseudo-maximum likelihood (PPML) as is often used to handle zeros in gravity equations (e.g. Silva and Tenrayo 2007). Given the oversize role of zeros, my preferred estimates come from the weighted specification in column (3) since these correspond most closely to the PPML results that are robust to this concern. Given the estimates of κ, these correspond to $\theta_H = 1.448$ and $\theta_L = 2.414$. These are similar to average value of 1.36 reported in the US in Hsieh et. al. (2016), as well as the estimates of 1.97 and 1.61 in the US from Lee (2015) for high school drop outs and college graduates respectively. This is despite these parameters being identified off commuting data alone in my setting. In the appendix I plot the residuals from this specification to show that the gravity equation fits commute flows well.[81]

[81] Finally, while the only parameter that matters for the computation of equilibria is the ratio $\theta_g = \tilde{\theta}_g/(1-\rho_g)$, I am also able to use wage data to seperate $\tilde{\theta}_g$ and ρ_g in my setting. I report the results and procedure in the appendix.

8.2.3 Parameters Estimated Solving the Full Model

It remains to estimate the parameters $\{\bar{h}, p_a, T_g, \eta_g, \mu_A, \mu_{U,g}\}$.

To begin, notice that given the parameter estimates in the previous section, for any value of T_g it is possible to solve for the full distribution of wages across the city. Since the vector T_g is only identified to scale, I normalize $T_L = 1$ and calibrate T_H so that the skill premium of wages in the model perfectly matches that observed in the data.[82] Given the distribution of wages and a value of the relative skill shifter, $\bar{h}$ and p_a are solved for so that the average expenditure share on housing and cars in the model exactly match those in the data. See the appendix for details.

The final step is to solve for the residential supply elasticity η_g and spillover parameters $\mu_A, \mu_{U,g}$. While the parameters estimated so far were identified using cross-sectional data, these require exogenous variation in the density of residence of skill groups and employment across the city. I therefore exploit the fact that changes in market access induced by TransMilenio provide a shock to the supply of labor and residents across the city.

Productivity Moment Taking logs of the expression for firm sales I find that

$$\frac{1}{\sigma_D - 1}\Delta \ln X_{js} = -(1-\alpha)\Delta \ln r_{Fj} - \alpha \Delta \ln W_{js} + \mu_A \Delta \ln \tilde{L}_{Fj} + \gamma_\ell + \gamma_F' \text{Cont}_j + \Delta \ln \bar{A}_{js}$$

where γ_ℓ and Cont_j are locality fixed effects and tract characteristics to (partially) control for factors driving growth in unobserved productivity, and $\Delta \ln \bar{A}_{js}$ (with an abuse of notation) is the residual. Given the parameters estimated above, all the variables in this regression equation are observed. For the calibrated values of σ_D and α, the key parameter of interest is μ_A. This is identified from the response of sales in a location to changes in total employment, where sales are model residuals. The model infers high sales when employment in a location is high relative to the observed price of commercial floorspace and the accessibility to workers through the commuting network (which determines wages). Intuitively, stronger agglomeration spillovers are identified when the composite productivity term depends more on local employment. The identification challenge is clear given that employment growth will be positively correlated with the error term.

I instrument for changes in employment using the LCP instrument for growth in firm commuter market access. Guided both by the model and the reduced form results of the previous section, this should be correlated with employment growth (by acting as a positive labor supply shock) but uncorrelated with growth in unobserved productivities. Letting Z_j denote this instrument,[83] the moment condition is given by

$$E\left[\Delta \ln \bar{A}_{js} Z_j\right] = 0.$$

[82] This involves jointly solving the for values of $\{T_H, w_{jg}\}$ such that the wages are an equilibrium of the model and the wage premium equals the observed value in the GEIH/ECH data.

[83] In practice, I use the change in the sum of market access across groups $Z_j = \Delta \ln \Phi_{Fj}$ where $\Phi_{Fj} = \sum_g \Phi_{Fjg}^{1/\theta_g}$ since this delivered the most precise estimates. I also include orthogonality conditions with each control variable, and demean each variable by locality prior to estimation, to estimate γ_F and control for fixed effects.

Amenities Moment Taking logs of the expression for residential populations delivers the following expression for residential population growth across skill groups

$$\Delta \ln L_{Riag} = \eta_g \Delta \ln V_{iag} + \eta_g \mu_{U,g} \Delta \ln \frac{L_{RiH}}{L_{Ri}} + \gamma_\ell + \gamma_R' \text{Cont}_i + \Delta \ln \bar{u}_{iag}$$

where $\Delta \ln V_{iag} \equiv \Delta \ln \left(T_g \Phi_{Riag}^{1/\theta_g} - r_{Ri} \bar{h} - p_a a \right) - (1-\beta) \Delta \ln r_{Ri}$ is the change in indirect utility from living in (i, a) net of changes in amenities. Once again, I include locality fixed effects and tract characteristics to (partially) control for factors driving growth in unobserved amenities. For each group, I require sources of exogenous variation in both the common component of utility from living in (i, a) ($\Delta \ln V_{iag}$) as well as the college share of residents living there ($\Delta \ln L_{RiH}/L_{Ri}$).

I instrument for the change in indirect utility using the LCP instrument for the change in residential commuter market access. I use two additional instruments to provide seperate variation in the share of college residents. First, I exploit the fact that conditional on a group's own RCMA shock within a location, tracts which experience a greater growth in commuter market access to high skill jobs relative to low skill jobs $Z_i^{Diff} = \Delta \ln \Phi_{RiH} - \Delta \ln \Phi_{RiL}$ should experience a larger increase in the share of college residents.[84] In practice I augment this by interacting the change in commuter market access for high skilled residents with the house price in the initial period $Z_i^{Rents} = \Delta \ln \Phi_{RiH} \times \ln r_{Ri}^{2000}$. That this should differentially predict entry of high-skilled residents is a direct consequence of log-linearizing the expression for residential populations (4).[85] The moment conditions are therefore given by[86]

$$E \left[\Delta \ln \bar{u}_{iag} Z_i \right] = 0, \quad Z_i \in \left\{ \begin{array}{c} \Delta \ln \Phi_{Riag} \\ Z_i^{Diff}, Z_i^{Rent} \end{array} \right\}$$

where I additionally include orthogonality conditions with controls and demean by locality fixed effects.[87]

8.2.4 GMM Results

[TO ADD: Values of $T_g, \bar{h}, p_a$. These are calibrated to perfectly match data, so no SEs, but add in footnote]

Main Results Table 9 presents the main results. Three comments are in order. First, the estimate of the productivity externality is large. Ahlfeldt et. al. (2016) obtain an estimate of 0.07 using a similar

[84] As when constructing the firm instrument, I define $\Phi_{Rig} = \sum_a \Phi_{Riag}$.

[85] See footnote 78 for exposition.

[86] For both moments, my baseline specification measures changes in outcomes between 2000 and 2015 and use the change in transit network due to the first phase of the system, since the raw population data at the tract level comes from 2005 (before using the 2015 UPZ totals to inflate to that year). I explore robustness to using both phases 1 and 2 in robustness checks.

[87] Since the unit of observation differs across firm outcomes (tract-industry) and residential outcomes (tract-skill group-car ownership), and there is no interdependencies between parameters across moment conditions (i.e. μ_A only affects the productivity moment condition while $\eta_g, \mu_{U,g}$ only affect the amenity moment conditions), I estimate the equations seperately via GMM rather than in one joint system, since it avoids the need to make arbitrary aggregation up to consistent units without any loss of information.

framework in Berlin, while the estimates in the literature have tended to lie within the 0.03-0.08 range reviewed in the survey by Rosenthal and Strange (2004). Other experimental approaches in the US have obtained estimates as high as 0.12 (Greenstone, Hornbeck, and Moretti 2010) and 0.2 (Kline and Moretti 2014), so this figure is not implausible. The majority of existing evidence is within the US and other developed countries. The returns to agglomeration may be higher in developing countries to do factors like a lack of road infrastructure or high crime, both of which are certainly at play in Bogotá. To my knowledge, this is the first intra-city estimate of agglomeration in less developed countries using quasi-experimental variation. However, in counterfactuals I turn spillovers off completeley as well as set them to a smaller concensus estimate for the US (0.06 from Ciccone and Hall 1996) to provide support that my quantitative results are not driven by this estimate alone.

Second, the residential population elasticity is greater for high skilled than low-skilled. I interpret this as reflecting other factors (such as capital constraints) that imply the poor face higher moving costs to relocate across the city than the rich. These elasticities are about 2-3 times larger than the commute elasticities, underscoring the benefits from estimating residential and employment location decisions as a two stage problem. Finally, the spillover parameters for residential amenities are large (though only twice as large as those in Ahfeldt et. al. (2016)) and slightly larger for high skilled (although the two coefficients are statistically indistinguishable from each other). In line with expectations, it appears the share of college educated residents in a tract increases the amenities from living there, and these endogenous forces appear stronger in Bogotá than existing evidence for developed countries.

Robustness In the appendix I test the robustness of these estimates. First, I control for log distance to the closest TransMilenio station (instrumented using the log distance to the LCP instrument). I find TransMilenio stations decrease both productivities and amenities, likely reflecting increased foot traffic (and potentially crime) near stations. Second, I show the estimates are qualitiatively similar when measuring the TransMilenio network as of 2006 (rather than 2003). Since the raw tract-level population data is from 2005, my preferred specification uses changes due to the first phase of the system, although I acknowledge the measurement error that this introduces. Finally, I vary the elasticity of substition of demand (from 4 to 9) and elasticity of substition between skill groups (from 1.3 to 2.5).[88] The point estimates are largely robust to the elasticity of substitution of labor, but the agglomeration point estimate is mechanically related to the demand elasticity as evident in the moment condition above. While my preferred estimate lies in the middle of the observed range, this underscores the need to check the robustness of my quantitative results to alternate values of this parameter.

8.3 Untargeted Moments: Model vs Data

In this section, I evaluate the performance of the model by comparing the model's predictions for moments not targeted in estimation.

[88] The value of 2.5 is estimated by Card (2009) for skill groups using regional data in the US.

Wages Figure 6 compares the average wage for each skill group earned by residents of each locality with that observed in 2014 in the GEIH data. The latter was not used in the procedure to estimate parameters determining wages. We see the two variables are highly correlated with values of 0.528 for non-college and 0.592 for college workers. However, while most observations lie along the 45-degree line for low-skilled workers, there is noticeable deviation for the richest localities amongst high skill workers. While the model is unable to capture all factors that drive differences in average income (which is unsurprising, given it iconsiders only labor income), the high correlation suggests that the spatial forces perform well in explaining income differences across the city.

Amenities and Productivities In the model, amenities and productivities represent characteristics that make locations more or less desirable to individuals and firms who might choose to locate there. Panel A of Table 10 shows that neighborhoods with less crime are associated with higher amenities, and that both skill groups have similar valuations over these characteristics. Panel B shows that productivities are higher in tracts with less crime , flatter slope and with a higher density of roads.[89] Overall, the model performs well at capturing features that affect the desirability of locations in the city.

Commute Flows The model solves for commute flows by first solving for wages that rationalize the observed distribution of residential population by skill and employment by industry, and then predicts the commute flows between origin, destination and car ownership pairs according to the gravity equation (2). I test the performance of the model's assumptions by comparing these implied commute flows with those observed within each cell in the 2015 Mobility Survey (again aggregating to the locality level). Other than the share of car owners in each UPZ, this data was not used in estimation or in solving for the model's unobservables. We see that the model performs very well matching the commute flows observed in the data, even when looking within car ownership groups (which will fit well by construction). Most importantly, the fit is even across college and non-college workers, suggesting the method used to back out wages by skill group using the location of employment by industries performs well in predicting commute flows.

Employment By Skill Group To provide more evidence that the model performs well in fitting the distribution of employment by skill groups in Bogotá, I compare the skill employment ratio $\ln(L_{FiH}/L_{FiL})$ within each UPZ in the model with that implied by trips to work in the 2015 Mobility Survey.

To show the importance of the ingredients in the model, panel (a) plots the results from a simplified baseline model in which labor skill groups are perfect substitutes and share the same commute elasticity (set to the average value). In this model, relative employment by skill group has an oddly smooth pattern that slowly declines as one moves further south in the city. This is because workers all receive the same wage across the city and have the same sensitively to commute costs, so the only differences in commuting decisions are caused by differences in their residential locations. Thus, the supply of high skilled workers

[89]Road density could of course affect productivity directly rather than through affecting the supply of labor as emphasized through commuting in the model. For example, better roads might make it easier to ship goods from or order supplies to offices. Slope might affect productivity through delivery accessibility in the same way.

is much greater in Northern UPZs close to where they live, and vice versa for the poor who live in the South. This pattern is clearly counterfactual to the distribution in the data shown in panel (c). By contrast, the baseline model with performs much better in matching this spatial distribution of the employment of relative skills (panel (b)): the correlation between the data in panel (b) and (c) is 0.406 compared to the correlation of 0.256 between that in panel (a) and (c).

9 Quantifying the Effect of TransMilenio

In this section, I use the estimated model the quantify the impact of TransMilenio by simulating the effect of its removal from the present day equilibrium.[90]

9.1 Removing the System

Main Results Table 11 presents the baseline effects of removing TransMilenio on GDP, total rents and welfare. Each entry reports the negative of the percentage change in each variable from removing the first two phases of the system. Panel A presents the closed city results, where the population of the city remains constant and utility adjusts in equilibrium. The effects on all outcomes are large, independent of whether spillovers are included or not: TransMilenio increases city GDP between 3.46%-4.39%, total city rents by 3.59%-3.57% and worker welfare by around 4%, the higher (lower) number referring to the case with (without) spillovers.

In the open city, there are no effects on worker welfare which is fixed at the reservation level in the wider economy. Instead, gains to workers can be read off changes in population. The effects of TransMilenio on population are large, increaseing the population of the low skilled by 9.49%-13.45% and the high skilled by 11.39%-15.33%. Given the increase in factor supply to the city, it is no surprise that the increase GDP between 11.8%-19.16% is much larger. This increase in population raises house prices which, recall, is in fixed supply. Gains are therefore shifted from workers to land owners who see total rents rise between 14.85-19.12% due to TransMilenio.[91]

Figure 9 plots the change in employment and population by tract characteristics. Panel (a) shows that tracts with the largest employment lose the most when TransMilenio is removed. By enabling productive locations (with high employment) to grow the most, the system's efficiency gains are driven by an improvement in the spatial allocation of labor. Panel (b) shows similar patterns hold for residence, but the effects are much more muted than for employment. Panel (c) shows that the model replicates the increase in college share in response to market access only in tracts with a high initial share. This response is exaggerated with positive spillovers, reflecting both the difference in spillover parameters between skill groups as well as the multiplier effect that increased inflows of high-skill workers have on amenities.

Returning the welfare effects in the closed city model, somewhat surprisingly we see that the welfare

[90]See the appendix for additional information on the algorithm used to solve for counterfactual equilibria.

[91]Note that the relative magnitude of the effects on outcomes such as output and floorspace values across open and closed city models is similar to those in Ahfleldt et. al. (2016) in the context of changing the subway network in Berlin.

of high skill workers (4.41-5.18%) rises more than for the low skilled (3.74%-4.61%), leading to a 0.692% rise in welfare inequality (defined as $\bar{U}_H/\bar{U}_L$) as a result of the system. I now turn to understanding the channels through which these differential welfare gains are accruing.

Decomposing the Channels Table 12 decomposes the channels starting from the simplest model and slowly incorporating each elements.[92]

In Model 1, worker skill groups take on the same values of the elasticities θ and η and enter as perfect subtitutes in production. Relative demand for skills across the city is constant, and labor supply from any given residential location is be identical between worker groups. Differences in relative employment by skill are driven solely by variation in residential locations.

In row (1), the system is removed holding all location decisions and prices fixed. The low skilled clearly benefit more from TransMilenio given their greater dependence on public transit. Rows (2) to (4) hold location decisions fixed but let prices (and land use) in different markets adjust in sequence. Rows (2) and (3) let rents adjust with and without the subsistence housing requirement. The welfare gap narrows substantially in both cases: the low-skilled live in the outskirts of the city which experience greater increases in commuter market access, and are therefore hurt more by the accompanying house price appreciation.[93] The results in the following rows are qualitatively similar as wages, employment location, residential location and car ownership decisions adjust. In sum, the poor benefit in this model due to their greater reliance on public transit, but are hurt more by house price appreciation in the outlying neighborhoods in which they live. The net effect is still that they benefit more from the system, with inequality falling 1.058% as a result of TransMilenio.

Model 2 extends the previous model by allowing skill groups to differ in their labor supply elasticities θ_g. In row (1), we see that this has an immediate impact on attenuating the relative welfare gain of the low-skilled relative to model 1 (inequality falls 0.556% vs 1.529%). This highlights the crucial role played by θ_g in the determining the welfare gains from transit: a lower elasticity θ_g implies these individuals are less willing to substitute between employment locations, and thus the incidence of commute costs falls more broadly on their shoulders.[94] The qualitative impact of the remaining channels remains similar. With spillovers shut down, it turns out that the differences in θ_g across groups are enough the reverse the sign of the change in inequality in this model (row (7)), but this reverses once more once productivity spillovers are included that disproportionately increases low skilled wages.

Model 3 allows for the residential supply elasticities η_g to differ between groups at their baseline

[92] For each model, both the baseline and counterfactual equilibria are recomputed. This involves solving for updated values for $T_H, \bar{h}, p_a$ and w_{jg}. Spillovers are shut down in all but the last two rows.

[93] The fact that the welfare change is similar with and without subsistence housing requirement shows that the majority of the costs of house price appreciation to the low-skilled comes from the fact that house prices appreciate more where they live rather than from their increased expenditures on housing. In fact with the subsistence housing requirement, the poor are actually hurt less since residential locations are fixed and their lower demand for housing dampens rises in house prices.

[94] Note that θ_g impacts both the substitution patterns as well as the calibrated wages. In results available on request, I compute the welfare gains in an intermediate scenario in the specification in row 1 by using the wages from model 1 but the θ_g from model 2 to compute the welfare gains. The majority of the observed drop in inequality observed in Table 12 is due to θ_g's role on substititability between employment destinations.

values. The low-skill gain slightly more in this case relative to model 2 since they are less able to substitute between residential locations (η is smaller), but overall the results remain qualitatively similar.

Finally, model 4 sets the elasticity of substitution between skill groups equal to its baseline value. Now the demand for worker skill groups differs across the city, due to differences in the location of high- and low-skill intensive industries. This last change has a large effect on inequality, and in combination with the differences in commute elasticities high skilled residents benefit more from the system.

To summarize, the primary channels determining the effects of improved transit on welfare inequality are as follows. First, the low-skilled benefit due to their relative reliance on public transit. Second, they are more able to subsistute away from more costly commutes than the high-skilled, who have more heterogeneous match productivites with firms across the city which increases their willingness to bear high commute costs. Improvements in transit benefit the high-skilled through this channel. Third, geographical features of the city and its transit network influence who gains. House prices appreciated most in the outskirts of the city, hurting the low-skilled. Differences in the relative demand for workers also created differences in wages across the city across worker group. The group for whom locations where many lived were connected to workplaces with high wages benefitted the most. In Bogotá, this operated in favor of the high-skilled whose residence is concentrated in the North and who receive highest wages in the high-skill intensive industries location in the center and North. By contrast, employment and residence of low-skill intensive industries is more spreadout through the South and West of the city. The net effect is that welfare inequality between the low- and high-skilled increased by 0.692% as a result of TransMilenio.

Costs vs Benefits How did the output gains from TransMilenio compare with the costs of the system? Panel A of Table 13 provides a breakdown of the costs and benefits of the system.[95] Even using the most conservative estimate in column (1), I find that the net present value of the net increase on GDP was about $56bn, or a net increase of 3.08% in the steady-state level of GDP. This suggests the system was a highly profitable investment for the city.

9.2 Robustness and Model Extensions

In the appendix, I report the robustness of these quantitative results to alternative parameters and modelling assumptions. The implications are robust to a variety of alternative parameter values, I omit discussion here and instead describe the model extensions and their results.

Incorporating Employment in Domestic Services As discussed in the appendix, domestic services is a a key source of employment for the low skilled in Bogota, constituting around 7% of their total employment.[96] This employment is not reported in either the census and CCB datasets. The model may therefore underestimate the benefits to the low skilled from TransMilenio to the extent that the system improves access to jobs as domestic servants in the houses of the college educated in the North. However,

[95] Add info on numbers either here in footnote, or in appx.

[96] By contrast, almost no college educated workers are employed as domestic servants.

the high-skilled consume domestic services directly and thus would have benefited from their increased supply through lower wages, making the net effect on inequality ambiguous.

In the appendix I extend the model to include employment in domestic services and discuss its calibration procedure. While the relative gains to the high-skilled fall slightly, the net benefit of the system to them is qualitatively unchanged. All in all, the benefit to the low-skilled from better access to jobs as domestic servants is roughly offset by the accompanying benefit to the high-skilled who have better access to domestic workers.

Incorporating Home Ownership [TO ADD]

Alternative Timing Assumption [TO ADD]

Preference Shocks [TO ADD]

10 Policy Counterfactuals

10.1 Impact of Different Lines and Planned Cable Car

Table 14 evaluates the effects of different portions of the system.

Column (2) simulates the effect of removing line H connecting the Southern most portions of the city with the CBD. Since this area has a much greater density of low skilled residents, the welfare effects are greater for the poor. However, since less people live there the total effect on both welfare and output is somewhat muted.

In contrast, column (3) examines the impact of removing line A connecting the Northern parts of the city to the CBD. The effects of this part of the system are substantial. This line connects high-skilled residents in the North with jobs in the two dense employment districts in the center and North of the city, and allows the low-skilled living in the South from access to jobs in the North. The effect of removing this line on welfare is about twice as large as line H; the effect on output is more than 6 times the size.

Finally, column (4) simulates the effect of removing the feeder system connecting outerlying areas with portals using buses that run on existing roadways. This system benefits both worker groups about the same, skewing the results towards the low-skilled who live in outlying areas. The large effects on both output and welfare are especially noteworthy given the minimal capital costs and low operational costs of running feeder buses when compared with BRT. This underscores the large benefits to providing solutions to the "last mile problem" to help ease connections betweeen final destinations and system stations.

Column (1) of Table 14 evaluates the impact of adding a Cable Car to the slums in the hills of Ciudad Bolivar in the South. This system is planned to open in 2019. Despite the unsirprisingly small aggregate effect of the line which has only 3 stops, the cable car disproportionately benefits the low-skilled.

10.2 Land Value Capture

One of the main criticisms of TransMilenio was that the city experienced such a large change in transit without any adjustment of zoning laws to allow housing supply to respond where it was needed. I show in the appendix that housing supply did not respond to the system's construction, consistent with other evidence on the restrictive role played by land use regulation (Suzuki et. al. 2013). Many cities, such as Hong Kong and Tokyo, have had success in implementing Land Value Capture (LVC) schemes which increase permitted densities around new stations but charge developers for the right to build there. These policies achieved the dual aim of increasing housing supply and raising revenue to finance the construction of the system.

In this section, I evaluate the impact on Bogotá if housing supply would have been able to respond in conjunction with the opening of the system. In extreme case, I assume that housing supplies are able to freely adjust, reflecting what might occur in the long-run adjustment. This provides a useful upper bound on the potential welfare loses from stringent zoning. I then simulate the effect of two potential land value capture schemes implemented by the government. First, I assume the government sells the rights to developers to increase floorspace by a maximum of 30% within a certain distance band from stations, mimicking the "air rights sales" undertaken in Asian cities. Second, I assume the government sells permits that allow for the same change in total floorspace, but instead allocates the permitted floorspace changes according to a location's predicted change in commuter market access.[97] Details are provided in the appendix.

Table 15 presents the results. In the closed city model, average welfare increases by 28% more than under fixed housing supply. Importantly, note that the housing supply response benefits the low-skilled since it tempers appreciation the most in the outskirts where they live. The same qualitative patterns and present in the open city model, but the numbers are larger as seen before.

Finally, panel B of Table 13 converts these percentage terms into fractions of overall capital costs of constructing the system. In the more conservative clsoed city case, the distance-band based permit measure recoups only 53.9% of the capital cost of the system, compared to the 127.7% earned using the commuter market access-based permit.

These results suggest both that complementary changes in zoning regulations can have large effects on welfare and can also be used to finance the system. Additionally, the comparison with the distance-based policy underscores how measures of commuter market access can be a used as parsimoneous tools for governments to guide the allocation of rezoning.

[97] In particular, I let the change in permitted FAR be proportional to $\vartheta_i \Delta \ln \Phi_{Ri} + (1 - \vartheta_i) \Delta \ln \Phi_{Fi}$ where ϑ_i are the residential floorspace shares in the initial equilibrium and $\Delta \ln \Phi$ are the intruemnts for the change in commuter market access holding population and employment at their initial values. Each of these values is based only on information the government has at the time of the policy change.

10.3 Optimal Route Placement

10.4 Bringing Jobs to the Low-Skilled

11 Conclusion

References

AHLFELDT, G. (2011), "If Alonso was Right: Modeling Accessibility and Explaining the Residential Land Gradient", *Journal of Regional Science*, 51(2): 318-338.

AHLFELDT, G.M., REDDING, S.J., STURM, D.M. AND N. WOLF (2015), "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 83(6): 2127-2189.

AKBAR, P., AND G. DURANTON (2017), "Measuring the Cost of Congestion in a Highly Congested City: Bogotá", Working Paper.

ALLEN, T., ARKOLAKIS, C. AND X. LI (2015), "Optimal City Structure," Working Paper.

ALLEN, T., ARKOLAKIS, C. AND Y. TAKAHASHI (2017), "Universal Gravity," Working Paper.

ALONSO, W. (1964), *Location and Land Use*. Harvard, Cambridge MA.

BARTELME, D. (2015), "Trade Costs and Economic Geography: Evidence from the U.S.", Working Paper.

BAUM-SNOW, N. (2007), "Did Highways Cause Suburbanization?", *Quarterly Journal of Economics*, 122: 775–805.

BAUM-SNOW, N., BRANDT, L., HENDERSON, J., TURNER, M., AND Q. ZHANG (2017) "Roads, Railroads and Decentralization of Chinese Cities", *Review of Economics and Statistics*, 99(3): 435-448.

BAYER, P., FERREIRA, F., AND R. MCMILLAN (2007), "A Unified Framework for Measuring Preferences for Schools and Neighborhoods" *Journal of Political Economy*, 115(4): 588— 638.

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

BOCAJERO, J-P., PEREZ, M.A., AND I. PORTILLA (2013), "Impact of Transmilenio on Density, Land Use, and Land Value in Bogotá", *Research in Transportation Economics*, 40:78-86.

BRT DATA (2017), *Global BRT Statistics*, Retrieved from http://brtdata.org.

CERVERO, R. (2013), "Bus Rapid Transit (BRT): An Ef cient and Competitive Mode of Public Transport", *Berkeley Institute of Urban and Regional Development Working Paper* 2013-01.

CERVERO, R., IUCHI, K., AND H. SUZUKI (2013), *Transforming Cities with Transit: Transit and Land-Use Integration for Sustainable Urban Development*. Washington, DC: World Bank.

CRACKNELL, J. (2003), "Transmilenio Busway-Based Mass Transit, Bogotá, Colombia", World Bank Working Paper

DIAMOND, R. (2016), "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000", *American Economic Review*, 106(3):479–524.

DONALDSON, D. (2012) "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, forthcoming.

DONALDSON, D. (2015), "The Gains from Market Integration", *Annual Review of Economics*, 7:619-647.

DONALDSON, D., AND R. HORNBECK (2015), "Railroads and American Economic Growth: A Market Access Approach", *Quarterly Journal of Economics*, 131(2): 799-858.

DURANTON, G., AND M. TURNER (2012), "Urban Growth and Transportation", *Review of Economic Studies*, 79: 1407-1440.

GIBBONS, S., AND S. MACHIN (2005), "Valuing rail access using transport innovations", *Journal of Urban Economics*, 57(1):148–169.

GLAESER, E., KAHN, M., AND J. RAPPAPORT (2008) "Why do the poor live in cities? The role of public transportation", *Journal of Urban Economics,* 63(1):1–24.

GONZALEZ-NAVARRO, M., AND M. TURNER (2016), “Subways and urban growth: Evidence from Earth”, Working Paper.

GUERRIERI, G., HARTLEY, D., AND E. HURST (2013), “Endogenous Gentrification and Housing Price Dynamics” *Journal of Public Economics*, 100:45-60.

GRAFTIEAUX, P., AND D. HIDALGO (2005), “A Critical Look at Major Bus Improvements in Latin America and Asia: Case Study TransMilenio, Bogotá, Colombia”, Working Paper.

GREENSTONE, M., R. HORNBECK, AND E. MORETTI (2010): “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings,” *Journal of Political Economy*, 118(3), 536–598.

HANSEN, W. (1959), “How Accessibility Shapes Land Use,” *Journal of the American Institute of Planners*, 25: 73–76.

HARRIS, C. (1954), “The Market as a Factor in the Localization of Industry in the United States,” *Annals of the Association of American Geographers*, 44(4), 315–348.

HEILMANN, K. (2016), “Transit Access and Neighborhood Segregation:A Study of the Dallas Light Rail System”, Working Paper.

HONG., Y., MURAKAMI, J., SUZUKI, H., AND B. TAMAYOSE (2015), *Financing Transit-Oriented Development with Land Values: Adapting Land Value Capture in Developing Countries*. Washington, DC: World Bank

KLINE, P., E. MORETTI (2014), “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority,” *Quarterly Journal of Economics*, 129(1), 275–331.

LEE, E. (2015), “Trade, Inequality, and the Endogenous Sorting of Heterogeneous Workers”, Working Paper.

LUCAS , R., AND E. ROSSI-HANSBERG (2002), “On the Internal Structure of Cities,” *Econometrica*, 70(4), 1445–1476.

MCARTHUR, D., L. OSLAND, AND I. THORSEN (2012), “Spatial Transferability of Hedonic House Price Functions”, *Regional Studies*, 46(5): 597-610.

MCKINSEY (2016), *Bridging Global Infrastructure Gaps*, McKinsey Global Institute

MENCKHOFF, G. (2005), “Latin American Experience with Bus Rapid Transit”, presented at Institute of Transportation Engineers Annual Meeting, Melbourne, Australia.

MILLS, E. (1967), “An Aggregative Model of Resource Allocation in a Metropolitan Centre,” American Economic Review, 57(2), 197–210.

MONTE, F., S. REDDING AND E. ROSSI-HANSBERG (2017), “Commuting, Migration and Local Employment Elasticities”, NBER Working Paper 21706.

MUTH, R. (1969), *Cities and Housing*, University of Chicago Press, Chicago.

OWENS, R., E. ROSSI-HANSBERG, AND P-D. SARTE (2017), “Rethinking Detroit”, Working Paper.

OSLAND, L., AND I.THORSEN (2008), “Effects on housing prices of urban attraction and labor-market accessibility”, *Environment and Planning*, 40: 2490-2509.

REDDING, S. (2010), “The Empirics of New Economic Geography”, *Journal of Regional Science,* 50(1), 297-311.

REDDING, S., AND D. STURM (2008), “The Costs of Remoteness: Evidence from German Division and Reunification,” *American Economic Review*, 98(5), 1766–1797.

REDDING, S., AND D. STURM (2016), “Estimating Neighborhood Effects: Evidence from War-time Destruction in London", Working Paper.

ROSENTHAL, S., AND W. STRANGE (2004), "Evidence on the Nature and Sources of Agglomeration Economics," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. Thisse, vol. 4. Elsevier North Holland, Amsterdam.

SEVEREN, C. (2016), "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification", Working Paper.

WRIGHT , L., AND W. HOOK (2007), *Bus Rapid Transit Planning Guide,* New York, NY: Institute for Transportation and Development Policy.

Tables

Table 1: College-Employment Shares by Industry

Industry	College Share	Employment Share
Domestic Services	0.085	0.050
Construction	0.181	0.052
Hotels & Restaurants	0.235	0.057
Wholesale, Retail, Repair	0.300	0.222
Manufacturing	0.315	0.173
Transport, Storage, Communications	0.341	0.089
Other Community, Social, Personal Serv	0.380	0.050
Real Estate	0.556	0.120
Social & Health Services	0.634	0.053
Public Administration	0.707	0.038
Education	0.810	0.052
Financial Brokerage	0.827	0.028

Note: Data is an average over 2000-2014 and comes from the GEIH and ECH. The first column shows the share of workers which have post-secondary education within each one-digit industry. The second column shows the industry's share of total city employment. Only industries accounting for at least 1% of employment reported.

Table 2: Commuting in 1995

	lnSpeed	lnSpeed	Bus	Bus
Bus	-0.353***	-0.305***		
	(0.021)	(0.016)		
Low-Skill			0.287***	0.163***
			(0.010)	(0.011)
R^2	0.06	0.76	0.18	0.47
N	14,841	12,877	18,843	16,461
UPZ O-D FE		X		X
Time of day Controls	X	X	X	X
Demographic Controls	X	X	X	X

Note: Low-Skill is a dummy for having no post-secondary education. Bus is a dummy for whether bus is used during a commute, relative to the omitted category of car. Data is from 1995. Time of day controls are dummies for hour of departure, and demographics are log age and a gender dummy. UPZ O-D FE are fixed effects for each upz origin-destination. Only trips to work included. Standard errors clustered at upz origin-destination pair. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 3: IV Results: Main Outcomes

	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV-LCP	(6) IV-LCP	(7) IV All	(8) IV All
Panel A: Residents								
ln(Res Floorspace Price)	0.470***	0.355***	0.208**	0.333***	0.463***	0.565***	0.435***	0.557***
	(0.091)	(0.074)	(0.097)	(0.079)	(0.152)	(0.118)	(0.158)	(0.122)
N	1,943	1,943	1,943	1,943	1,943	1,943	1,943	1,943
F-Stat					851.31	852.24	410.63	418.15
Over-ID p-value							0.52	0.16
ln(Residential Pop)	0.271**	0.283**	0.203	0.209	0.325*	0.348*	0.303*	0.327*
	(0.136)	(0.138)	(0.138)	(0.140)	(0.178)	(0.185)	(0.181)	(0.188)
N	1,997	1,997	1,997	1,997	1,997	1,997	1,997	1,997
F-Stat					1,747.80	1,706.02	859.94	844.56
Over-ID p-value							0.83	0.95
Panel B: Firms								
ln(Comm Floorspace Price)	0.223*	0.228*	0.259**	0.277**	0.261	0.278*	0.338**	0.338**
	(0.122)	(0.128)	(0.124)	(0.128)	(0.162)	(0.168)	(0.165)	(0.171)
N	1,884	1,884	1,884	1,884	1,884	1,884	1,884	1,884
F-Stat					1,265.27	1,114.61	812.06	733.28
Over-ID p-value							0.09	0.19
Comm Floorspace Share	0.164***	0.162***	0.163***	0.166***	0.124**	0.122*	0.131**	0.130**
	(0.043)	(0.045)	(0.044)	(0.046)	(0.059)	(0.064)	(0.061)	(0.065)
N	1,981	1,981	1,981	1,981	1,981	1,981	1,981	1,981
F-Stat					1,361.75	1,183.55	876.75	772.27
Over-ID p-value							0.52	0.57
ln(Establishments)	1.423***	0.880**	1.185***	0.780**	1.573***	1.488**	1.090*	0.974*
	(0.365)	(0.354)	(0.374)	(0.361)	(0.598)	(0.604)	(0.592)	(0.591)
N	1,724	1,724	1,724	1,724	1,724	1,724	1,724	1,724
F-Stat					224.50	264.90	333.56	293.00
Over-ID p-value							0.12	0.01
Locality Fixed Effects	X	X	X	X	X	X	X	X
CBD X Region Controls	X	X	X	X	X	X	X	X
Basic Tract Controls	X	X	X	X	X	X	X	X
Historical Controls		X		X		X		X
Init. Land Controls		X		X		X		X
Init. Demographic Controls		X		X		X		X
Distance to Tram Controls							X	X

Note: Observation is a census tract. Each entry reports the coefficient from a regression of the variable in each row on firm or residential commuter market access in first differences. Each column corresponds to a specification. Only tracts further than 500m from a portal and the CBD (and less than 3km from a station) are included. Controls are as described in previous table, other than distance to tram which is a dummy for whether a tract is closer than 500m from the historical tram line. Columns (1) and (2) run an OLS specification. Columns (3) and (4) instrument for the change in CMA holding residence and employment fixed at their initial levels and changing only commute costs, excluding the census tract itself from the variable construction. Kleinberg-Paap F-statistics are very high (>10,000) and not reported for brevity. Columns (5) and (6) instrument using the change in CMA induced by the LCP route, while (7) and (8) include both the LCP instrument and the change induced by the tram instrument. Robust standard errors reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 4: Commute Distance

Outcome	(1) OLS UseTM	(2) OLS lnDist	(3) IV lnDist	(4) IV-LCP lnDist	(5) IV All lnDist	(6) IV All lnDist
lnRCMA	0.957***	0.541**	0.383	0.904**	0.951**	0.310
	(0.212)	(0.252)	(0.300)	(0.404)	(0.404)	(0.466)
lnRCMA X High Skill						-0.147**
						(0.058)
N	9,088	22,119	22,119	17,212	17,212	19,920
R^2	0.07	0.10				
F-Stat				102.91	72.88	18.26
Over-ID p-value					0.52	0.55
UPZ FE	X	X	X	X	X	X
Locality FE X Post FE	X	X	X	X	X	X
Log Dist CBD X Region FE X Post FE	X	X	X	X	X	X
Trip Controls X Post FE	X	X	X	X	X	X
Tract Controls X Post FE	X	X	X	X	X	X
Historical Controls X Post FE	X	X	X	X	X	X
Educ X Post FE						X

Note: Observation is a trip, only trips to work are included. Column (1) reports coefficients from a regression of the probability an individual uses TransMilenio in 2015 on the change in lnRCMA in the origin UPZ. The other columns run difference-in-difference specifications using data from 2015 (Post) and 1995 (Pre), examining how changes in commute distances vary with changes in RCMA. RCMA is measured at the UPZ level using the pre-TM network in the pre-period and the 2006 network in the post period. Trip controls include hour of departure dummies and demographic characistics (sex, log age, hh head dummy, occupation dummies). Tract controls include log area, log distance to a main road and log population density in 1993. Historical controls include quartile dummies of 1918 population, dummy for whether closer than 500m to main road in 1933, and when the tram instrument is used a dummy for whether a tract is closer than 500m from the historical tram line. Last column includes education level dummies interacted with Post FE. Standard errors clustered by origin UPZ are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 5: College Share

Outcome: Change in College Share	(1) OLS	(2) OLS	(3) IV	(4) IV-LCP	(5) IV-All
Δ lnRCMA	-0.011 (0.030)	-0.046 (0.029)	-0.040 (0.029)	-0.065 (0.050)	-0.062 (0.047)
Δ lnRCMA X HighColl		0.051* (0.027)	0.063* (0.033)	0.107** (0.051)	0.111** (0.047)
N	1,886	1,886	1,886	1,886	1,886
R^2	0.27	0.27			
F-Stat				89.85	123.61
Over-ID p-value					0.55
Locality FE	X	X	X	X	X
HighColl FE		X	X	X	X
Log Dist CBD X Region FE	X	X	X	X	X
Tract Controls	X	X	X	X	X
Historical Controls	X	X	X	X	X

Note: Outcome is the change in a census tract's share of residents older than 20 with post-secondary education between 1993 and 2005. Dependent variable is change in RCMA between these years using the pre-TM and phase 1 of the system to measure commute times, interacted with a dummy for whether a tract is high college. The high college measure is constructed by first computing the share of college residents within a 1km disk around each tract centroid (excluding the tract itself) and then setting high college dummy equal to one for tracts in the top two terciles of its distribution. Specifications with interactions include an intercept to allow growth to differ across low and high college tracts (HighColl FE). Tract controls include log area, log distance to a main road and log population density in 1993; all other controls are as described in previous tables. Final column includes additional control for whether tract is closer than 500m from historical tram route. Columns (1) and (2) run OLS. Column (3) instruments for the change in CMA holding residence and employment fixed at their initial levels and changing only commute costs, excluding the census tract itself from the variable construction. Column (4) instruments using the change in CMA induced by the LCP route, while column (5) additionally includes the tram instrument. Only tracts further than 500m from a portal and the CBD (and less than 3km from a station) are included. Robust standard errors reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 6: Wages

Outcome: $\ln$Wage	(1) OLS	(2) OLS	(3) IV	(4) IV-All	(5) IV	(6) IV-All
lnRCMA	0.479*** (0.162)	0.202* (0.108)	0.282** (0.129)	0.221 (0.221)	0.188 (0.140)	0.185 (0.236)
lnRCMA X College					0.304*** (0.055)	0.298*** (0.054)
N	75,981	75,981	75,981	75,981	75,981	75,981
R^2	0.35	0.47	0.47	0.47	0.47	0.47
F-Stat				30.94		16.41
Over-ID p-value				0.94		0.64
UPZ FE	X	X	X	X	X	X
Region X Post FE	X	X	X	X	X	X
Log Dist CBD X Region FE X Post FE	X	X	X	X	X	X
Tract Controls X Post FE	X	X	X	X	X	X
Historical Controls X Post FE				X	X	X
College FE					X	X

Note: Dependent variable is the log hourly wage for full-time workers reporting more than 40 hours worked per week. Data covers 2000-2005 in the pre-period and 2009-2014 in the post period. RCMA is measured at the UPZ-level using the pre-TM network in the pre-period, and using the 2006 network in the post-period. IV specification uses both the LCP and Tram instruments. Region are dummies for the North, West and South of the city. Tract controls are dummies for terciles [FILL]. College is a dummy for having post-secondary education. Remaining controls are as described in previous tables. Standard errors are clustered by UPZ and period. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 7: Mode Choice Model Estimates

	(1)
Time	-0.012**
	(0.005)
Bus	-0.086*
	(0.050)
Car	0.837***
	(0.292)
TM	-0.216**
	(0.105)
λ	0.140**
	(0.064)
Time of Day Controls	X
Demographic Controls	X

Table shows estimation from nested logit regression on trip-level data from the 2015 Mobility Survey. λ is the correlation parameter for the public nest. Demographic controls include a sex dummy as well as dummies for quintiles of the age distribution, while time of day controls include dummies for the hour of trip departure. Each have choice-varying coefficients. Only trips during rush hour to and from work are included. Robust standard errors are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 8: Gravity Regression

	(1)	(2)	(3)	(4)
HighSkill X ln Commute Cost	-0.024***	-0.017***	-0.019***	-0.019***
	(0.004)	(0.003)	(0.003)	(0.003)
LowSkill X ln Commute Cost	-0.033***	-0.031***	-0.022***	-0.033***
	(0.004)	(0.003)	(0.003)	(0.003)
N	1,444	879	879	879
Origin-Skill-Car Ownership Fixed Effects	X	X	X	X
Destination-Skill Fixed Effects	X	X	X	X
Weighted		X		X
Method	PPML	PPML	OLS	OLS

Note: Outcome is the conditional commuting shares in cols 1-2, and its log in cols 3-4. Skill corresponds to college or non-college educated workers. Only trips to work during rush hour (5-8am) by heads of households included. Weighted columns weight by the number of observations in each origin-destination-skill-car ownership combination. Robust standard errors reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 9: GMM Results

Parameter	Estimate
Panel A: Firms	
μ_A	0.238**
	(0.118)
Panel B: Workers	
η_L	3.701**
	(1.515)
η_H	4.436***
	(0.985)
μ_U^L	0.305***
	(0.054)
μ_U^H	0.330***
	(0.043)
P-value $\eta_L = \eta_H$	0.50
P-value $\mu_U^L = \mu_U^H$	0.77

Note: Estimates are from 2-step GMM procedure separately for firms at the tract-industry level with 6137 observations and for workers at the tract-group-car ownership level with 7036 observations. Controls include log distance to CBD interacted with region fixed effects, commercial floorspace share in 2000, and log population density and college share in 1993 for employment moment conditions. Spillover parameter estimates obtained via delta method: original parameter clusters $\eta_L \mu_U^L$ and $\eta_H \mu_U^H$ are 1.128 (0.422) and 1.396 (0.280) respectively. Only tracts within 3km of the network and those more than 500m from portals and the CBD are included. Standard errors clustered at the tract reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 10: Amenities and Productivities: Model vs Data

Panel A: Amenities	(1) ln Crime Density	(2) ln Crime Density	
Amenity Elasticity	-0.160***	-0.196***	
	(0.012)	(0.016)	
Skill	College	Non-College	
R^2	0.30	0.22	
N	503	499	
Panel B: Productivities	(1) ln Crime Density	(2) ln Slope	(3) ln Roads
Productivity Elasticity	-0.046	-0.196***	0.109***
	(0.031)	(0.013)	(0.015)
R^2	0.01	0.24	0.08
N	504	615	615

Note: Estimates show coefficients from regressions of log (composite) productivities and amenities on variable given in each column. Observation is a sector. Crime is measured either as total homicides in a sector between 2007 and 2012. In column (2) of Panel B, the dependent variable is log of the average slope of land. In column (3), the dependent variable is log of 1 plus the kilometers of primary and secondary roads within a disk of 1.5km radius around the sector centroid. Standard errors clustered by sector reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 11: Effect of Removing Phases 1 and 2 of TransMilenio

	No Spillovers	Spillovers
Panel A: Closed City		
GDP	3.467	4.349
Rents	3.598	3.573
Welfare Low	3.741	4.606
Welfare High	4.407	5.182
Inequality	0.692	0.604
Panel B: Open City		
GDP	11.801	19.158
Rents	14.849	19.123
Population Low	9.489	13.445
Population High	11.388	15.334
Relative Population	2.097	2.182

Note: Table shows the (negative of the) value of the percentage change in each variable from removing phases 1 and 2 of the TransMilenio network from the 2012 equilibrium, with and without spillovers.

Table 13: Cost vs. Benefits of TransMilenio

	Closed City		**Open City**	
	No Spillovers	**Spillovers**	**No Spillovers**	**Spillovers**
Panel A: Costs & Benefits				
NPV Increase GDP (mm)	63,758	79,979	217,022	352,317
Capital Costs (mm)	1,137	1,137	1,137	1,137
NPV Operating Costs (mm)	5,963	5,963	5,963	5,963
NPV Total Costs (mm)	7,101	7,101	7,101	7,101
NPV Net Increase GDP (mm)	56,658	72,878	209,921	345,217
Annual Net Increase GDP	3.08%	3.96%	11.41%	18.77%
Panel B: Land Value Capture				
LVC Band Revenue (mm)		613		763
As share of capital costs		53.91%		67.08%
LVC CMA Revenue (mm)		1,453		1,718
As share of capital costs		127.78%		151.08%

Note: All numbers in millions of 2016 USD. NPV calculate over a 50 year time horizon with a 5% discount rate. Each column describes to a different model. Row (1) reports the increase in NPV GDP from phases 1 and 2 of the TransMilenio network from the baseline equilibrium in 2012 (calculated as the fall in GDP from its removal). Row (2) reports the capital costs of constructing the system, averaging 12.23mm per km over 93km of lines. Row (3) reports the NPV of operating costs, defined conservatively as farebox revenue in 2012. Row (4) reports the NPV of total costs, while row (5) reports the difference between row (1) and row (4). Row (6) reports this difference as a percent of the NPV of GDP in 2012. Row (7) reports the government revenue from the distance band-based land value capture scheme as described in the text, while row (8) reports this as a percentage of capital costs. Rows (9) and (10) report the same figures for the commuter market access-based LVC scheme.

Table 12: Effect of TransMilenio: Decomposing the Channels

	Model 1: Same $\eta, \theta, \sigma_L \to \infty$			**Model 2:** Diff $\theta, \sigma_L \to \infty$			**Model 3:** Diff $\eta, \theta, \sigma_L \to \infty$			**Model 4:** Full Model		
	Low	High	Ineq.	Low	High	Ineq.	Low	High	Ineq.	Low	High	Ineq.
Partial Eqbm	6.169	4.734	-1.529	6.002	5.479	-0.556	6.029	5.417	-0.651	5.797	5.596	-0.214
Rents Adjust (no subsistence)	4.085	3.140	-0.986	3.901	3.708	-0.200	3.916	3.668	-0.258	3.758	3.817	0.062
Rents Adjust	4.079	3.091	-1.030	3.852	3.682	-0.177	3.870	3.637	-0.242	3.702	3.794	0.096
Rents & Wages Adjust	4.083	2.990	-1.140	3.841	3.640	-0.208	3.856	3.600	-0.267	3.637	3.837	0.208
Emp Decisions Adjust	4.073	2.866	-1.259	3.820	3.569	-0.261	3.836	3.528	-0.320	3.737	3.681	-0.058
Emp & Car Decisions Adjust	4.125	2.950	-1.226	3.886	3.663	-0.232	3.909	3.630	-0.290	3.710	3.877	0.173
All Decisions Adjust	4.210	3.364	-0.883	3.982	4.089	0.112	3.992	4.074	0.086	3.741	4.407	0.692
Res Spillovers	4.470	3.634	-0.875	4.243	4.254	0.011	4.255	4.241	-0.014	3.939	4.636	0.726
Res & Prod Spillovers	5.080	4.076	-1.058	4.819	4.673	-0.153	4.999	4.760	-0.252	4.606	5.182	0.604

Note: Table shows the (negative of the) value of the percentage change in welfare from removing phases 1 and 2 of the TransMilenio network from the 2012 equilibrium across different models. For each model, the first column reports the percentage change in low-skill worker welfare, the second column reports the percentage change in high-skill worker welfare, and the third column reports reports the percentage change in welfare inequality (defined as the ratio of high-skilled to low-skilled welfare). In model 1, both worker groups are assigned the same (average) η and θ parameters and are assumed to be perfect substitutes in production (i.e. $\sigma_L \to \infty$). In model 2, worker groups differ by their estimated θ parameters. In model 3, worker groups differ both by their estimated θ and η parameters. In model 4, workers differ both by their estimates θ and η parameters and are imperfectly substitutable within firms in the way described in the text. Each row corresponds to the welfare change from the observed equilibrium allowing different margins of adjustment. In row (1), location decisions and prices are held fixed and spillover parameters are set to zero. In row (2), rents are allowed to adjust where the subsistence housing requirement is set to zero in the observed and counterfactual equilibrium. In row (3), rents adjust using the estimated value of the subsistence housing requirement. In row (4), rents and wages are allowed to adjust but all location decisions continue to be held constant. In row (5), in addition to prices changing worker employment location decisions adjust. In row (6), employment and car decisions adjust. In row (7), all prices and location decisions (employment, car ownership and residential location) adjust but spillovers remain shut down. In row (8) residential spillovers are set to their estimated values. Finally, in row (9) both residential and productivity spillovers are set to their estimated values.

Table 15: Effect of Adjusting Housing Supply, and Land Value Capture Scheme

Panel A: Closed City

	Output	Housing	Rents	Welfare Low	Welfare High	Gvt Revenue
Fixed Supply	4.270		3.706	4.828	5.466	
Free Adjustment	4.419	1.600	1.525	6.045	6.392	
LVC, Bands	4.650	1.397	2.208	5.258	5.923	0.670
LVC, CMA	4.570	2.468	0.568	6.333	6.674	0.952

Panel B: Open City

	Output	Housing	Rents	Pop. Low	Pop. High	Gvt Revenue
Fixed Supply	23.311		23.645	15.533	18.111	
Free Adjustment	53.713	17.914	25.387	38.282	40.813	
LVC, Bands	30.903	5.442	23.174	20.222	23.624	1.588
LVC, CMA	37.308	7.629	25.569	26.472	28.895	2.114

Note: Table shows the percentage change in each outcome going from the equilibrium without TransMilenio to that with TransMilenio under the housing supply conditions indicated in each row. Row (1) is the case with fixed housing supply. Row (2) is the case of freely adjusting housing supply. Row (3) is the distance-band based land value capture (LVC) scheme, where the government sells rights to construct up to 30% new floorspace in tracts closer than 750m from stations. Government revenue from the scheme is given in column (6) as a percentage of pre-TM GDP. Row (4) shows the results of the scheme based on predicted changes in commuter market access as described in the text.

Table 14: Effect of Different System Components

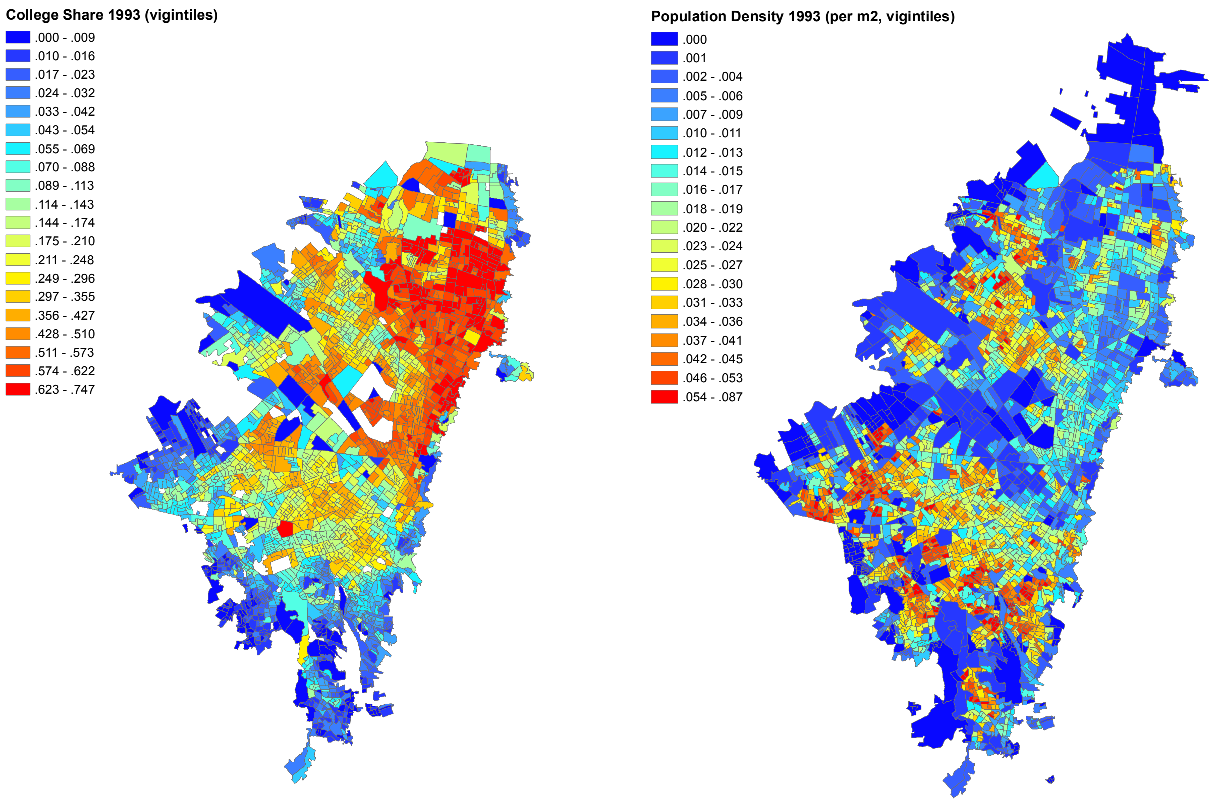
	Add Cable Car			**Remove Line South**			**Remove Line North**			**Remove Feeders**		
	Welfare Low	Welfare High	Output	Welfare Low	Welfare High	Output	Welfare Low	Welfare High	Output	Welfare Low	Welfare High	Output
Partial Eqbm	0.095	0.028		1.956	1.329		4.935	7.412		2.685	2.080	
Rents Adjust	0.066	0.017		1.362	0.929		2.988	5.438		1.836	1.456	
All Prices Adjust	0.063	0.035	0.035	1.334	1.228	1.452	3.704	4.688	5.361	1.747	1.640	1.874
Emp Decisions Adjust	0.063	0.035	0.036	1.361	1.151	1.330	3.769	4.631	5.150	1.774	1.582	1.783
Full Adjust, No Spillovers	0.057	0.051	0.028	1.546	1.626	0.866	4.277	4.983	5.077	1.713	1.927	1.321
Full Adjust, Spillovers	0.088	0.075	0.031	2.040	2.034	0.990	3.366	4.263	6.527	2.333	2.440	1.595

Note: Table shows the (negative of the) value of the percentage change in welfare from removing a piece of the TransMilenio (existing or future) network. These counterfactuals are adding the Cable Car system (columns (1)-(3)), removing line H in the south (columns (4)-(6)), removing line A in the north (columns (7)-(9)), and removing the feeder system. Each row corresponds to allowing different margins to adjust (see previous table description for details).

Figures

Figure 2: Population Density and Demographic Composition in 1993

(a) College Share (b) Population Density

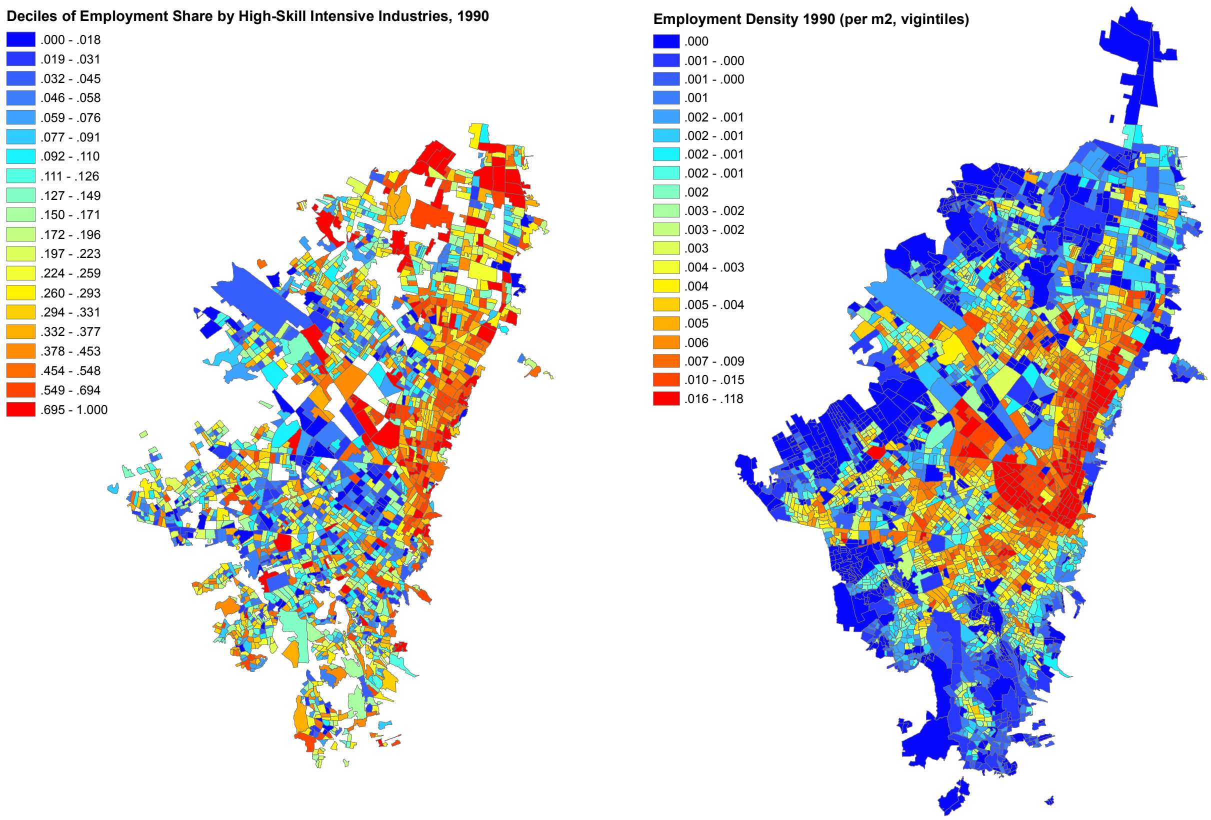


Note: Data is from 1993 Census.

Figure 3: Employment Density and Industry Composition in 1990

(a) High-Skill Industry Share

(b) Employment Density



Note: Data is from 1990 Economic Census. High-skill industries defined in text.

Figure 4: TransMilenio Routes

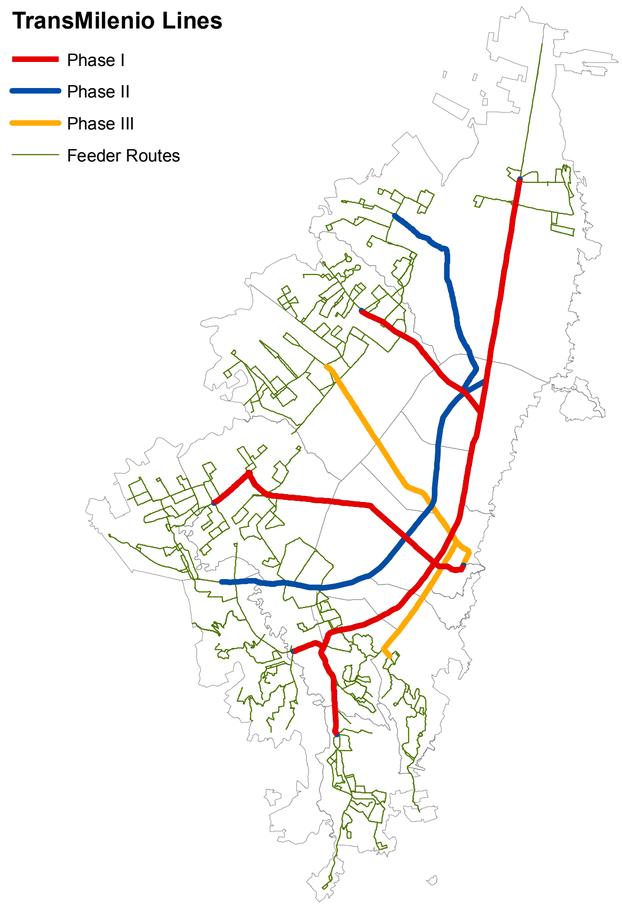
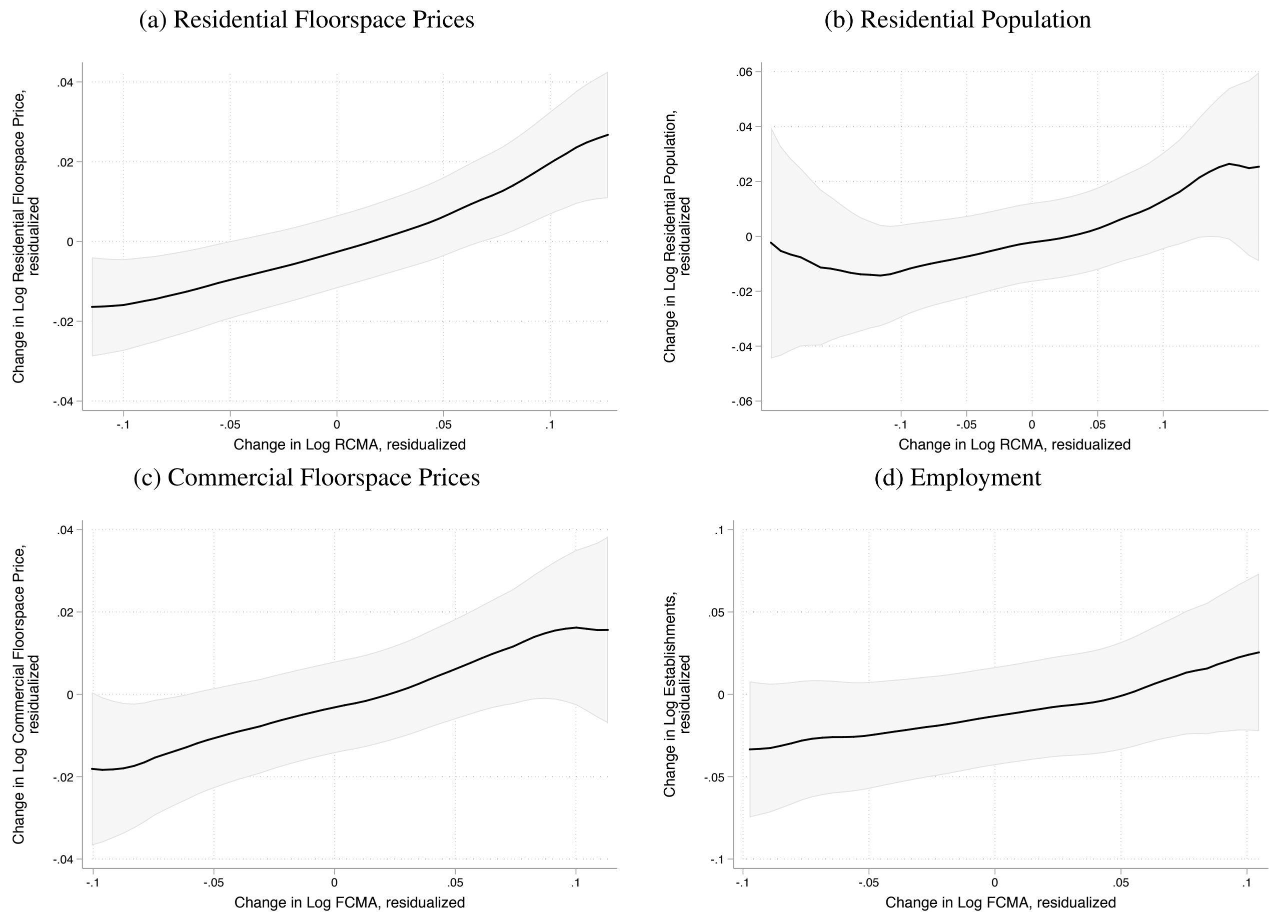
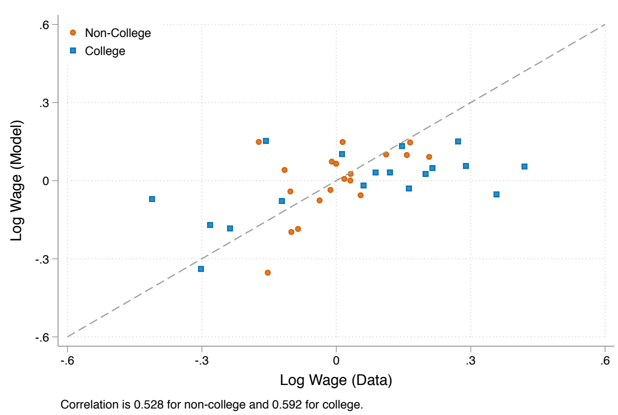


Figure 5: Non-Parametric Relationship Between Outcomes and Commuter Market Access



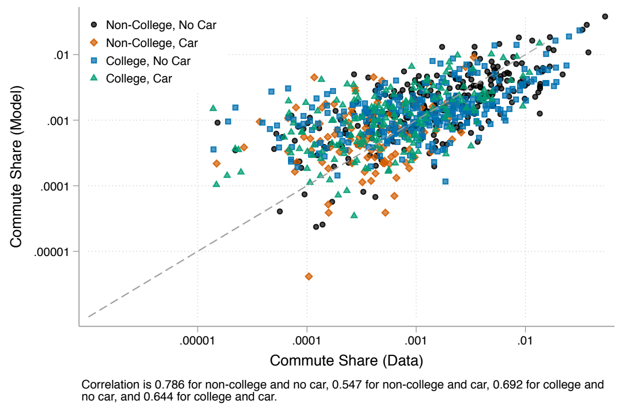
Note: Plot shows the non-parametric relationship between outcomes and commuter market access. Specifications correspond to the reduced form from column (4) of main table in which commuter market access is measured holding population and employment fixed at their initial levels, with the full set of baseline controls included, and is regressed directly on outcomes.

Figure 6: Wages: Model vs. Data



Note: Plot compares the average wage by skill group in each locality as predicted by the model with that observed in the GEIH data (not used in estimation).

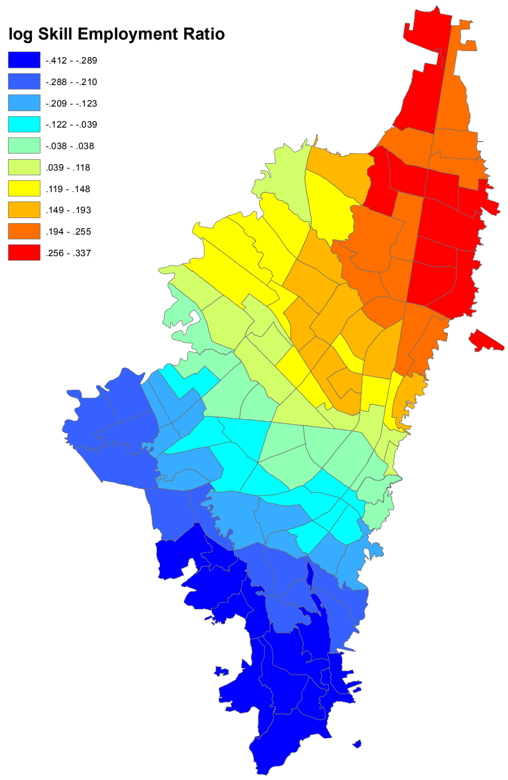
Figure 7: Commute Flows: Model vs. Data



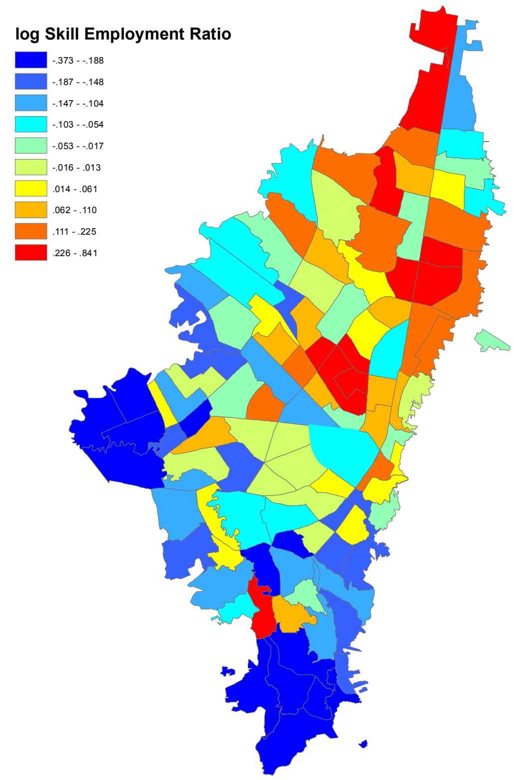
Note: Observation is a locality origin-destination pair, skill group and car ownership combination. Plots shows relationship between share of commuters choosing each (i, j, a) pair in the model vs those doing so in the 2015 Mobility Survey.

Figure 8: Relative Employment by Skill by UPZ: Model vs Data

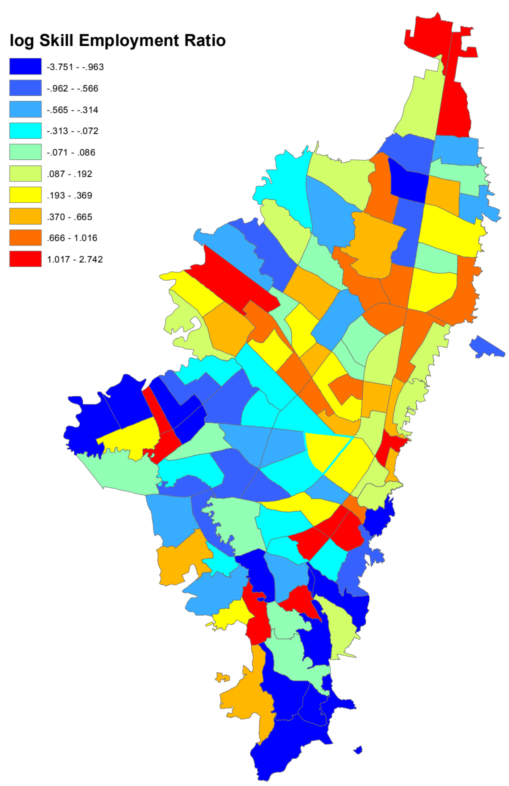
(a) Model: Perfect Subsitutes & Same θ



(b) Model: Baseline Estimates



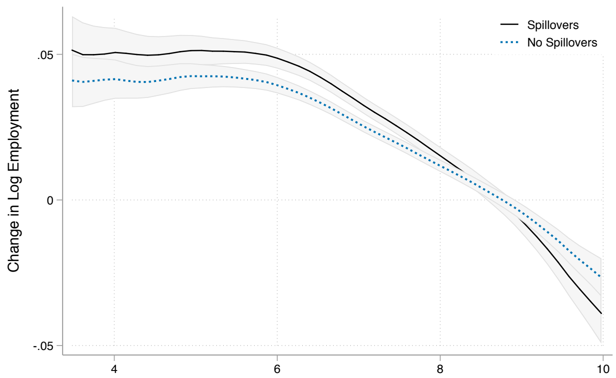
(c) Data



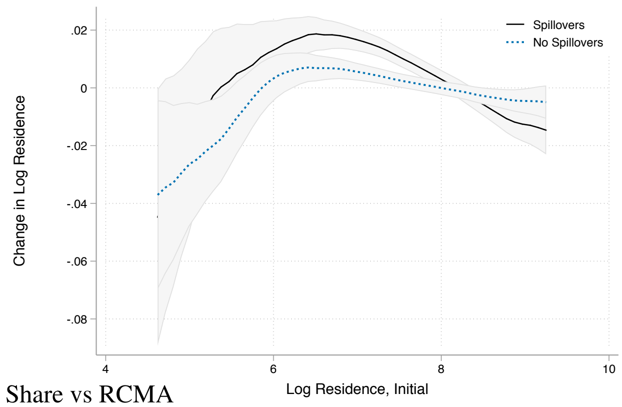
Note: Panel (a) shows the deciles of the distribution of the log skill employment ratio $\ln L_{FjH}/L_{FjL}$ by UPZ in the model when skill groups are perfect substitutes in production and have the same value of θ (equal to the average value in the population. Panel (b) shows the distribution for the baseline model. Panel (c) shows the distribution in the 2015 Mobility Survey. Corr between data in panel (a) and (c) is 0.256, while that between panel (b) and (c) is 0.406.

Figure 9: Simulated Changes in Outcomes

(a) Employment



(b) Residential Population



(c) Change in College Share vs RCMA