Slum growth in Brazilian Cities

Guillermo Alves
PhD Candidate
Brown University

Working Paper

October, 2016

*This work is part of my thesis to obtain the degree of PhD in Economics at Brown University. I thank Mathew Turner, Nathaniel Baum-Snow, Vernon Henderson, Jesse Shapiro, Andrew Foster, Aurora Ramírez, Heitor Pellegrina, Lucas Scottini, Fernando Álvarez, Sebastián Fleites, and participants at IEB/UEA 2016 Summer School in Urban Economics, 2016 CAF Research Seminar at Di Tella University, Iecon-Udelar seminar, and Brown University Development and Applied Micro breakfasts and Applied Micro lunch seminar for helpful suggestions and conversations. This project received financial support from CAF-Development Bank of Latin America research funds on topics of "Habitat and Urban Development in Latin America". Any mistakes are mine.

Abstract

I study slum growth in contemporary urbanization processes by estimating a spatial equilibrium model with houses with and without basic water and sanitation services in Brazilian cities between 1991 and 2010. Slum growth results from households moving to cities following higher wages (elasticity of 1.7), this movement impacting cities' serviced housing rents much more (elasticity of 0.4) than unserviced ones (elasticity of 0.1), and these rent changes impacting households' location decisions more for serviced (elasticity of -0.5) than for unserviced houses (elasticity of -0.4). I show that the effect of urban economic growth on cities' slum incidence depends critically on what happens in other cities. When a few cities grow, they experience higher slum incidence because they are the focus for migrants coming from rural areas and less dynamic cities. When all cities grow, slum incidence declines in all cities as a result of two forces. First, each individual city faces less housing demand pressure as migration between cities becomes more balanced and rural migrants flow to all cities. Second, generalized economic growth improves households' incomes nation-wide, allowing households to switch to higher quality non-slum housing. In terms of common slum policies, I show that the effects of slum repression on any individual city are mild and decrease with the number of other cities repressing slums. If all cities repress slums by making unserviced housing 20% more expensive, this lowers aggregate urbanization by 0.4% and low income households' welfare by 1.1%. On the other hand, a generalized slum upgrading policy turning 10% of cities' 1991 unserviced housing stock into serviced housing, increases aggregate urbanization by 1.1%, low income households' welfare by 4.0%, and high income households' welfare by 3.6%.

Keywords: Urbanization, Slums, Brazil.

JEL Codes: O18, R13, R31

1 Introduction

One third of developing world's urban households live with either no adequate water, sanitation, durable construction materials, or experience overcrowding (UN, 2012). With rapid urbanization taking place across the developing world, the number of urban households in those living conditions keeps increasing, having reached 880 million in 2010 (UN, 2015). Local and national governments invest heavily in policies dealing with slum growth, ranging from strong repression, including evictions, to slum upgrading programs (UN, 2003).[1] In contexts of high spatial mobility of households, both between rural areas and cities and between different cities, slum policies implemented in a few cities may affect slum growth in other cities as well as aggregate urbanization rates. These reallocation effects may in turn have significant welfare impacts on different types of households.

This paper estimates a spatial equilibrium model of a system of cities with slum and non-slum houses to look at the effects of changes in cities' slum policies and economic fundamentals on households' spatial allocation and welfare. I estimate the model by looking at the growth in the number of houses with and without basic water and sanitation in Brazilian cities between 1991 and 2010. The basic structure of the model features two elements. First, low and high income households choose between serviced or unserviced houses in a set of cities or living in the countryside. Second, cities provide housing with one supply function for each type of house.

A quick look at the estimated structural elasticities illustrates the basic mechanics of slum growth according to the model. First, as developing world cities become more productive places and offer higher wages, low income households flow rapidly to cities, increasing the demand for urban houses (elasticity of 1.6). Second, these increases in housing demand impact serviced housing rents much more (elasticity of 0.4) than unserviced ones (elasticity of 0.1). Third, low income households react to these changes in relative housing rents by increasingly choosing unserviced houses (price elasticity of -0.4) over serviced ones (price elasticity of -0.5). As a result of the interplay of these elasticities, urban economic growth pushes slum incidence upwards. However, urban growth has another effect operating in the opposite direction, which tends to dominate in the long run as countries become wealthier. This other effect is explained by a steep observed positive gradient between serviced housing consumption and households' income. As cities grow and countries become richer, wealthier households increasingly consume serviced housing over unserviced one, and this pushes slum

[1] For instance, own processing of Brazil's MUNIC survey indicate that around 50% of Brazilian cities in my sample were implementing some type of slum upgrading program in the mid 1990's.

incidence downwards in the long run.[2]

The impact of urban economic dynamism on slum incidence depends then on the balance between those two forces. In particular, I show that the impact of economic growth on slum growth in any given city depends critically on what happens in other cities. For instance, when I simulate a 20% extra wage growth shock between 1991 and 2010 in a single medium-size city, this city's unserviced housing incidence increases by 1.3% and its number of unserviced houses increases by 28.0%. However, when the same extra wage growth takes place in all cities, unserviced housing incidence in that same medium-size city goes down by 3.3% and the number of unserviced houses in that city goes down by 3.5%. Two reasons explain the different outcomes between the two scenarios. First, in a general equilibrium context, a single growing city attracts more migrants when no other city grows than when all cities grow. Second, when all cities grow, the national share of low income households goes down, and this mechanically reduces the demand for unserviced urban houses given the steep gradient referred to above.

I further use the estimated model to study the reallocation and welfare effects of slum repression and slum upgrading policies. An average city repressing slum growth by making the supply of unserviced houses 20% more expensive reduces slum incidence only by 1.5%. Since the average city size is small, slum repression in one city does not have significant effects outside the city. However, if all cities repress slums making unserviced urban housing 20% more expensive everywhere, aggregate urbanization in 2010 goes down by 0.4% and low income households' welfare goes down 1.1%. I implement slum upgrading policies as cities' targeting to turn 10% of their 1991 stock of unserviced houses into serviced ones.[3] I find that this improves low income households' welfare by 4.0% and high income households' welfare by 3.6%. Also, I find that this policy makes cities more attractive such that the 2010 aggregate urbanization rate becomes 1.1% higher.

This paper's methodological approach follows a recent set of works modeling households' location choices with a discrete choice model and featuring log-linear functions to characterize cities' housing supply side (Diamond, 2016; Serrato & Zidar, 2016). I take this literature a step forward by adding two types of houses and by solving for the general equilibrium of the estimated model for a set of counterfactuals.[4] This paper's empirical strategy uses

[2]Several authors have documented how slums were common in today developed countries' cities and slums disappeared as countries developed. See for example World Bank (2009) and Glaeser (2012).

[3]Although I do not know of any systematic international measures of the magnitude of slum upgrading policies, survey evidence I processed for Brazil indicates around 50% of cities in my sample were doing some type of slum upgrading by the end of the 90's.

[4]For instance, although Diamond (2016) features a full general equilibrium model, her model does not have a closed form solution. Since there is a trade-off between writing down a more complex model and

1991 and 2010 Census data on wages, housing prices, and population quantities for a set of 272 Brazilian cities. The main goal of this paper's empirical strategy is to identify the set of elasticities mentioned above. Empirically identifying these parameters suffers from the typical simultaneity problem of estimating supply and demand systems. I tackle this problem by building a set of moment conditions based on Bartik (1991) wage shocks and Card (2001) migration shocks, both computed separately for low and high income households. The Bartik wage shock computed for each type of household shifts the respective labor demand and then identifies how households respond to changes in cities' real wages (wages net of housing rents). In terms of identifying housing rents' responses to housing demand shocks, Bartik wage shocks and Card migration shocks attract more people to cities, thus shifting cities' housing demand and identifying the two housing supply functions.[5]

This paper features several methodological improvements over the relatively scarce literature on the determinants of slum growth (World Bank, 2009; Feler & Henderson, 2011; Marx, Stoker & Suri, 2013; Alves, 2014; Castells-Quintana, 2016; Jedwab, Christiaensen & Gindelsky, 2016). First, the empirical strategy exploits housing quantities and housing rents' census data for the universe of cities in a country with great spatial heterogeneity. This allow me to empirically study slum growth's determinants by looking at how the relative evolution of cities' productivities leads to population reallocation across the space and across the housing quality dimension. Second, the paper features a general equilibrium methodology that looks into cities' dual housing markets, links them with what is happening in rural areas and in other cities, and is capable of aggregating city-level outcomes into national level urbanization and slum incidence statistics.

While there is a long tradition of urban studies adopting this spatial equilibrium perspective for the US system of cities (Rosen, 1979; Roback, 1982; Glaeser, 2008; Hsieh & Moretti, 2015; Diamond, 2016), no studies have used this approach to analyze slum growth in developing countries' systems of cities.[6] By looking at the local and national effects of city-level slum policies, this paper also contributes to the literature on the impacts of place based policies (**?**Kline & Moretti, 2014).

keeping a closed form solution for the model's equilibrium, my model's discrete choice structure is simpler than Diamond's.

[5]The exclusion restriction for these housing demand shifters requires that they affect housing rents only through housing quantities. This might not hold if these instrument affect for instance construction wages. I explicitly test for this and show that the two instruments do not affect local construction wages.

[6]A recent literature uses a spatial equilibrium approach to study developing countries' system of cities (Harari, 2015; Chauvin et al., 2016). Importantly, Chauvin et al. (2016) have recently provided evidence that using standard Rosen-Roback spatial equilibrium tools seems reasonable for the case of Brazil (not for India for instance).

The paper starts with a description of the data in Section 2. Section 3 presents evidence on Brazil's general context between 1991 and 2010, studies Brazilian households' spatial mobility, and documents the empirical relationship between slum growth, wages, and rents. Section 4 introduces the spatial equilibrium model. Section 5 describes the estimation methodology, discussing the main identification concerns and presenting the details of the paper's identification strategy. Section 6 presents the estimation results and Section 7 solves for the general equilibrium of the model for a set of counterfactual scenarios. Section 8 concludes.

2 Data

The paper's empirical exercise looks at 1991-2010 changes in the allocation of two types of households (high and low income) across two types of houses (unserviced and serviced) in 272 Brazilian cities and the countryside. I study the determinants of these changes in households' spatial allocation by looking at three main variables: population quantities, wages, and housing prices. In this section, I describe the construction of these variables. All the data come from own processing of 1991 and 2010 Brazilian censuses' microdata.[7]

I adapt UN's slum definition to the Brazilian context and data availability and define serviced houses as those with both proper water and sanitation services UN (2003). A house has proper water services if it is connected to the local water network, with connection inside the house. Proper sanitation services imply being connected to the local sewerage network or having a septic tank.

I restrict the paper's relevant population to those households with working household heads between 14 and 70 years old. I compute average wages from household heads' earnings in their main occupation and average rents from self-declared monthly rents by renting households.[8] I express both wages and rents in constant 2010 prices.

I classify households in low and high income in order to capture the wide underlying income inequality and the existence of stark differences in serviced housing consumption across the income distribution. I draw the income cutoff at the 75th percentile of the distribution of

[7]The 1991 data is a 25% sample and the 2010 a 10% sample. Although Brazil has censuses approximately every 10 years, I do not use the 2000 census because it does not have data on housing prices.

[8]Because the data are samples of the total population, average rents by type of housing get noisy as the number of renting households with or without services in some small municipalities gets very small. For instance, 34 cities have less than 30 observations of renting households without services in 1991 and 63 in 2010. 36 cities have less than 30 observations of renting households with services in 1991 and 6 in 2010.

wages.[9] Figure 1 plots the 1991 gradient of serviced housing consumption and wages and shows how this cutoff defines a rather homogeneous high income group in terms of serviced housing consumption. Very low unserviced housing incidence among high income households in Figure 1 implies that the population share of high income households in unserviced houses in many cities is too small to allow for any empirical study of its determinants.[10] I then assume through the paper that high income households only live in serviced houses.[11]

I define an urban universe of 272 cities formed by Brazil's 66 official metro areas plus those municipalities that are not part of any metro area but had an urban population of at least 50,000 people in 2010.[12] Throughout the paper, I refer to those 272 cities as Brazil's urban areas and to everything else outside those cities as an homogeneous rural area.[13] Because Brazilian municipalities' borders changed during this period, I use Ipums constant geographies to get time-coherent spatial definitions.

I use two additional pieces of data from Brazilian censuses to construct the set of instrumental variables. First, I construct a 1991-2010 compatible classification of 169 industries for household heads' main occupation to build the Bartik instrument. Second, I use household heads' municipality of residence five years before each census to compute the migration instrument based on Card (2001).

3 Background

As noted in the Introduction, this paper uses a spatial equilibrium approach to study slum growth with cities' wages and housing rents playing a central role in understanding slum dynamics. This section starts by describing the general economic context of Brazil between 1991 and 2010 and then turns to presenting evidence on two key aspects supporting the

[9]This 75th percentile corresponds to a gross monthly wage of 1,140 Reais measured in 2010 prices, approximately 650 US dollars.

[10]The median share of high income households in unserviced houses in a given city (with respect to the national population of high income households) is 0.0085%.

[11]Besides the empirical reasoning behind this assumption, a more conceptual reason is that, given the high prevailing inequality, it seems safe to assume that those high income households living in unserviced houses according to the Census have unobservable ways of dealing with the disamenities of unserviced houses.

[12]Metro areas in Brazil are defined as a set of municipalities. I take the 2010 definition of metro areas. Municipalities include rural and urban population. I then further use the censuses' classification of households as urban or rural to identify the urban population within each municipality. The criteria described above yields a total of 278 cities. I lose 6 cities that do not have any observation for serviced housing prices in 1991. I remove those 6 cities from the sample.

[13]For instance, the rural population is the one: living in municipalities not included in any official metro area and with an urban population less than 50,000 in 2010 plus those households living in rural areas within the municipalities belonging to some metro area or having more than 50,000 people in 2010.

paper's methodological approach. First, I look at evidence of households' spatial mobility and of spatial equilibrium notions being relevant for Brazil's system of cities.[14] Second, I explore the role of wages and housing rents for understanding unserviced housing growth in Brazil.

3.1 Brazil between 1991 and 2010

Brazil is an early urbanizer by the historical standards of today's developed countries (Chauvin, Glaeser, Ma & Tobio, 2016). This feature, shared by most Latin American countries, is well illustrated by Brazil being today more urbanized than the US, although Brazil has less than half of US per capita GDP. Table 1 presents some basic descriptive statistics for Brazil between 1991 and 2010. The general picture is of moderate but generalized progress: per capita GDP grew 41%, inequality measured by the Gini index went down slightly by 3 points, and typical welfare indicators such as infant mortality and illiteracy improved notably.[15] The number of urban households without basic water and sanitation in Table 1 grew from 5.3 million in 1991 to 5.7 million in 2010. This growth rate was well below total urban population growth rate, leading to an overall reduction in the share of urban unserviced houses, which fell from 32.7% to 23.4% among urban working households.

The lower half of Table 1 shows population and wage growth rates for low and high income households. Average wages grew more for the low than for the high income group. As a result, a composition effect made the number of households in the high income group grow by 56.9%, compared to 11.8% for the low income group. Given the steep gradient between serviced housing consumption and income we saw in Figure 1, these changes in households' composition by income play a key role in explaining the observed reduction in unserviced housing incidence in Brazil.

Brazil's economy experienced two big nation-wide productive shocks between 1991 and 2010: trade liberalization in the 1990's and commodities boom in the 2000's.[16] These two shocks impacted the allocation of resources both across industries and across the space, favoring the

[14] This aspect has been recently explored by Chauvin, Glaeser, Ma & Tobio (2016).

[15] Successful macroeconomic stabilization plans in the 1990's coupled with the expansion of public education, public health, and monetary transfers are usually identified as the main drivers of the improvement in Brazil's welfare indicators in this period (Lustig, Lopez-Calva & Ortiz-Juarez, 2013).

[16] In the early 1990's the Brazilian economy went through a process of commercial liberalization that brought down tariffs differentially for different economic sectors, with manufacturing in particular facing big cuts in commercial protection. A set of works have studied the spatially heterogeneous impacts of this opening process on local economies' wage and employment growth (Kovak, 2013; Dix-Carneiro, 2014; Dix-Carneiro & Kovak, 2015). The commodity boom of the late 2000's reinforced the 'pro-primary sector' impact of 1990's trade liberalization policies.

expansion of cities close to the agricultural frontier in the Midwest and North regions of the country, as well as of a few other cities scattered around the country linked to mineral and oil extraction. Figure 2 allows for a full visualization of this spatial variation with a heat map of Brazilian cities' population growth rates. The map shows high growth rates for cities close to the west and northern borders of the country as well as high variation in growth rates within all regions. In order to better appreciate these growth rates' magnitudes, Figure 11 plots them on a histogram. Heterogeneity in population dynamics shows up in Figure 11 with very few cities keeping their size roughly constant and several cities more than doubling their size. These positive urban growth rates contrast heavily with the slightly below-zero growth rate in the number of rural working households (vertical red line to the left in the graph).

3.2 Households' spatial mobility

A key element of the paper's methodological approach is to use spatial equilibrium tools to study slum growth in Brazil's system of cities. Chauvin, Glaeser, Ma & Tobio (2016) have recently argued that several pieces of data point to standard spatial equilibrium tools being relevant for Brazil. A first point these authors make is that Brazil's migration data depicts a country with high internal mobility.[17] Table 2 looks at this by showing that more than half of household heads living in cities in 2010 were not born in the municipality where they currently live.[18] Also, Table 2 shows that although most migrant household heads were born in rural areas, around 10% of them came from another city, which points to the relevance of population flows between cities to explain cities' growth.

A second sign of spatial equilibrium being relevant is that wage and rent growth are positively correlated. This is another point made by Chauvin et al. (2016) and points to the classic Rosen-Roback idea that cities with higher wages must exhibit higher housing prices such that households remain indifferent between cities. Figure 4 confirms this spatial equilibrium sign for Brazil by plotting average wage and rent growth for the 272 cities between 1991 and 2010 and shows a strong positive correlation of 0.54.[19]

[17]In contrast, international migration is relatively limited.

[18]This ratio is a bit higher for households in unserviced houses than for serviced ones.

[19]Chauvin et al (2016) look at this correlation in the cross section. I am looking at it in terms of growth rates.

3.3 Slum growth, wages, and rents

Wages and rents are the two main pecuniary incentives that households face when choosing between living in different cities and play a central role in the formal framework of the paper. Figure 5 looks at the relationship between slum growth, wages, and rents by plotting unserviced housing growth rates against average wage and average rent growth. A first thing to note in Figure 5 is that there is great variation in cities' unserviced housing growth rates. This variation is essential for empirically disentangling the determinants of slum growth.

Figure 5 shows strong positive correlations of unserviced housing growth with both wage and rent growth. I further explore these correlations by running OLS regressions of cities' unserviced housing growth on wage and rent growth. Regression results in Table 3 show that the positive correlation between unserviced housing growth and wage and rent growth remains positive and strong when controlling by either rent of wage growth and by changes in Gini index and by initial population size. These patterns fit common narratives referring to slum growth as taking place in booming cities and cities with tight housing markets. The paper formal structure will clarify the role that cities' economic dynamism and housing rents play in explaining slum growth.

4 Conceptual framework

This section presents the paper's formal structure to study slum growth in a system of cities. The model serves three main purposes. First, it provides a set of estimatable equations characterizing households' location decisions and cities' housing supply capacities. Second, it allows for a formal discussion of those equations' main identification concerns and how the paper's identification strategy deals with them. Third, the model features a closed form solution which allows to study the general equilibrium effects of a set of counterfactuals on changes over time in households' spatial allocation and welfare.

The basic structure of the model follows the recent work by Diamond (2016) with the key extension of allowing for general equilibrium computation. In the model, households' location decisions follow a discrete choice multinomial logit formulation in the spirit of McFadden (1973).[20] These decisions depend on observed wages and rents, and on unobserved type of house and city amenities. The model's housing supply side features a specific housing supply function for each type of house. Cities' production sector features two infinitely elastic housing

[20] A series of recent works have used discrete choice multinomial structures to model households' location choices in system of cities Serrato & Zidar (2016); Morten & Oliveira (2016); Diamond (2016).

demands with low and high income households' wages following city-specific productivity shocks. The model is static and I interpret it as capturing the long run equilibrium of Brazil's system of cities. In particular, I take the model to the data by assuming that 1991 and 2010 are two different long run equilibria of the model.

4.1 Households' location decisions

Each household i is either of low L or high H income. In order to simplify the exposition, I first present households' location decision structure for a generic household and then indicate which aspects of this structure differ between low and high income types.

At each time t, a fixed number of households $\overline{N}_t$ chooses to live in serviced or unserviced housing $m \in \{u, s\}$ at some city j or in the countryside c. Denote this set of alternatives as O. At each specific location choice, households maximize a Cobb Douglas utility defined over a composite non-housing good X, housing Z_m, and a city and type of house specific amenity A_{imjt}. Households face a budget constraint given by wages W_{jt}, housing prices P_{mjt}, and by the price of the non-housing good being normalized to 1. The maximization problem for each location is then:

$$\begin{aligned} \max_{Z_m, X} \quad & \alpha_m \ln Z_m + (1-\alpha_m)\ln X + \ln A_{imjt} \\ & s.t. \ P_{mjt} Z_m + X = W_{jt} \end{aligned} \tag{1}$$

With lower cases denoting variables in logarithms, the indirect utility function V_{imjt} for a household i from choice m, j is:

$$V_{imjt} = w_{jt} - \alpha_m p_{mjt} + a_{imjt} \tag{2}$$

Since only differences in indirect utility between alternatives matter for households' location choices, I normalize indirect utility in the countryside to zero.[21]

City and type of house amenities a_{imjt} have a generic component common to all households $\bar{a}_{mjt}$ and a household-specific shock ε_{imjt} such that $a_{imjt} = \bar{a}_{mjt} + \varepsilon_{imjt}$, with ε_{imjt} distributed iid type I Extreme Value with dispersion parameter σ. This dispersion parameter measures how much real wages and amenities matter for households' location decisions. For instance, high dispersion in idiosyncratic preferences implies that households do not react much to

[21] This normalization is also necessary for performing Berry (1994) procedure which allows to estimate the demand system with linear regression equations. I discuss estimation in the next Section.

changes in wages and housing rents.

Households choose the type of house and city with the highest indirect utility in the set of all alternatives. Define the component of indirect utility which is common to all households as v_{mjt}, such that $V_{imjt} = v_{mjt} + \varepsilon_{imjt}$. Then, the number of households N_{mjt} in type of house m and city j at time t is:

$$N_{mjt} = P(V_{imjt} = \max_{O}\{V_{imjt}\}) = \sum_{i \in \overline{N}_t} \frac{exp(v_{mjt}/\sigma)}{\sum_O exp(v_{mjt}/\sigma)} \tag{3}$$

Low and high income households largely share this common choice structure but differ in two things. First, as explained in Section 3, I assume that high income households do not live in unserviced houses. Then, their choice set O^H consists only of urban serviced houses or the countryside. Second, low and high income households may have different population sizes $\overline{N}^L, \overline{N}^H$, taste parameters $\alpha_m^L, \alpha_s^H, \sigma^L, \sigma^H$, earn different wages w_{jt}^L, w_{jt}^H, and have different values for type of house and city amenities $\bar{a}_{mjt}^L, \bar{a}_{sjt}^H$.

4.2 Production

I model cities' labor demand side following **?**. Two types of firms operate in perfectly competitive factor and output markets with constant returns to scale Cobb-Douglas production technology using labor and capital and a city-specific productivity term. One type of firm uses only low income labor and the other type of firm uses only high income labor. Labor shares for each of the two types of firms are β^L, β^H and city-specific productivities are $\theta_{jt}^L, \theta_{jt}^H$. The supply of capital is infinitely elastic at a national interest rate i_t and output prices P_t^L, P_t^H are exogenous and set nationally.

Under those assumptions, profit maximization yields a perfectly elastic labor demand for each type of household such that wages are a function of city-specific productivities and a constant C_t capturing national level factors, including the national interest rate and output prices. The labor demand expressed in logarithms is:

$$w_{jt}^L = C_t^L + \frac{1}{\beta^L}\theta_{jt}^L \tag{4}$$

$$w_{jt}^H = C_t^H + \frac{1}{\beta^H}\theta_{jt}^H \tag{5}$$

The assumptions above imply then that L and H types face completely separated labor markets. This rather extreme assumption seems reasonable for the context of extreme inequality of Brazilian cities. For instance, the correlation between both types of households' wage growth between 1991 and 2010 is -0.05.

4.3 Housing Market

Competitive firms produce the two types of housing such that the price of each type of house equals its marginal cost. Housing costs have two components: land and construction costs. Land costs LC_{mjt} increase with the amount of each type of housing supplied Z^S_{mjt}. I allow the land cost gradient dLC_{mjt}/dZ^S_{mjt} to differ between serviced and unserviced houses and assume that all housing rents go to absentee landlords. Construction costs C_{mjt} might also differ between types of houses. I parametrize the corresponding inverse housing supply function for each type of house as:

$$\ln P_{mjt} = \gamma_h \ln Z^S_{mjt} + \ln C_{mjt} \tag{6}$$

The demand for each type of housing in each city Z^D_{mjt} depends on how many households decide to live in each city and type of house (extensive margin) and how much housing each household consumes (intensive margin). The extensive margin is given by the number of households choosing each city and type of house (Equatino 3) and the intensive margin comes from households' Cobb-Douglas optimization problem above. Then, the two housing demands are:

$$Z^D_{ujt} = N^L_{ujt} \frac{\alpha^L_u W^L_{jt}}{P_{ujt}} \tag{7}$$

$$Z^D_{sjt} = N^L_{sjt} \frac{\alpha^L_s W^L_{jt}}{P_{sjt}} + N^H_{sjt} \frac{\alpha^H_s W^H_{jt}}{P_{sjt}} \tag{8}$$

4.4 Equilibrium

The system of cities' equilibrium is defined by an allocation of the country's population between types of houses in cities and the countryside $(N^{*L}_{mjt}, N^{*L}_{ct}, N^{*H}_{sjt}, N^{*H}_{ct})$ and by a vector of housing prices and wages $(P^*_{ujt}, P^*_{sjt}, W^{*L}_{jt}, W^{*H}_{jt})$ such that each city's labor and housing markets are in equilibrium and all households live somewhere.

Formally, the labor market supply for each type of household is defined by Equation 3. This labor supply must equal the labor market demands defined by Equations 4 and 5 in each city:

$$N_{mjt}^{*L} = \sum_{i \in \overline{N_t^L}} \frac{exp((w_{jt}^{*L} - \alpha_m^L p_{mjt}^* + \bar{a}_{mjt}^L)/\sigma^L)}{\sum_{O^L} exp((w_{jt}^{*L} - \alpha_m^L p_{mjt}^* + \bar{a}_{mjt}^L)/\sigma^L)} \tag{9}$$

$$N_{sjt}^{*H} = \sum_{i \in \overline{N_t^H}} \frac{exp((w_{jt}^{*H} - \alpha_s^H p_{sjt}^* + \bar{a}_{sjt}^H)/\sigma^H)}{\sum_{O^H} exp((w_{jt}^{*H} - \alpha_s^H p_{sjt}^* + \bar{a}_{sjt}^H)/\sigma^H)} \tag{10}$$

$$w_{jt}^{*L} = C_t^L + \frac{1}{\beta^L}\theta_{jt}^L \tag{11}$$

$$w_{jt}^{*H} = C_t^H + \frac{1}{\beta^H}\theta_{jt}^H \tag{12}$$

Also, the housing supplies for each type of house and city must equal the respective housing demands:

$$\ln P_{mjt}^* = \gamma_m \ln Z_{mjt}^{*D} + \ln C_{mjt}, \tag{13}$$

$$Z_{ujt}^{*D} = N_{ujt}^{*L} \frac{\alpha_u^L W_{jt}^{*L}}{P_{ujt}^*} \tag{14}$$

$$Z_{sjt}^{*D} = N_{sjt}^{*L} \frac{\alpha_s^L W_{jt}^{*L}}{P_{sjt}^*} + N_{sjt}^{*H} \frac{\alpha_s^H W_{jt}^{*H}}{P_{sjt}^*} \tag{15}$$

Finally, all households must be somewhere:

$$N_{ct}^{*L} + \sum_{O^L} N_{mjt}^{*L} = \overline{N}_t^L \tag{16}$$

$$N_{ct}^{*H} + \sum_{O^H} N_{sjt}^{*H} = \overline{N}_t^H \tag{17}$$

The equilibrium defined by Equations 9 to 17 is non-linear and does not feature a closed form solution. In order to study the general equilibrium effects of changes in slum policies and economic fundamentals, in Section 7 I solve for a first differences' version of the model which features a closed form solution.

5 Estimation strategy

In this section I describe how I estimate the set of structural parameters $(\sigma^H, \sigma^L, \gamma_U, \gamma_S)$. This estimation, together with calibration of the housing expenditure shares' parameters (α_m^L, α_s^H), fully characterizes the equilibrium of the model above. The estimation strategy consists in constructing a set of first-time differences' linear equations from the model above and estimate them with IV techniques.

For exposition purposes, I divide the set of estimating equations in two groups. The first group of equations estimates σ^L and σ^H following Berry (1994)'s method for estimating the discrete choice problem above with a set of linear regression equations. Second, the estimation of the housing supply parameters γ_U and γ_S regresses, for each type of house, 1991-2010 changes in housing rents on changes in the respective housing demands.

Estimating the set of parameters above suffers from the typical simultaneity problem of estimating supply and demand systems. In this section, I specifically discuss these identification concerns for each of the estimating equations and describe how I use a set of instruments to deal with those concerns. I start this section by introducing the set of instruments and, in the second part of the section, I derive the estimating equations and discuss the endogeneity concerns and identification assumptions.

5.1 Instruments

5.1.1 Bartik Wage Instruments

Bartik (1991) wage instruments predict cities' wage growth by interacting 1991-2010 national wage growth rates by industry with cities' 1991 industrial employment composition. I compute three versions of this instrument: one for low income workers ΔB_j^L, one for high income workers ΔB_j^H, and one considering all workers in each city ΔB_j. These Bartik wage instruments fit the conceptual framework outlined in Section 4 by acting as proxies for changes in the local productivity shocks $\theta_{jt}^L, \theta_{jt}^H$ in Equations 4 and 5 (Bartik, 1991; Bound & Holzer, 2000; Notowidigdo, 2011; Diamond, 2016).

Formally, let $W_{ind,j,t}$ denote the average wage paid by industry ind in city j at time t and $W_{ind,-j,t}$ denote the average wage paid by that industry in all other cities except j. With these definitions in mind, the three Bartik instruments are:

$$\Delta B_j = \sum_{ind} (\ln W_{ind,-j,2010} - \ln W_{ind,-j,1991}) \frac{N_{ind,j,1991}}{N_{j,1991}} \tag{18}$$

$$\Delta B_j^L = \sum_{ind} (\ln W^L_{ind,-j,2010} - \ln W^L_{ind,-j,1991}) \frac{N^L_{ind,j,1991}}{N^L_{j,1991}} \tag{19}$$

$$\Delta B_j^H = \sum_{ind} (\ln W^H_{ind,-j,2010} - \ln W^H_{ind,-j,1991}) \frac{N^H_{ind,j,1991}}{N^H_{j,1991}} \tag{20}$$

Note that when computing the Bartik wage for a given city, national average wage growth rates by industry do not include wage growth in that city. Otherwise, local shocks affecting labor supply in a city could trivially affect the instruments and invalidate their interpretation as proxies for changes in local labor demand.

Table 4 presents the first stages for Bartik instruments predicting actual wage growth. The first two columns show strong first stages for each type of household Bartik predicting each type of household wage growth. Those two columns in Table 4 also show how Bartik for one type of household does not predict actual wages for the other type of household. This is additional evidence in favor of the assumption on labor markets for low and high types being fairly independent.

5.1.2 Migration Instruments

The migration instrument proposed by Card (2001) is based on the evidence that migration networks matter for migration decisions and thus previous migration flows can be used to predict future flows.[22]

Denote the number of migrants from origin k to destiny j at time t as $M_{k,j,t}$ and the number of migrants from origin k to destinies other than j as $M_{k,-j,t}$. Then, the idea behind the instrument is that $M_{k,j,t}$ is affected by migration influxes from k to j which occurred before t. For instance, if a given city j had a big share of migrants from a given origin k in 1991 and that origin is, for whatever reason, generating substantial out-migration between 1991 and 2010, then the city j will get a positive migration shock.

[22] See for example Munshi (2003).

Analytically, defining the set of all possible origins as K and the set of all possible destinies as J (i.e. the 272 cities in the sample), the pair of instruments $(\Delta M_j^L, \Delta M_j^H)$ is:[23]

$$\Delta M_j^L = \ln\left(\sum_{k\in K} M^L_{k,-j,2010}\frac{M_{k,j,1991}}{\sum_{j\in J} M_{k,j,1991}}\right) - \ln\sum_{k\in K} M^L_{k,j,1991} \tag{21}$$

$$\Delta M_j^H = \ln\left(\sum_{k\in K} M^H_{k,-j,2010}\frac{M_{k,j,1991}}{\sum_{j\in J} M_{k,j,1991}}\right) - \ln\sum_{k\in K} M^H_{k,j,1991} \tag{22}$$

The instrument uses the 1991 distribution of migrants by place of origin and destiny and the total outflow of migrants by place of origin in 2010 to predict the 2010 number of migrants for each destiny j. Based on the available migration data in the 1991 and 2010 Brazilian censuses, I define the set of origins K as the municipality of residence 5 years before each census. Note that the procedure excludes flows from k to j when computing the total out-migration flows used to predict the 2010 number of migrants in j. I do this to prevent local labor market conditions from affecting the instrument.

In terms of the identification strategy of the paper, these migration shocks bring more people to cities and thus increase the demand for houses. Specifically, I use the migration shock for high income households ΔM_j^H to predict growth in serviced housing demand. Column 4 of Table 4 regresses serviced houses' growth on ΔM_j^H and shows a positive first stage relationship.

5.2 Location decision equations

The estimation strategy to recover the parameters characterizing households' location decisions follows the procedure developed by Berry (1994). Berry showed how to estimate the discrete choice problem above with linear regressions of cities' population quantities on the components of the log indirect utility function except the extreme value term, which is integrated out. In order to see how Berry's procedure works, start by taking logs on both sides of Equation 3:

$$\ln N_{mjt} = v_{mjt}/\sigma - \ln\sum_{O} exp(v_{mjt}/\sigma) \tag{23}$$

[23] Following Card (2001) the relevant migrant share is not skill specific (note that the quotient does not have L, H notation) based on the intuition that migration networks matter across income groups.

By noting that the indirect utility of the outside option c has been normalized to zero and by doing some simple algebra, the difference between the log population of mj and c is:

$$\ln N_{mjt} - \ln N_{ct} = v_{mjt}/\sigma \tag{24}$$

The next step consists in applying the first time differences operator Δ to both sides of Equation 24 and substituting the indirect utility function by its components (except for the idiosyncratic error term). Since the indirect utility components differ between L and H types, the procedure yields the following two linear equations, one for each type of household:

$$\Delta n^L_{mj} = \Delta n^L_c + \frac{1}{\sigma^L}(\Delta w^L_j - \alpha^L_m \Delta p_{mj} + \Delta \bar{a}^L_{mj}) \tag{25}$$

$$\Delta n^H_{sj} = \Delta n^H_c + \frac{1}{\sigma^H}(\Delta w^H_j - \alpha^H_s \Delta p_{sj} + \Delta \bar{a}^H_{sj}) \tag{26}$$

In these two linear equations, I observe population quantities, wages, and housing rents and I do not observe amenities.[24] As I said above, I calibrate the housing expenditure parameters α^L_m, α^H_m. Then, I estimate these two equations by regressing population growth on real wage growth, with unobserved amenities being the error term.[25]

Because changes in unobserved amenities affect equilibrium housing prices in the model, OLS estimation of Equations 25 and 26 would yield biased estimates for σ^L and σ^H. I then instrument for real wages in each equation with the respective Bartik wage shock (i.e. ΔB^L_j in Equation 25 and ΔB^H_j in Equation 26).[26] The identification assumption is that Bartik shocks are uncorrelated with changes in local amenities and then the corresponding moment conditions are:

$$E(\Delta B^L_j \Delta \bar{a}^L_{mj}) = 0$$

$$E(\Delta B^H_j \Delta \bar{a}^H_{sj}) = 0,$$

[24]Note that the first RHS term is population growth in rural areas. This term does not vary across observations and thus is part of the constant.

[25]Note that the unit of observation in both equation is the location alternative. These are type of house-city combinations for low income households and cities for high income households.

[26]As I discussed before, I associate the Bartik shocks with the productivity shocks $(\Delta\theta^L_j, \Delta\theta^H_j)$ and then the first stage for this instrument is defined by first time differences versions of Equations 4 and 5.

5.3 Housing Supply equations

The two regression equations for estimating γ_U, γ_S come from taking first time differences on the housing market equilibrium Equation 13 for each type of house $m \in \{u, s\}$:

$$\Delta p_{uj} = \gamma_u \Delta z_{uj} + \Delta c_{uj}, \tag{27}$$

$$\Delta p_{sj} = \gamma_s \Delta z_{sj} + \Delta c_{sj}, \tag{28}$$

In these two equations I observe growth in rents and all the components of each housing demand growth.[27] The growth in construction costs is unobserved and makes the residual of the two regression equations. Consistent estimation of these two (inverse) housing supply equations needs housing demand shifters to overcome the standard simultaneity problem. Both Bartik and migration shocks shift cities' housing demands by bringing more people into cities and thus potentially identify the two equations.[28] Specifically, I use the average wage Bartik shock to identify the unserviced supply equation and the migration shock for high income workers to identify the serviced supply equation.[29] The corresponding exclusion restriction states that these housing demand shifters must be uncorrelated with the growth in unobserved construction costs:

$$E(\Delta B_j \Delta c_{uj}) = 0 \ , \ E(\Delta M_j^H \Delta c_{sj}) = 0$$

A natural concern is that wages and migration shocks might affect construction costs through their impact on construction wages. I do two things to deal with this concern. First, when using the average Bartik shock to identify the unserviced housing supply equation, I compute the instrument without considering construction workers' wages. Second, in Table 5 I show that both the Bartik shock computed without considering construction workers and the Migration shock for high income households do not affect construction wages.

[27] See Equations 14 and 15.

[28] Figure 6 provides a graphical representation of the identification mechanism in place using identification of the serviced housing supply as an example.

[29] In principle I could use both Bartik and Migration shocks (one for each type of household) in each of the two equations. The specific choice has to do with the empirical relevance of each instrument for predicting changes in housing demand. See First Stage F's in Table 6.

5.4 Estimation summary

Summarizing the discussion above, estimation consists in running IV regressions for Equations 25, 26, 27, and 28 to estimate the set of structural parameters $\sigma^L, \sigma^H, \gamma_u, \gamma_s$. All equations feature one endogenous variable and one instrument and are thus exactly identified. Note that all the regression errors have a structural interpretation in the context of the model and are then inputs for the general equilibrium computation in Section 7.

6 Estimates

Table 6 presents the 2SLS regression results for the four estimating equations. The first column estimates Equation 25 by running low income households' population growth on real wage growth and identifies the parameter σ^L. The point estimate $1/\hat{\sigma}^L = 1.67$ gives a picture of highly mobile households, which is coherent with the evidence on households' spatial mobility in Section 3.[30]

Multiplying $1/\hat{\sigma}^L$ by the calibrated housing expenditure shares $(\alpha_U^L, \alpha_S^L = 0.25, 0.30)$ yields low income households' elasticity with respect to housing rents for each type of house.[31] Those coefficients imply that a 1% increase in unserviced housing rents reduces the number of low income households demanding unserviced houses by 0.4% and a 1% increase in serviced housing rents reduces the number of low income households demanding serviced houses by 0.5%.

High income households' reaction to real wages' given by $1/\hat{\sigma}^H$ in Column 2 is much noisier and seems smaller in magnitude. This is coherent with high income households exhibiting a much higher urbanization rate (83% compared to 64% for low income households), which limits their rural-urban migration margin. High income households' housing consumption share is 0.16, which implies that their responses to housing rents' shocks are very small.

Columns 3 and 4 in Table 6 show estimates for $\hat{\gamma}^U$ and $\hat{\gamma}^S$. 2SLS regression estimates show two positively sloped housing supplies, with unserviced housing prices reacting much less to housing demand shocks than serviced housing ones. For instance, a 1% increase in the demand for unserviced housing leads to less than 0.1% higher unserviced housing rents. The

[30] Note that this estimate does not give yet the equilibrium effect of higher wages because it does not account for the effect of increased housing demand on housing prices and for the impact of changing housing prices on households' housing demand.

[31] Although I will be referring generically to estimated parameters as 'elasticities', keep in mind that in the context of the underlying multinomial logit, parameters should be interpreted as elasticities for "small" cities. For instance, given the log-linear indirect utility function, the wage elasticity for choice mj is $(1 - N_{mj}) * 1/\sigma^L$.

same increase in the demand for serviced housing leads to 0.4% higher serviced housing rents.

In terms of external validation of these housing supply-side estimates, my serviced housing (inverse) supply elasticity is very similar to the 0.47 reported by Saiz (2010) for (serviced) housing in the US.[32] Moreover, the fact on unserviced housing supply being relatively more elastic than serviced one has been one of the usual suspects in the slum growth and urbanization literature when trying to explain why economic dynamism leads to slum growth (UN, 2003). This is, to my best knowledge, the first empirical study confirming this hypothesis.

7 General Equilibrium and Counterfactuals

In this section I solve for 1991-2010 changes in the general equilibrium of Brazil's system of cities for a set of different scenarios. This exercise yields the population reallocation and welfare effects of changes in slum policies and economic fundamentals. In particular, I contrast the effect of these changes when they take in a few cities versus when they take place in all cities simultaneously.

7.1 Benchmark General Equilibrium

I solve for the 1991-2010 changes in the general equilibrium of the model. Intuitively, this implies finding population growth and rent growth in each type of house and city such that two conditions hold. First, growth in housing supply must equal growth in housing demand in all housing markets. Second, for each type of household, the sum of weighted population growth in all cities and the countryside must add up to the national population growth rate.

The general equilibrium computation takes as inputs the point estimates for the set of structural parameters $(\sigma^L, \sigma^H, \gamma_u, \gamma_s)$, the calibrated structural parameters $(\alpha_u^L, \alpha_s^L, \alpha_s^H)$, the set of regression residuals $(\Delta\bar{a}_{uj}^L, \Delta\bar{a}_{sj}^L, \Delta\bar{a}_{sj}^H, \Delta c_{uj}, \Delta c_{sj})$, and a set of exogenous variables from the data $(\Delta w_j^L, \Delta w_j^H, \Delta\overline{n}^L, \Delta\overline{n}^H)$. These elements fully characterize the set of linear equations of the model expressed in first time differences. The resulting system of equations is fully linear and has the same number of equations as endogenous variables. The endogenous variables are the population growth rates $\Delta n_{mj}^L, \Delta n_c^L, \Delta n_{sj}^H, \Delta n_c^H$ and the housing rents growth rates Δp_{mjt}. The set of linear equilibrium equations are the four equations estimated

[32]The unserviced housing estimate could be interpreted as coherent with the recent finding by Henderson, Venables & Regan (2016) that slum prices in Nairobi do not decreasing with distance from the central business district. This gradient is the typical microfoundation for upward sloping housing supplies in the Alonso-Mills-Muth urban model (Fujita, 1989).

above plus two equations guaranteeing that the weighted sum of local population growth rates add up to national exogenous population growth.[33]

Table 7 presents some aggregate statistics for the system of cities for the actual data and for a set of simulated equilibrium scenarios. The table summarizes the system's endogenous variables by showing Brazil's aggregate urbanization rate and unserviced urban housing share in 2010 as well as unserviced and serviced average rent growth between 1991 and 2010.[34] Columns 1 and 2 display the actual values of those variables in the data for 1991 and 2010, Column 3 shows how well the model's equilibrium replicates the data for 2010, and the remaining columns show 2010 statistics under a few counterfactual scenarios. The bottom part of Table 7 presents some summary statistics for the exogenous drivers of the model in each counterfactual scenario.

A quick look at the exogenous drivers in the lower part of Table 7 gives some intuitions on the story behind the changes observed between 1991 and 2010 in Brazil's urbanization and unserviced housing incidence. A first thing to note is that despite Brazil's relatively high level of urbanization, almost half of low income households lived outside of the sample of 272 cities in 1991. This leaves ample space for substantial reallocations of households across the system of cities. In this context, growing urban incomes, coupled with households being highly elastic in their responses to urban income growth, are then part of the explanation for the rise in urbanization.

Section 3 established that this was a period of "pro-poor" economic growth in Brazil, with incomes improving both for low and high income households but much for the former than for the latter. This trend in the level and dispersion of incomes led to changes in the composition of the population between low and high income households. Specifically, there was a ten point shift in the national population share of high income households. In the context of the model above, this population change mechanically reduces the share of urban residents in unserviced houses.

Column 3 in Table 7 presents the model's general equilibrium computed for the actual exogenous parameters. The model does well in matching the four (endogenous) aggregate statistics in Table 7. The goodness of fit of the model is also illustrated by Figure 7, which

[33]Note that the linear decomposition of a growth rate into a weighted sum of its components is exact for the exponential growth rates but not for log-growth rates. I then rely on the approximation between log and exponential growth rates to linearize the two national-level equations as well as each serviced housing market equation.

[34]Here I define urbanization as the share of households in my sample of 272 cities. These urbanization rates differ from the official ones which I reported in Table 1. This discrepancy results from official urbanization rates considering small towns as urban and this paper considering urban only those municipalities with at least 50,000 people in 2010.

plots the actual versus the predicted values for the three endogenous population growth variables: low income households in urban unserviced houses, low income households in serviced houses, and high income households in serviced houses.

7.2 Slum growth, Urbanization, and cities' economic dynamism

The first counterfactual exercise looks at the mechanics of urban economic growth and slum growth in developing countries' system of cities. By looking at the long-run historical trajectories of today's developed countries' cities, a series of authors have noted that slum incidence disappears in the long run as countries become richer (World Bank, 2009; Glaeser, 2012). This seems to hold true for Brazil between 1991 and 2010. As noted in Section 3, in this period Brazil's per capita GDP grew by 41% and unserviced housing incidence went down from 28% to 18% (Table 7).

In order to see how cities' economic growth affect slum incidence in the context of the model, I simulate an extra wage increase of 20% for both types of households in all cities. This type of shock has two main effects in the context of the model. First, higher wages bring more low income households to cities (elasticity of 1.7). These migration flows imply higher housing demand and thus make both housing rents grow. In particular, given the elasticities estimated above, serviced rents grow much more (0.4) than unserviced ones (0.1). Also, because serviced houses are more expensive, changes in serviced housing rents impact low income households' housing demand more than changes in unserviced housing rents. These housing demand and supply mechanics define a first effect of higher wages, which pushes unserviced housing incidence upwards. The second effect operates in the opposite direction. When all wages grow by an extra 20%, the 2010 share of high income households goes up 35.4% to 45.5%. Since high income households' unserviced housing incidence is very low, this change in the population's composition mechanically pushes slum incidence downwards.

In order to quantify the role of each of these two opposite effects of urban led economic growth on slum incidence, Column 2 in Table 7 shows the set of summary statistics for 2010 without changing the population composition. When population composition does not adjust, both housing rents increase reflecting higher housing demand for both types of houses. Under this scenario, the equilibrium unserviced share is slightly higher and the number of households in unserviced houses (not shown in the Table) is 7% higher compared to the benchmark of Column 1. Column 3 shows the full effect of higher urban wages on the system of cities by including the changes in population composition implied by the extra 20% wage shock. Urbanization in Column 3 goes up by 7.8% and unserviced housing goes down by 2% with respect to the benchmark.

The exercise in the previous paragraph shows how national income increases are key in explaining long run reductions in slum incidence. Another way to look at this is to contrast the effect of generalized economic growth with the effects of spatially unbalanced growth (i.e. some cities growing much faster than others). A common place in the slum growth and urbanization literature is that rapidly growing cities experience enormous growth in the number of slum households in periods of two or three decades. In order to evaluate this idea for the case of Brazil between 1991 and 2010, Table 8 regresses unserviced housing growth (Columns 1 and 2) and changes in unserviced housing incidence (Columns 3 and 4) on exogenous local economic shocks captured by the average Bartik instrument. Regression coefficients show reduced-form evidence on how economic dynamism leads to increases both in the number of unserviced houses (Column 1) and in unserviced housing incidence (Column 3). This relationship is robust to controlling for cities' initial population size and income levels (Columns 2 and 4).

This paper's general equilibrium approach helps to understand why unbalanced economic growth might lead to higher slum growth and slum incidence in economically dynamic cites. First, when balanced urban economic growth takes place, all cities attract rural households at the same time and this decreases the housing demand that any single city faces. In contrast, when a few cities grow, they become the focus of all rural migrants and they also attract households from other, less dynamic, cities. Second, when balanced urban economic growth takes place, it activates the composition effect by which households become wealthier and switch to serviced housing in all cities. In order to illustrate this process using the structure of the model, I consider what happens to an average city of around 100,000 people when all cities' wages grow by 20% versus when only that city's wages grow by 20%. First, when economic growth takes place in all cities, unserviced housing incidence in that city goes down by 3.3% and the number of unserviced households goes down by 3.6%. Second, when only this city grows, unserviced housing incidence in the city grows by 1.2% and the number of unserviced households grows by 26.6%.

7.3 Slum policies

Turning to the role of policies, I model slum repression and slum upgrading as exogenous shifts in the supply of each type of housing. In particular, I analyze what happens to households' spatial allocation and welfare if a few versus all cities implement these policies.

Starting with slum repression, this policy may take the form of evictions once houses have been already built but also can operate ex-ante by making it harder for households to build in land without services (UN, 2003). In any of these cases, slum repression substantially increases the cost of producing unserviced housing. Then, I model it as a generic backwards shift in the supply for unserviced houses. Specifically, I implement a supply shift increasing the Δc_{uj} term by 20 points.[35]

When a single medium-size city implements slum repression, the model indicates that this city reduces its number of unserviced houses by 7.4%. The mechanism in place involves rent elastic low income households reacting to higher housing costs by moving to unserviced houses in other cities where there is no repression. However, if this policy generalizes to all cities, unserviced housing becomes more expensive everywhere and the reduction in the number of unserviced households in that single medium city goes down to 6.3%. The generalization of slum repression to all cities also brings in significant nation-wide changes. Column 4 in Table 7 shows the national summary statistics for the counterfactual scenario in which all cities repress slum formation. Slum repression in all cities shows up as a huge spike in equilibrium unserviced rents in Column 4. Households react to this price shock by moving both to serviced houses and rural areas. The first movement shows up as higher serviced housing rents, which grow 1.2% more than in the baseline due to increased demand from those households leaving unserviced houses. The second movement, from unserviced houses to rural areas, shows up as a lower equilibrium urbanization rate. In Column 4 of Table 7 urbanization in 2010 goes down by 0.4%. These reallocation effects have welfare consequence since households are moving away from what was their best location choice in terms of wages, housing prices, and amenities. Specifically, low income households' welfare is 1.1% lower with respect to the benchmark after this policy.[36] Slum repression's impact on high income households' welfare and spatial allocation is negligible.

Slum upgrading consists in bringing services and other amenities to previously unserviced houses. In terms of the two housing markets in the model, slum upgrading can be thought as withdrawing substantial numbers of unserviced houses and simultaneously adding an equivalent number of serviced houses. I then model this policy as a shift of the unserviced supply backwards and a simultaneous shift of the serviced supply forward. The magnitudes of the housing supply shifts are such that the equilibrium number of withdrawn unserviced houses equals the number of added serviced houses. Specifically, I simulate a scenario in

[35]This shock would increase Δp_{uj} by 20 points if housing quantity were fixed. Note that given that the estimated inverse housing supply elasticity is almost horizontal, this shock will translate to almost 20 points higher equilibrium unserviced housing prices.

[36]See Appendix A.2 welfare calculation's details.

which cities target to reduce the 1991 stock of unserviced houses by 10%.[37]

In the context of the model, the subsidy in favor of serviced houses and against unserviced ones impacts serviced rents downwards and unserviced rents upwards and this makes households substitute from one type of housing to the other one. The effects of this policy on any single city depend on whether this city is the only one implementing this policy or not. For instance, when a single medium-size city does slum upgrading, it reduces its number of unserviced houses by 2.3% and increases the number of serviced houses by 1.7% in comparison with when all cities implement it.

Column 5 in Table 7 shows aggregate statistics for the counterfactual scenario when all cities do slum upgrading. Although this policy reallocates households away from their benchmark location choices and thus could potentially hurt low income households' welfare, I find that welfare improves for both types of households when all cities implement slum upgrading policies. This contrasts with what I find for slum repression and has to do with households attaching a higher amenity value to serviced houses in comparison to unserviced ones. Specifically, welfare improves 4.0% for low income households and 3.6% for high income households with respect to the benchmark.

8 Concluding Remarks

This paper contributes to a better understanding of developing world's contemporary urbanization processes with a particular focus on the housing quality dimension. I do this by characterizing households' location decisions across the unserviced/serviced and the urban/rural margins and cities' serviced and unserviced housing production capacities in reaction to housing demand shocks. I use this characterization in the context of a spatial general equilibrium model to study the effects of changes in economic fundamentals and common housing supply-side policies on the allocation of households across housing types, cities, and the countryside. This methodology explains how unbalanced urban economic growth leads to slum growth in dynamic cities and how this is not inconsistent with long run slum incidence going down as countries become richer. I show how the two main paradigms in terms of slum policy, slum repression and slum upgrading, have opposite effects on households' welfare. Also, I show how the reallocation effects of these slum policies on any given city depend on

[37] Exactly targeting that 10% reduction is a hard problem for cities since it involves a general equilibrium calculation. I assume the policy shifts both supply functions in opposite directions in order to achieve a 10% reduction under a partial equilibrium scenario. In any case, the exact magnitude of the policy is not meaningful.

what other cities are doing.

I would like to close the paper with a few remarks on directions for future work. The provision of water and sanitation services in cities features huge economies of scale and thus involves collective action problems. In order to keep the problem tractable, my paper abstract from this dimension but the economics of providing these and other urban amenities should be further explored. In another paper (Alves, 2014), I explore the local political economy of the problem and show that the political sign of the local government matters for which slum policies are implemented and for the local dynamics of slum incidence. A second issue to be further explored are the efficiency implications of slums' location in the internal structures of cities since slums are usually built in land that could have more efficient uses. This aspect has been recently studied by Henderson, Venables, Regan & Samsonov (2016) for the case of Nairobi.

References

Alves, G. (2014). On the determinants of Slum Formation: the role of Politics and Policies.

Bartik, T. J. (1991). Boon or Boondoggle? The debate over state and local economic development policies.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Bound, J. & Holzer, H. J. (2000). Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s. *Journal of Labor Economics*, *18*(1), 20–54.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, *19*(1), 22–64.

Castells-Quintana, D. (2016). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics*.

Chauvin, J. P., Glaeser, E., Ma, Y., & Tobio, K. (2016). What is Different About Urbanization in Rich and Poor Countries? Cities in Brazil, China, India and the United States. Working Paper 22002, National Bureau of Economic Research.

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.

Dix-Carneiro, R. (2014). Trade Liberalization and Labor Market Dynamics. *Econometrica*, *82*(3), 825–885.

Dix-Carneiro, R. & Kovak, B. K. (2015). Trade liberalization and the skill premium: A local labor markets approach. Technical report, National Bureau of Economic Research.

Feler, L. & Henderson, J. V. (2011). Exclusionary policies in urban development: Under-servicing migrant households in Brazilian cities. *Journal of Urban Economics*, *69*(3), 253–272.

Fujita, M. (1989). *Urban Economic Theory: Land Use and City Size*. Cambridge University Press.

Glaeser, E. L. (2008). *Cities, Agglomeration, and Spatial Equilibrium* (1 edition ed.). Oxford: Oxford University Press.

Glaeser, E. L. (2012). *Triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier, and happier*. New York, New York: Penguin Books.

Harari, M. (2015). Cities in bad shape: Urban geometry in india. *Job market paper, Massachusetts Institute of Technology.*

Henderson, J. V., Venables, A. J., & Regan, T. (2016). Building functional cities | Science. Technical Report 196, SERC.

Henderson, J. V., Venables, A. J., Regan, T., & Samsonov, I. (2016). Building functional cities. *Science*, *352*(6288), 946–947.

Hsieh, C.-T. & Moretti, E. (2015). Why do cities matter? Local growth and aggregate growth. Technical report, National Bureau of Economic Research.

Jedwab, R., Christiaensen, L., & Gindelsky, M. (2016). Demography, urbanization and development: Rural push, urban pull and ... urban push? *Journal of Urban Economics.*

Kline, P. & Moretti, E. (2014). People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs. *Annual Review of Economics*, *6*(1), 629–662.

Kovak, B. K. (2013). Regional Effects of Trade Reform: What Is the Correct Measure of Liberalization? *American Economic Review*, *103*(5), 1960–76.

Lustig, N., Lopez-Calva, L. F., & Ortiz-Juarez, E. (2013). Declining Inequality in Latin America in the 2000s: The Cases of Argentina, Brazil, and Mexico. *World Development*, *44*, 129–141.

Marx, B., Stoker, T., & Suri, T. (2013). The Economics of Slums in the Developing World. *Journal of Economic Perspectives*, *27*(4), 187–210.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

Morten, M. & Oliveira, J. (2016). Paving the Way to Development: Costly Migration and Labor Market Integration. Working Paper 22158, National Bureau of Economic Research.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the u.s. labor market. *Quarterly Journal of Economics*, 549–599.

Notowidigdo, M. J. (2011). The Incidence of Local Labor Demand Shocks. NBER Working Paper 17167, National Bureau of Economic Research, Inc.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy*, *90*(6), 1257–1278.

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, *3*.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, *125*(3), 1253–1296.

Serrato, J. & Zidar, O. (2016). Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms. *American Economic Review*, *106*(9), 2582–2624.

Train, K. E. (2009). *Discrete Choice Methods with Simulation* (2 edition ed.). Cambridge ; New York: Cambridge University Press.

UN (2012). *The Millennium Development Goals Report 2012*. United Nations Publications.

UN (2015). Millennium development goals. Technical report.

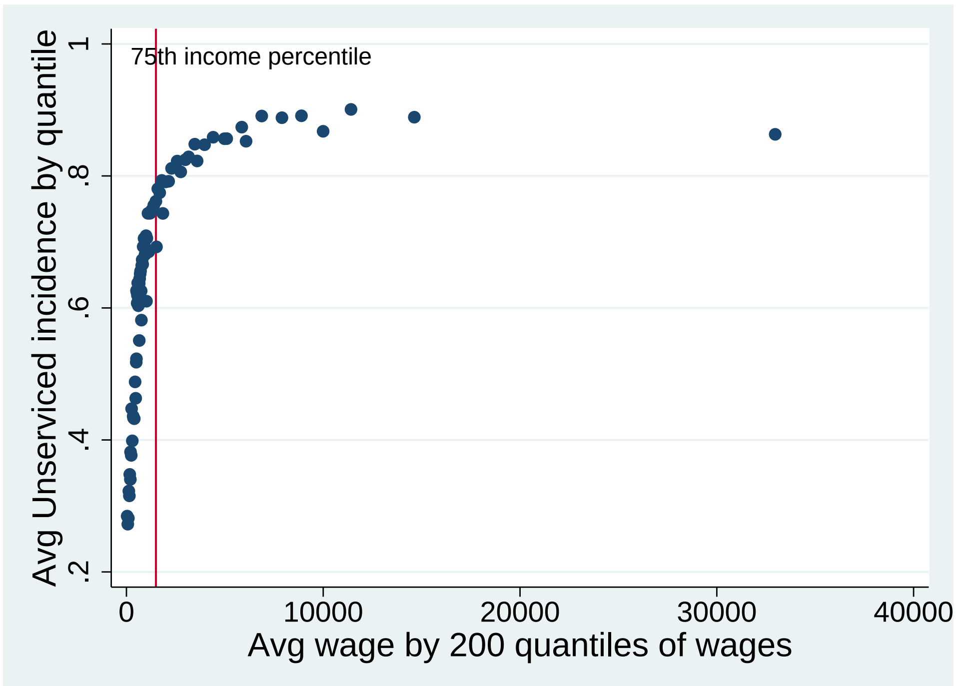
UN, U. N. H. S. (2003). *The Challenge of Slums: Global Report on Human Settlements 2003*. Routledge.

World Bank (2009). *World Development Report 2009: Reshaping Economic Geography*. World Bank.

A Appendix

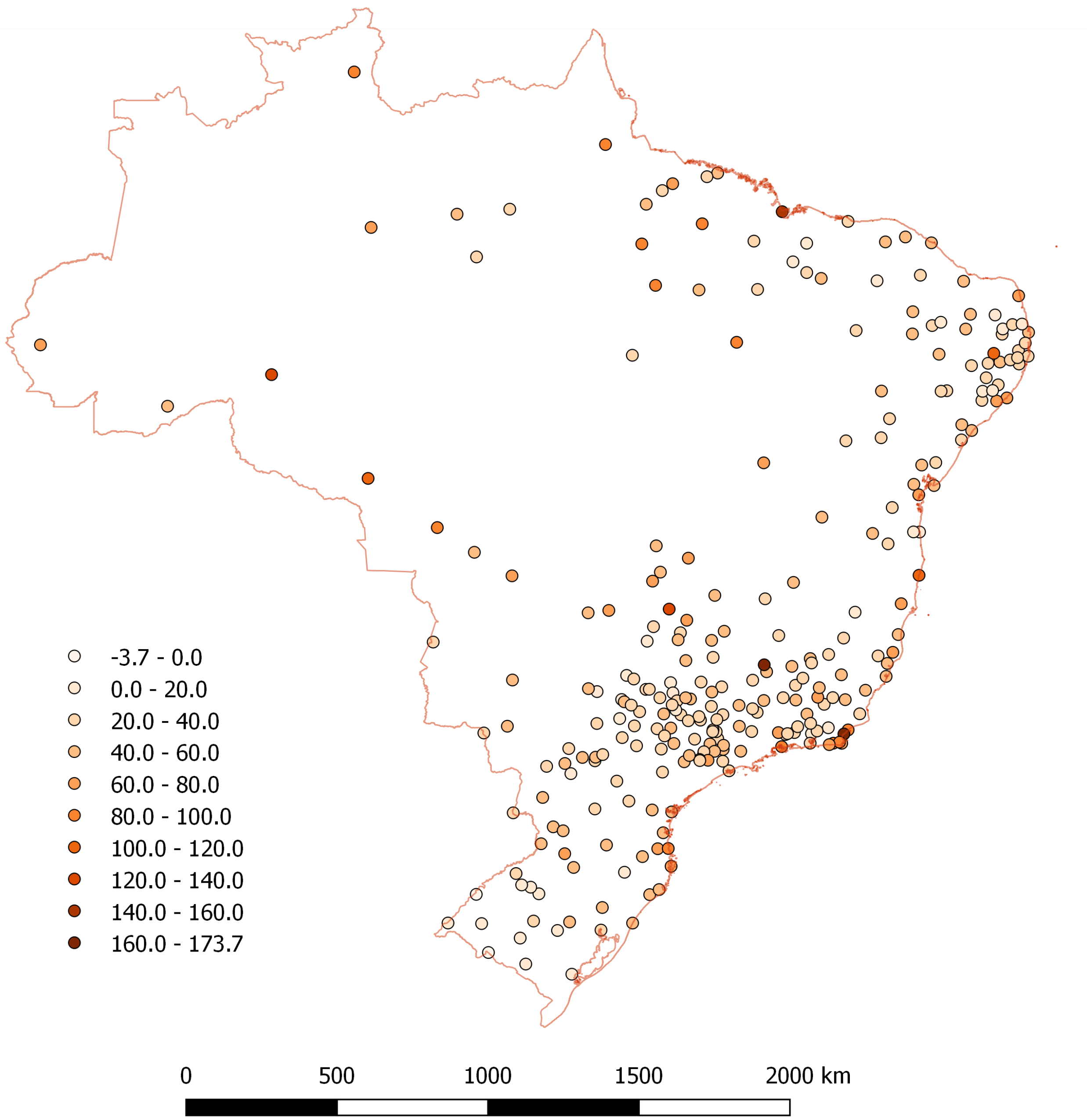
A.1 Graphs and Tables

Figure 1: Serviced housing and wages



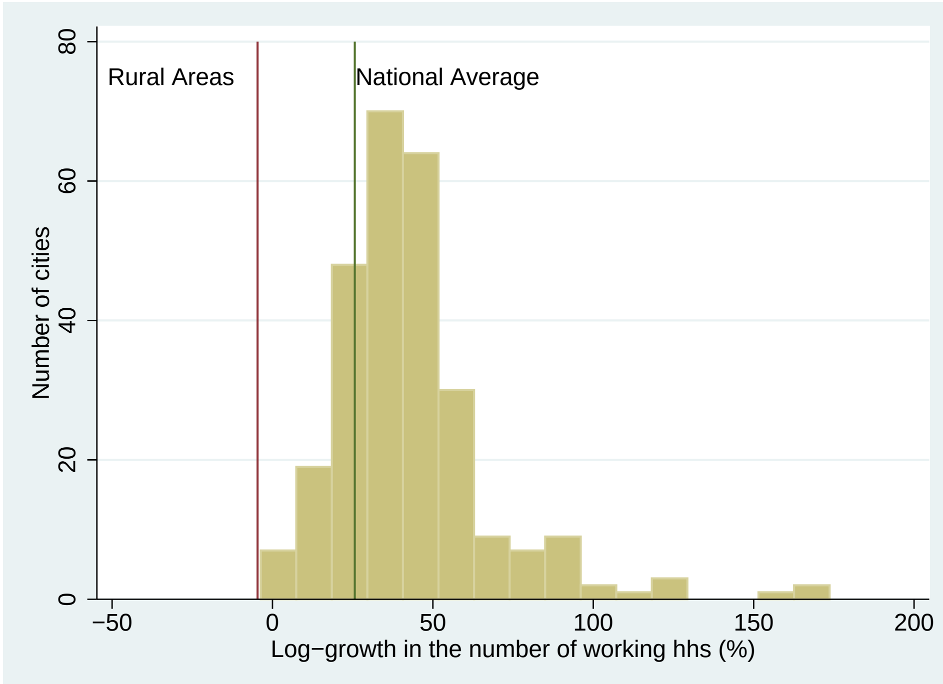
Each dot represents averages of serviced housing and wages by 200 quantiles of wages. Wages measured in Brazilian currency (Reais) and expressed in 2010 prices. Own processing of Brazilian Census data. See data section for details.

Figure 2: Map of Cities' Population Growth



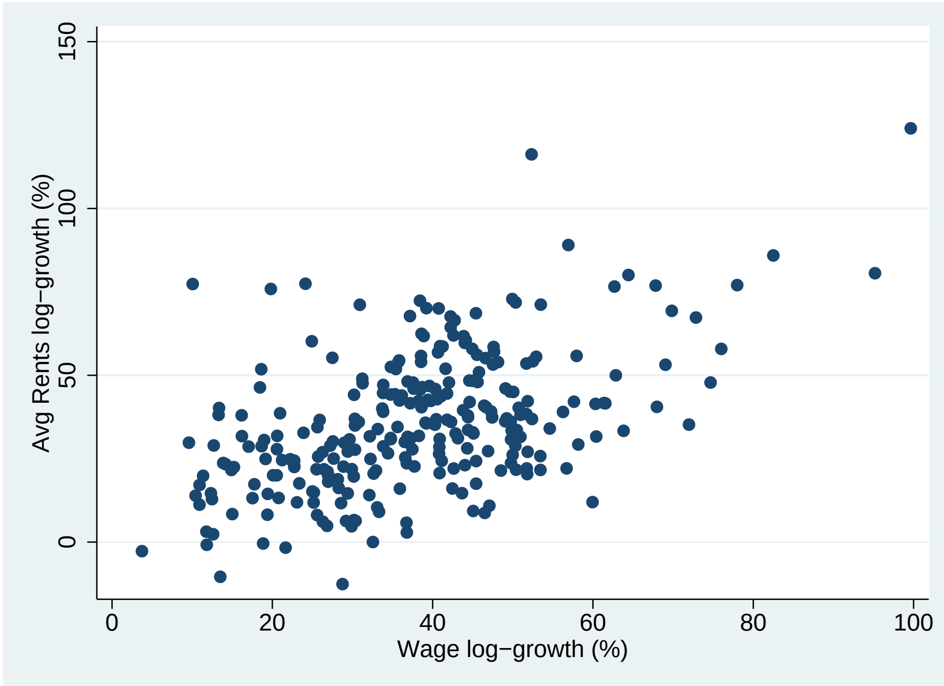
272 cities. Growth in the number of working housholds. Own processing of Brazilian Census data. See data section for details.

Figure 3: Histogram of Cities' Population Growth between 1991 and 2010



272 cities. Growth in the number of working housholds. Own processing of Brazilian Census data. See data section for details.

Figure 4: Wages and Rents growth



Each dot represents one of 272 cities. Own processing of Brazilian Census data. See data section for details.

Figure 5: Unserviced Housing growth, Wages and Rents

272 cities. Growth in the number of working housholds. Own processing of Brazilian Census data. See data section for details.

Figure 6: Identification of Housing Supply equations

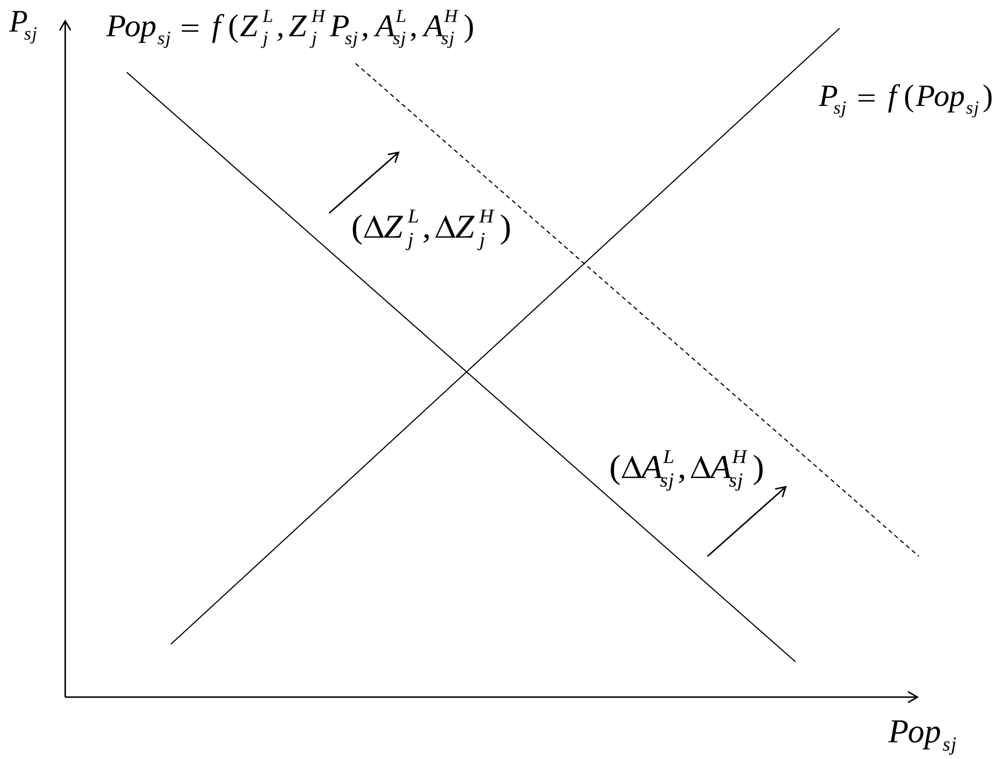


Figure 7: Model's Fit

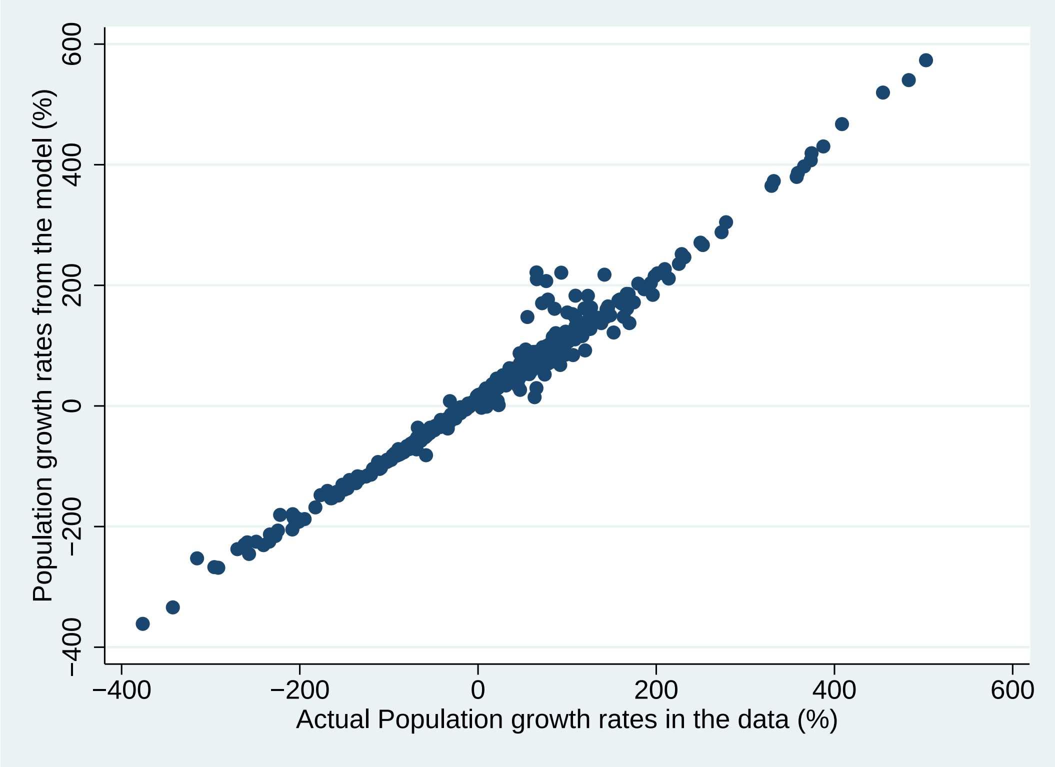


Table 1: National statistics for Brazil between 1991 and 2010

	1991	2010
Per Capita GDP (2005 US$ PPP)	10,263	14,409
Gini Index	0.64	0.61
Total Population (millions)	145.7	181.9
Urbanization (%)	74.7	84.3
Working Households (millions)	26.5	34.2
Living in the 272 cities	16.2	24.4
In houses without services	5.3	5.7
Wage growth (%)		
Low income	29.4	
High income	16.3	
Population growth (%)		
Low income	11.8	
High income	56.9	

Source: World Bank (for GDP and Urbanization), IBGE for Gini Index, and own processing of Census data for the rest.

Table 2: Share of urban households' heads not born in the city where they currently live (2010 census)

	Unserviced	Serviced
Born rural area	48.5	45.9
Born in other cities	10.7	8.7
Born elsewhere	59.2	54.6

Source: own processing of Census data.

Table 3: Unserviced housing growth, wages and rents

	(1)	(2)	(3)	(4)	(5)
	Unserviced housing growth				
Dwage	1.43***		1.07***	0.97***	1.16***
	(0.33)		(0.38)	(0.37)	(0.39)
Drentavg		0.92***	0.49*	0.55**	0.58**
		(0.23)	(0.26)	(0.28)	(0.28)
Dgini				0.59	0.09
				(0.66)	(0.71)
lnhhs_1					10.55**
					(4.08)
Constant	-88.53***	-66.78***	-92.34***	-84.96***	-202.49***
	(14.52)	(10.82)	(14.95)	(15.81)	(52.12)
Observations	272	272	272	272	272
R-squared	0.07	0.05	0.08	0.08	0.10

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Source: own processing of Census data.

Table 4: Instruments' First Stages

	(1)	(2)	(3)	(4)
VARIABLES	DW_L	DW_H	DN_U	DN_S
Bartik_L	2.06***	-0.04		
	(0.26)	(0.35)		
Bartik_H	0.22	0.96***		
	(0.16)	(0.16)		
Bartik_avg			3.17***	
			(0.76)	
Migration_H				0.17***
				(0.06)
Constant	-35.60***	1.68	-111.53***	103.92***
	(9.03)	(11.77)	(19.76)	(9.88)
Observations	272	272	272	272
R-squared	0.19	0.11	0.05	0.03
F	31.77	18.10	17.62	8.715

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
The average Bartik shock is computed without construction workers.

Table 5: Instruments and construction wages

	(1)	(2)
VARIABLES	Dwage_c	Dwage_c
B_nc	-0.31	
	(0.26)	
M_H		0.05
		(0.05)
Constant	63.23***	55.82***
	(6.19)	(1.44)
Observations	272	272
R-squared	0.01	0.00

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. B_nc is the average Bartik shock computed without construction workers.

Table 6: 2SLS estimates

	(1)	(2)	(3)	(4)
	Δn^L	Δn^H	Δp_u	Δp_s
$\Delta w^L - \alpha_m * \Delta p_m$	1.67*			
	(0.88)			
$\Delta w^H - \alpha_s * \Delta p_s$		0.46		
		(0.54)		
Δz_u^D			0.07	
			(0.11)	
Δz_s^D				0.37*
				(0.22)
Observations	544	272	272	272
1st stage F	61.9	30.9	19.0	15.6

Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
See Section 5 for details on moment conditions.

Table 7: General Equilibrium and Counterfactuals

	Data		Model				
	1991	2010	(1)	(2)	(3)	(4)	(5)
Urbanization	60.6	71.1	70.4	75.0	78.2	70.0	71.0
Share Unserviced Urban	28.4	17.5	15.2	15.4	13.4	14.4	13.4
Unserviced Rent Growth	47.6		48.7	49.5	48.3	68.3	72.5
Serviced Rent Growth	18.7		22.1	24.0	29.3	22.3	-1.8
Wage growth L's	29.4		29.4	49.4	49.4	29.4	29.4
Wage growth H's	16.3		16.3	36.3	36.3	16.3	16.3
Pop growth L's	11.8		11.8	11.8	-4.7	11.8	11.8
Pop growth H's	56.9		56.9	56.9	81.8	56.9	56.9
Share of H types	25.0	35.4	35.4	35.4	45.5	35.4	35.4

1 - Benchmark. 2 - Extra 20% urban wage growth without changing population composition. 3 - Extra 20% urban wage growth changing population composition 4 - "Slum repression": 20% higher cost of supplying unserviced housing. 5 - "Slum upgrading" : Turn unserviced houses into serviced (See Section 7 for details).
Rent growth rates are national averages weighted by city population.

Table 8: Productivity shocks and unserviced housing growth

	(1)	(2)	(3)	(4)
	Unserviced housing growth		Change in unserviced incidence	
Bartik	4.5***	3.2***	0.2	1.0***
	(0.9)	(1.1)	(0.2)	(0.2)
Constant	-169.6***	664.9*	-21.5***	-209.8***
	(27.8)	(356.8)	(5.8)	(73.9)
Observations	272	272	272	272
R-squared	0.1	0.1	0.0	0.2
Controls	No	Yes	No	Yes

Robust s.e. in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Controls are logs of 1991 number of households in the city, low income wages, and high income wages.

A.2 Welfare Calculation

The logit structure features a closed form solution for the welfare associated to a set of alternatives. Since the full indirect utility is not observed, one must calculate an expected consumer surplus by integrating over the known probability of the extreme value idiosyncratic error. Following Train (2009), define the expected consumer surplus CS_i for a generic household i from the set of location choices O as:

$$E(CS_i) = \sigma E[\max(v_{mjt}/\sigma + \epsilon_{imjt})]$$

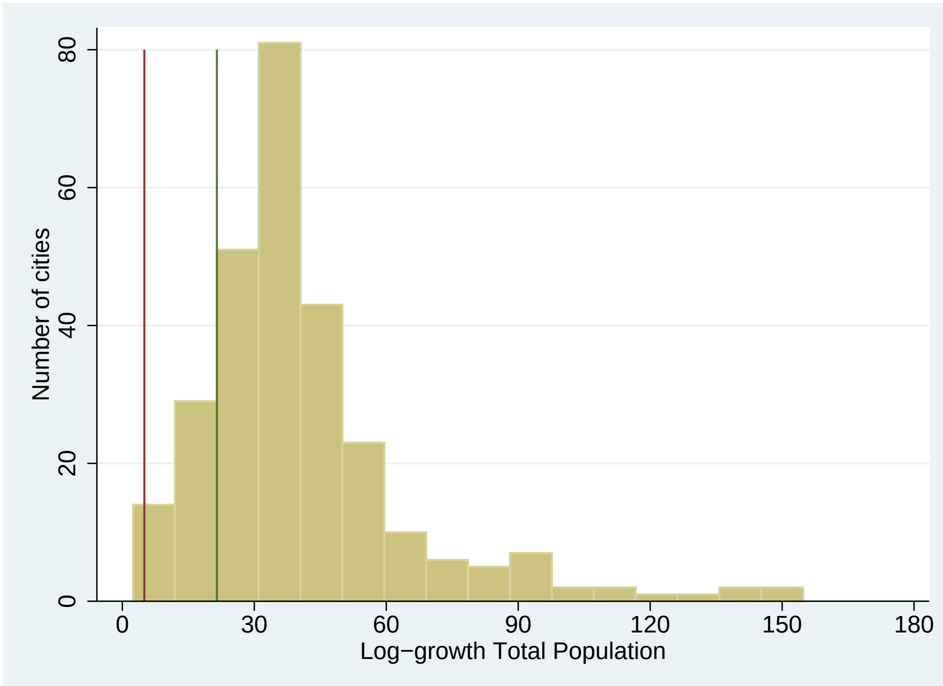
The expectation term yields the expected utility from the discrete choice problem and pre-multiplying the inverse of the marginal utility of wages σ express expected utility in monetary terms. Given the extreme value distribution assumption, the last expression simplifies to:

$$E(CS_i) = \sigma \ln \left(\sum_O v_{mjt}/\sigma \right) + C,$$

where C is a constant capturing the absolute level of utility which is not identified in the discrete choice model. This constant drops when I look at the difference in welfare before and after changes in policies.

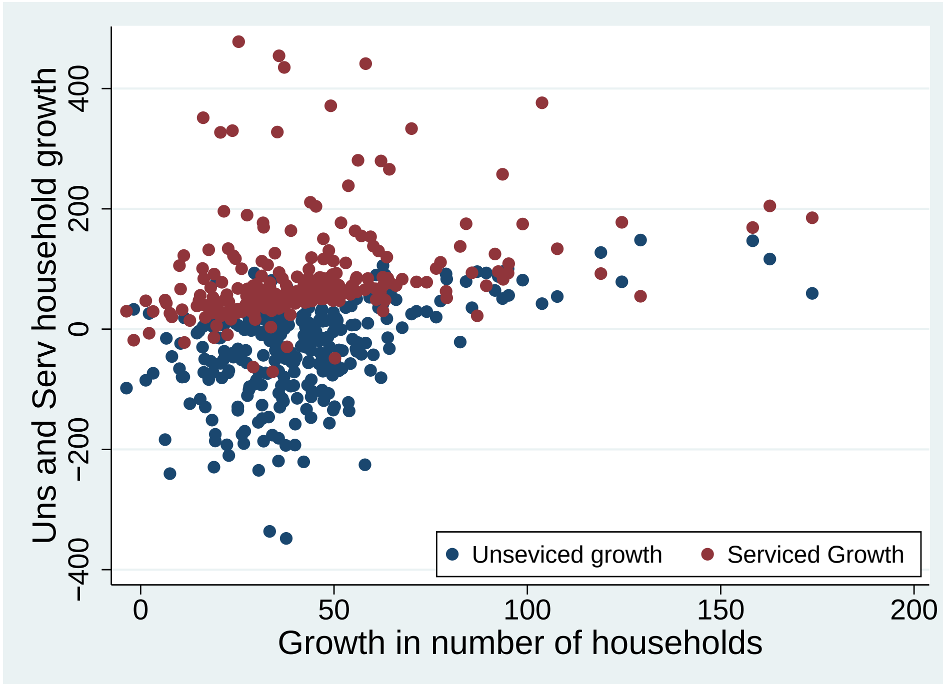
A.3 Appendix CAF

Figure 8: Histogram of Cities' Population Growth between 1991 and 2010



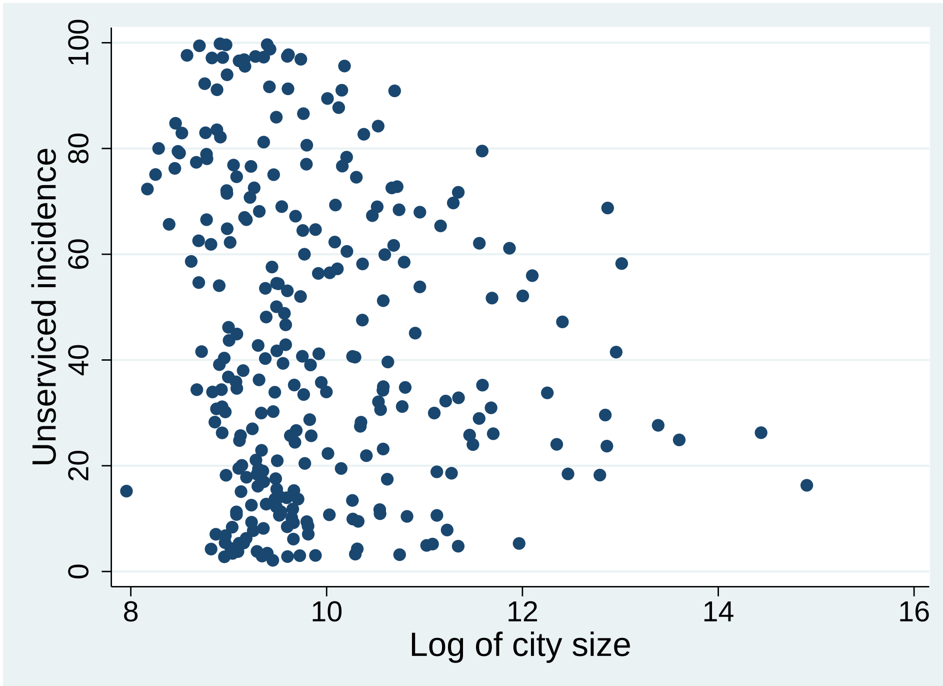
Red vertical line corresponds to population growth of rural areas. Green vertical line corresponds to national population growth. Own processing of Brazilian Census data. See data section for details.

Figure 9: Growth in the number of households by type of house. 1991-2010



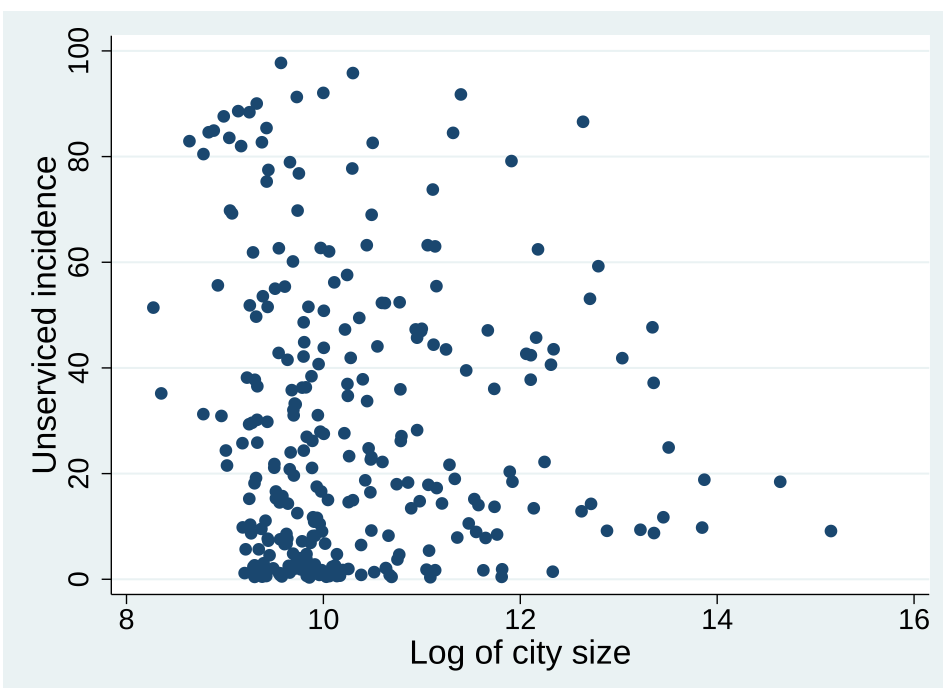
272 cities. Growth in the number of working households. Own processing of Brazilian Census data. See data section for details.

Figure 10: Unserviced share and city size. 1991



272 cities. Working households. Own processing of Brazilian Census data. See data section for details.

Figure 11: Unserviced share and city size. 2010



272 cities. Working households. Own processing of Brazilian Census data. See data section for details.

Table 9: Unserviced housing growth. Young and Old households

	Unserviced Housing Growth							
	Young				Old			
Dwage	2.12***		1.65***		1.16***		0.87**	
	(0.41)		(0.46)		(0.34)		(0.38)	
Drentavg		1.29***	0.65**			0.74***	0.40	
		(0.28)	(0.31)			(0.23)	(0.26)	
Bartik				3.87*				-0.16
				(2.29)				(1.73)
Constant	-155.67***	-120.65***	-160.71***	-204.16***	-79.53***	-61.80***	-82.60***	-29.98
	(18.87)	(13.35)	(19.32)	(76.39)	(15.00)	(10.87)	(15.41)	(58.06)
Observations	269	269	269	269	272	272	272	272
R-squared	0.10	0.07	0.11	0.01	0.04	0.03	0.05	0.00

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