# FAR Regulations and the Spatial Size of Brazilian Cities

Ricardo Carvalho de Andrade Lima*

Leonardo Monasterio†

August 2, 2022

**Abstract**

This paper evaluates the influence of the maximum allowed floor area ratio (FAR) on the spatial size of cities. We built a novel database on building height restrictions for the 325 largest Brazilian cities and combined it with recent satellite data. Our estimations show that, as predicted by theory, tighter constraints lead to more urban sprawl, and this result is robust to several specifications, including the framework proposed by Cinelli and Hazlett (2020). Using the share of homeowners among high-income households as an instrumental variable for FAR stringency, we find that the decrease of one standard deviation in the maximum allowed FAR increases the spatial area of a city by 12.4% (or 8km$^2$). Additionally, increasing the stringency of maximum FAR generates an annual cost of about US$ 2.36 million per year per average city.

***Keywords:*** welfare costs, land-use restrictions, urban sprawl, Brazil.

***JEL Code:*** H11, H77, R11.

*Catholic University of Brasília, Brazil, e-mail: ricardocarvalho2009@gmail.com.

†Institute for Applied Economic Research (Ipea), IDP- Brasília, and National Council for Scientific and Technological Development, e-mail: leonardo.monasterio@gmail.com. The first author acknowledges the financial support of Enap (National School of Public Administration- Brazil).

# 1 Introduction

The increase in the spatial size of cities and urban sprawl is a worldwide phenomenon that has raised significant concerns for policymakers. Although it may be a natural market equilibrium with potential benefits, expanding the urban area with sprawl tends to provoke a set of negative externalities. These include rising commuting costs, higher levels of car pollution, congestion on highways, loss of open spaces and natural amenities, and difficulties in the provision of public goods and infrastructure (Brueckner and Sridhar, 2012; Nechyba and Walsh, 2004). In recent years, Brazilian cities have expanded their urban area impressively. According to satellite images collected by MapBiomas, between 1985 to 2020, the size of the country's total urban area grew by 96.36%.

Local governments traditionally implement urban growth boundaries[1] (UGBs) policies to directly control the size of urban areas and avoid the negative externalities associated with sprawl. However, this instrument is just one of several local land use regulations with the potential to influence the spatial size of cities (Geshkov and DeSalvo, 2012). Regulations restricting the height and density of buildings (such as the maximum allowed floor area ratio[2]) can generate an unintentional stimulus to urban sprawl. According to the theoretical framework developed by Bertaud and Brueckner (2005), by imposing limits on the density and height of buildings in the city center, FAR regulations encourage land occupation in more distant areas and, consequently, accommodate population growth with the increase in urban land use. Therefore, regulations that limit FAR may offset the potential effects of UGB on urban sprawl, making the general consequences of land use restrictions on urban size uncertain and at odds with the policymaker's objectives.

The main objective of our paper is to investigate the role of maximum-allowed FAR regulations on the spatial size of Brazilian cities. Thus, we will check if the theoretical predictions of Bertaud and Brueckner (2005) are empirically corroborated for Brazil. For this purpose, we built a new database by collecting information about building-height limits for the largest 325 Brazilian cities. We combined this with recent satellite data that accurately captures the urban area size of each municipality. Our secondary objective is to calculate the potential welfare costs associated with the hypothetical adoption of a stringer FAR limit. To calculate the financial burden, we considered two classical negative externalities from urban sprawl: increased transport costs and increased pollutant emissions (effects on human health and the level of carbon dioxide).

Our paper directly relates to the empirical literature that investigates the effect of land use regulations on the shape of urban areas. Evidence for the impact of UGBs on US cities shows that this type of instrument is effective in reducing land consumption and the size of urban areas (Howell-Moroney, 2007, Wassmer, 2006, Paulsen, 2013), discouraging plot development (Dempsey and Plantinga, 2013) and reduces the urban blight (Hortas-Rico, 2015). On the other hand, evidence for FAR

[1]UGBs establish artificial boundaries in the city where lands inside the border can have urban use, and outside land can only have agricultural use or remain preserved.

[2]The Maximum Floor Area Ratio (FAR) is a number multiplied by the area of a lot that indicates the maximum amount of built-up area that developers can construct in the lot. For example, on a plot of 100 $m^2$ and a maximum FAR of 4, the maximum building area must be equivalent to 400 $m^2$.

regulations indicates that these policies can have unintended consequences. For example, Moon and Ahn (2022) show that a lower maximum-allowed FAR is associated with higher chances of early demolition and lower density of new construction in New York City. Also, for the United States, Geshkov and DeSalvo (2012) indicate that, unlike other land-use restrictions, the adoption of FAR regulation tends to generate significant expansions of urban areas. The literature evaluating this issue is still scarce in developing countries. For example, Zhou et al. (2017) analyze the consequences of land use planning policies in Chinese cities and find that these effectively limit the growth of built-up land. Exploring the Indian case, Brueckner and Sridhar (2012) show that cities that adopt stricter FAR limits have a larger urban spatial size. In addition, the flexibilization of these restrictions can improve the population's well-being by reducing commuting costs. Therefore, the empirical evidence is mixed, since the observed results strongly depend on the type of land use regulation and the context.

Our paper intends to contribute to this literature in three different ways. Firstly, we provide unprecedented evidence for the case of Brazil, a country with a recent urbanization process accompanied by rapid urban sprawl. Thus, we obtain a first assessment of the consequences of maximum-allowed FAR regulations on the spatial size of Brazilian cities and simultaneously estimate the potential economic costs associated with such restrictions. Existing empirical evidence is restricted to urban areas in the United States and India and, for this reason, cannot be easily generalized. Secondly, we deal with the potential existence of endogeneity by using two alternative strategies. We implement a set of sensitivity tools recently developed by Cinelli and Hazlett (2020) that indicate how OLS estimates changes in the face of threatening unobserved confounders. And we explore the variability of the local share of homeowners among the high-income group of households to construct an instrumental variable for the maximum FAR regulations. Most of the papers discussed above ignore the possibility of endogeneity due to omitted variable bias, which makes the set of previous evidence questionable. Finally, as discussed earlier, we developed and presented a new database summarizing information about the different types of building-height regulations for a sample of the 325 largest Brazilian cities.

We also contribute a new piece of evidence to a strand of literature that empirically tests the main drivers of urban sprawl based on the standard monocentric model (Brueckner and Fansler, 1983; McGrath, 2005; Paulsen, 2012; Deng et al., 2008; Spivey, 2008; Santos, 2020). The standard monocentric urban model predicts that income and population size positively impact the spatial dimension of cities because they drive the demand for housing and, consequently, for the built-up areas. Conversely, the value of agricultural land and commuting costs limit urban sprawl. The first increases the opportunity cost of rural-urban land conversions, and the second reduces consumers' disposable income for housing expenditures.

Our results show that cities that adopt a more stringent FAR regulation have larger spatial area sizes. This conclusion is robust to potential unobservable confounders and is maintained in different estimation strategies (OLS and 2SLS/IV) and alternative robustness tests. In our preferred 2SLS/IV specification, we show that a reduction of one standard deviation in the maximum allowed FAR value is associated with an average 12.4% increase in the spatial size of cities. We also note that cities with higher average income and larger populations have larger urban areas. These results confirm the theoretical predictions of the Bertaud and Brueckner (2005) model. Finally, increasing the FAR stringency generates the following economic costs per year for an average city: U$$1.9 million due to higher commuting costs, U$$29.7 thousand due to higher carbon emissions, and $449 thousand due to higher health costs associated with air pollution.

The remainder of the paper is organized as follows. Section 2 briefly describes the standard monocentric model that guided our empirical analysis. Section 3 presents the data and explains how we built the new database containing information about the building-height restrictions. In section 4, we describe the details of the empirical strategy. Section 5 presents the main results and the robustness tests. Section 6 reports the estimates of welfare costs associated with a hypothetical increase in the stringency of FAR regulations. Finally, Section 7 presents the conclusion and discusses some policy implications.

# 2 Theoretical Framework

Our empirical analysis will be based on the standard monocentric urban model developed by Alonso-Muth-Mills following the treatment Brueckner and Fansler (1983) proposed and the respective extension to include build-height restrictions set by Bertaud and Brueckner (2005).

## 2.1 The Spatial Size of a City Without Build-Height Restrictions

Firstly, the urban economy comprises $N$ identical consumers who work in the Central Business District (CBD) and earn a labor income equal to $y$. They must commute to the CBD daily for work at a cost equal to $t$ per mile of travel. The utility function of these consumers is denoted by $u(q, c)$, where $q$ is the consumption of housing in square footage and $c$ is a numeraire good. The labor income can be spent on the consumption of numeraire goods, on commuting expenses, or on the rent price per m$^2$ of housing. Defining $x$ as the radial distance in miles of travel to the CBD and $p$ the rent price per m$^2$, the budget constraint faced by consumers is given by: $y = c + pq + tx$. The rent price has spatial variation to ensure that all individuals have the same utility level (given by $\overline{u}$) regardless of their location in the city. Choosing the value of $q$ that maximizes the utility function subject to the budget constraint, i.e., $\max_q u(q, y - tx - pq) = \overline{u}$, it is possible to obtain the value of $p$ and $q$ as a function of $x, t, y$ and $\overline{u}$.

On the production side, developers produce housing with constant returns of scale through the combination of land and capital. The housing space per unit of land can be defined as $h(S)$, where $S$ is the capital-land ratio used in the production process. Furthermore, taking $i$ as the rental price of a unit of capital and $r$ the rental price of a unit of land, the developer chooses the value of $S$ that maximizes the profit function (per unit of land), i.e., $\max_S \pi = ph(S) - iS - r$. As $S$ and $r$ depend on the value of $p$, it is also possible to obtain the value of $S$ and $r$ as a function of $x, t, y$, and $\overline{u}$[3]. Finally, note that $h(S)$, measures the unrestricted Floor Area Ratio (FAR).

[3]It is possible to demonstrate that the housing price falls with increasing distance to the CBD, i.e.,$\frac{\partial p}{\partial x} < 0$. Lower prices discourage capital intensive buildings in peripheral areas so that $\frac{\partial S}{\partial x} < 0$.

It is possible to obtain the urban equilibrium conditions to determine the utility level $\overline{u}$ and the spatial city size $\overline{x}$ (defined as a distance between the CBD to the maximum limit of the urban edge). To obtain equilibrium, the following conditions must hold: (1) The value of the land rent in $\overline{x}$ is precisely equal to the agricultural land rent (denoted by $r_a$), and (2) The entire urban population must fit within the area of the city, $\overline{x}$. Formally, the urban equilibrium conditions can be written as:

$$r\left(\overline{x}, y, t, \overline{u}\right) = r_a \tag{1}$$

$$\int_0^{\overline{x}} 2\pi x \frac{h(S)}{q} dx = N \tag{2}$$

Equation (1) is an equilibrium condition because if the land rental price $r$ is higher than the agriculture land rent $r_a$, there would be incentives to convert land use in the rural-urban direction. In relation to equation (2), it should initially be noted that the ratio between the housing space per unit of land (defined by $h(S)$) and the housing space per square foot (defined by $q$) is the same as the number of dwellings per unit of land. If each individual lives in a house, the ratio $h(S)/q$ can be interpreted as the density of the city. By multiplying this density by the accumulated area of each concentric ring ($2\pi x$) and performing the integration, we obtain the population size.

Based on conditions (1) and (2) and on the equality of demand and supply of housing, it is possible to obtain the equilibrium values for $\overline{x}$ and $\overline{u}$. Through a comparative static analysis, Wheaton (1974) showed that, at equilibrium, the spatial size of the city $\overline{x}$ is influenced by the following variables:

$$\overline{x} = f\left(N, y, t, r_a\right) \tag{3}$$

It's possible to demonstrate based on equation (3) that the spatial size of the city increases with higher average income and population size and decreases with higher commuting costs or higher agriculture land rent. A larger urban population requires an expansion of the city area to accommodate new housing. Similarly, increases in income generate increases in demand for housing space and, consequently, expands the urban spatial size. On the other hand, increases in the commuting cost cause a reduction in the disposable income of consumers, which ends up generating a reduction in the demand for housing. Finally, higher agricultural land prices are associated with an increase in the opportunity cost of converting rural land for urban purposes, which also tends to generate compact urban areas.

## 2.2 The Spatial Size of a City with Build-Height Restrictions

In an urban economy constrained by FARs, the local government imposes an upper limit on the housing space that can be developed effectively on a unit of land. Following the Bertaud and Brueckner (2005) approach, this restriction can be written as $h\left(S\right) \leq h^*$, where $h\left(S\right)$ is the unrestricted FAR and $h^*$ is the maximum FAR stipulated by local officials. The practical influence of the maximum FAR tends to be stronger in the city's central areas since the capital-land ratio used in the housing production is higher in these places due to the higher land prices. The maximum FAR does not impose any real restrictions in the proximity of the urban edge area because the buildings already have a lower capital-land ratio.

Denoting by $\hat{x}$ the distance between the CBD and the area of effective influence of the maximum FAR and by $\overline{x}_1$ and $\overline{u}_1$ the urban area and the utility level of a city with build-height restriction, it is observed that the equilibrium conditions of the urban economy become the following:

$$r\left(\overline{x}_1, y, t, \overline{u}_1\right) = r_a \tag{4}$$

$$h\left(S\left(\hat{x}, \overline{u}_1\right)\right) = h^* \tag{5}$$

$$\int_0^{\hat{x}} 2\pi x \frac{h^*}{q} dx + \int_{\hat{x}}^{\overline{x}_1} 2\pi x \frac{h(S)}{q} dx = N \tag{6}$$

Equation (5) indicates that the FAR adopted in the area of influence of the maximum FAR ($\hat{x}$) will be precisely equal to the maximum limit imposed by the local government, given by $h^*$. Equation (6) is very similar to equation (2). The main difference is that the CBD density up to $\hat{x}$ becomes limited by the maximum FAR regulation, and from $\hat{x}$ to the edge of the city ($\overline{x}_1$), the density is the same as in equation (2). Based on equations (4), (5) and (6) it is possible to obtain the equilibrium values of $\hat{x}$, $\overline{x}_1$ and $\overline{u}_1$. Bertaud and Brueckner (2005) demonstrated that, compared to the unrestricted city discussed in subsection 2.1, the city that imposes building-height restrictions has a larger spatial size and a lower level of individual utility, i.e., $\overline{x}_1 > \overline{x}$ e $\overline{u}_1 < \overline{u}$.

# 3 Data

## 3.1 Measure of Maximum FAR Regulation

As a variable of interest that measures the stringency of land use restriction, we will use the maximum allowed FAR. This instrument can restrict the potential density and building height, as it limits the build-up area given the size of the lot. It is a regulation widely adopted in cities around the world because it avoids the negative externalities associated with excessive density and, at the same time, allows flexibility in the design of buildings. In the case of Brazil, the municipalities (the smallest administrative unit of the federation) define the urban planning rules, including the implementation and value of the maximum-allowed FAR.

As there is no official information about the implementation of building-height restrictions in each municipality, we implemented web scraping techniques to find information about the usage and the maximum allowed FAR value in cities with more than 100,000 inhabitants. In addition, we also collect information on the maximum height and number of floors allowed in each city and the regulations that establish these restrictions[4]. In this way, we develop a new database composed of a cross-section[5] including information about build-height limits for the 325 largest

[4]Usually, the local zoning law or the municipal master plan defines the building-height limits.

[5]It is not feasible to build a panel database with the maximum-allowed FAR due to the impossibility of tracking the historical evolution of the maximum FAR adopted in each city.

Brazilian cities, which contain 57.53% of the country's population. It should be noted that there is no rule or guideline for determining the appropriate maximum-allowed FAR, so urban planners and policymakers arbitrarily define its choice. As the maximum-allowed FAR value varies in the different zoning areas of the city, we follow Brueckner and Sridhar (2012) and adopt the highest maximum FAR within the entire urban space as our reference.

Figure 1 shows the distribution of the maximum allowed FAR value in the 325 largest Brazilian cities. It is possible to observe that most cities implement a relatively flexible maximum-allowed FAR but with substantial variability (ranging from 1 to 27). The range between 4 and 4.8 is the most common (20.6% of cities). The average is 5.17, and 12.92% of the cities do not adopt the maximum allowed FAR as a land use instrument.

[*Figure 1 here*]

It is expected that this set of cities that do not implement a maximum-allowed FAR regulation to have an almost unrestricted density and building heights[6]. Excluding these observations from our empirical analysis can lead to misleading conclusions because these cities may exhibit nationally relevant land-use dynamics and urban sprawl. Thus, to take advantage of this set of cities with an unrestricted FAR (which in theory would be equal to infinity) and not exclude them, we imputed the value of 27 to the maximum-allowed FAR of these observations (corresponds to the maximum value of our sample).

## 3.2 Other Variables and Descriptive Statistics

To measure the spatial size of Brazilian cities, we exploit information on the land use coverage area with urban infrastructure (in hectares) for the year 2020. This variable is collected annually by the Annual Mapping Project for Land Use and Coverage in Brazil (MapBiomas) through detailed satellite images from Landsat that indicate the aggregate coverage/use class area for each Brazilian municipality or state in the period from 1985 to 2020. Land cover data are classified into 20 categories using machine learning techniques. These include urban infrastructure, forest formation, grassland, other natural formations (separated by biome type), mining, and hydrographic use (lakes, rivers, and ocean).

As predicted by the urban monocentric model described in section 2, the population, income, commuting costs, and agricultural land value are the main drivers of cities' spatial size. We used data on population and average household income (measured in 2010 Brazilian Reais, R$) from the 2010 Demographic Census collected by the Brazilian Institute of Geography and Statistics (IBGE) to measure the population and average h. Unfortunately, no consolidated municipal database has information on agricultural land prices and commuting costs. Therefore, as a proxy for the value of agricultural land, we follow Santos (2020) and consider the ratio between the value of agricultural production (in 2010 Brazilian Reais, R$) and the harvest area (in hectares). We call this variable "agricultural income". IBGE collects both variables through the Municipal Agricultural Production survey. As detailed in Section 4.2, we will adopt the local share of homeowners among

[6] Of the 42 cities that do not adopt the maximum-allowed FAR regulation, only two implement regulations that impose restrictions on the number of building height or floors.

the high-income group of households as the instrumental variable for the maximum FAR regulation. We used three concepts from the 2010 Demographic Census to define the high-income subgroup: the share of homeowners with more than 10, 20, or 30 Brazilian minimum wages (MWs).

In some specifications, we included a set of geographic controls with the potential to influence the spatial size of cities: distance to the state capital, average altitude, average terrain ruggedness index, coastal city dummy, size of the hydrogeographic area, the number of natural conservation units and a dummy for cities localized in metropolitan areas. We will also consider a control vector of alternative land-use regulations: implementation of a minimum lot size (MLS) higher than that recommended by federal law 6799/79 (greater than 125 m$^2$), land subdivision regulation, building code law, and urban growth boundary. These variables were collected through the 1999 Survey of Basic Municipal Information carried out by the IBGE. They all have a dichotomous format and assume one if the municipality adopts the land-use restriction and 0 otherwise. Finally, we also calculate a simple regulation index to summarize the degree of stringency of all land-use instruments in the municipality. This regulatory index is calculated by simply summing up the number of regulations adopted in the cities and is traditionally used in previous studies to measure the land-use restrictiveness environment (See, for example, Quigley and Raphael, 2005). Table 1 shows descriptive statistics for the set of variables described above.

[*Table 1 here*]

Figure 2 presents scatter plots showing the relationship between the log of urban area size and its main determinants: log of average household income, log of population size, log of agricultural income, and the maximum allowed FAR. It is possible to notice that - except for agricultural income - all correlations are consistent with the theoretical predictions of Bertaud and Brueckner (2005) urban monocentric model. Income and population are positively correlated with the spatial size of the city, and the maximum allowed FAR is negatively correlated.

[*Figure 2 here*]

# 4 Empirical Strategy

## 4.1 OLS Regressions

To estimate the impact of maximum allowed FAR regulations on the spatial size of Brazilian cities, we took advantage of the substantial variability in the degree of land-use restrictions at the local level. Therefore, we estimated the following specification by the ordinary least squares method (OLS):

$$log(SpatialSize_i) = \alpha + \beta FAR.Max_i + \mu X_i' + \delta_s + \varepsilon_i \tag{7}$$

Where $SpatialSize_i$ is the outcome variable that measures the spatial size of the urban area of the city $i$, $FAR.Max_i$ is our variable of interest that captures the adopted maximum-allowed FAR by city $i$ and, $X_i'$ corresponds to the vector of controls that includes the main theoretical drivers of urban spatial size (average income, population size and, rural income) and the geographical controls. The term $\delta_s$ is the state fixed-effects, which controls non-parametrically for unobserved state-level variables fixed in time and that affect the size of cities. Finally, $\varepsilon_i$ is the error term. The key parameter is $\beta$, which measures the effect of the maximum allowed FAR regulation on the spatial size of cities conditioned by our set of controls. In our main regressions, we use heteroscedasticity-robust standard errors.

## 4.2 Sensitivity Analysis

For the parameter of interest in equation 7 to be correctly identified, it is necessary to assume the validity of the conditional independence assumption (CIA). The CIA implies that after conditioning the control variables, the variable of interest is entirely independent of potential outcomes. However, the possible existence of unobserved variables (confounders) that affect the adopted maximum FAR and the spatial size of the urban area can violate the CIA. If the CIA is violated due to the existence of a confounder, we will have an omitted variable bias that can bring our parameter of interest to zero. Although the CIA is untestable by construction, it is possible to check how the coefficient of interest is changed due to specific violations in the CIA using a set of sensitivity tests. The present paper will use the sensitivity analysis framework developed by Cinelli and Hazlett (2020).

Defining as bias the difference between the estimate of the parameter of interest empirically obtained (given by $\hat{\beta}$) and the corresponding estimate if there was the hypothetical inclusion of unobserved variables (given by $\hat{\beta}_{\prime}$), Cinelli and Hazlett (2020) derived the following expression for the bias generated by omitted variables:

$$|\hat{bias}| = |\hat{\beta} - \hat{\beta}_{\prime}| = se(\hat{\beta})\sqrt{\frac{R^2_{Y\sim Z|X,D}R^2_{D\sim Z|X}}{1 - R^2_{D\sim Z|X}}df} \quad (8)$$

Where $Y, D, Z$ and $X$ denote the outcome variable, the variable of interest, the set of unobservable variables, and the set of observable variables, respectively. The term $se(\hat{\beta})$ is the standard deviation of $\hat{\beta}$, and $df$ is the number of degrees of freedom. $R^2_{Y\sim Z|X,D}$ is the proportion of the variation of the outcome variable that is explained by $Z$ after the inclusion of $X$ and $D$, and $R^2_{D\sim Z|X}$ is the proportion of the variation of the variable of interest that is explained by $Z$ after the inclusion of $X$. Although it is not possible to directly estimate the value of $R^2_{Y\sim Z|X,D}$ and $R^2_{D\sim Z|X}$, equation 8 allows us to evaluate the strength that a confounder must have to bring our coefficient of interest to zero and allows the construction of sensitivity intervals.

With this framework, Cinelli and Hazlett (2020) recommend the use of three summary statistics to characterize the potential fragility of an OLS result in the face of threat confounders: I) The partial $R^2$ of the treatment with the outcome (denoted by $R^2_{Y\sim D|X}$) which can be interpreted as the value of $R^2_{D\sim Z|X}$ needed to bring the estimated effect to zero considering an extreme scenario where the confounder explains 100% of the residual variation of the outcome; II) Robustness Value (denoted by $RV$) that measures the minimum explanatory power (in %) that a confounder needs to have simultaneously with $Y$ and with $D$ to bring the coefficient of interest to zero; III) Robustness Value where the coefficient of interest is no longer statistically different from zero at a confidence level equivalent to $\alpha$ (denoted by $RV_\alpha$).

In general, these summary statistics reveal the strength required for the existence of confounders to be problematic and fully explain the results of the OLS estimation. However, a practical difficulty is knowing whether the calculated values are high or low. To clarify this point, Cinelli and Hazlett (2020) suggest using relative claims: to assess whether the unobservable variable has a greater explanatory force compared to the observable variables that are important in the research context.

In this sense, we use the average income as a benchmark variable to compare with the potential confounder. In our application, it is not easy to imagine any confounder that is as strong as income in explaining the size of the urban area and, at the same time, in the adoption of stricter FAR regulations. As seen in section 2, income is one of the main theoretical mechanisms to explain urban spatial size, since it drives the demand for housing and, consequently, expands urban land use. This result is empirically corroborated in Figure 2. Furthermore, previous empirical evidence show that the local income is one of the main factors that guide the adoption of strict land-use regulations in Brazilian municipalities (Lima and Silveira-Neto, 2019; Avila, 2006). Our data also show that average income negatively correlates with maximum allowed FAR regulation with a Pearson correlation coefficient of -0.212.

## 4.3 Instrumental Variable Estimation

We also rely on instrumental variables approach as an alternative way of dealing with endogeneity concerns. We adopt the local proportion of homeowners among the high-income households[7] as an instrument for maximum-allowed FAR. The rationale behind this instrument is the homevoter hypothesis developed by Fischel (2005). According to that, homeowners are politically engaged and influence local government, including the rules of zoning ordinances. In addition, as houses are their main assets and are not easily diversifiable, homeowners try to act to counteract risks. In this way, homeowners support projects and policies that appreciate the value of their houses and tend to be against projects that have the potential to devalue their properties. The group of homeowners belonging to the higher income categories has a more remarkable ability to influence local policy direction. Documentary evidence for Brazil indicates that elites played a crucial role in the historical formulation of zoning laws (Nery Júnior and Villaça, 2002).

Land-use regulations are the main practical instruments to accommodate the desire of homevoters in the intended patterns of urban development (Been, Madar and McDonnell, 2014). Traditionally, homevoters are against high density due to the negative externalities generated by high-rise buildings (increased housing supply, reduced open spaces, and increased congestion, for example). Therefore, they vote in favor of more restrictive land-use regulations. This behavior is commonly known as "Not in my back yard" or NIMBY. However, it is also possible to observe the exact opposite: homevoters may have the perception that new developments and denser buildings can trigger positive externalities (creation of jobs or increase the local tax collection, for example) that appreciate their homes and, in that sense, they vote in favor of flexible land-use regulations. This behavior is known as 'Yes, in my backyard' or YIMBY.

To estimate instrumental variable regression, we adopted the two-stage least

[7] This measure is calculated by the ratio between the number of high-income homeowners and the total number of high-income households.

squares method (2SLS). The first stage regression is given by the following specification:

$$FAR.Max_i = \eta Homeowners_i + \rho X_i^{'} + \delta_s + u_i \tag{9}$$

Where $Homeowners_i$ is the proportion of homeowners among the high-income group in city i. After estimating equation (1), we obtain the predicted values of the dependent variable (denoted by $\widehat{FAR.Max_i}$) and replace it with the endogenous variable of the second-stage equation that is similar to equation (7).

The relevance and exogeneity conditions must hold for an instrument to be valid. The relevance condition indicates that the instrumental variable must be correlated with the endogenous variable conditionally to other controls. The validity of this assumption can be easily checked through the statistical significance of the coefficient $\eta$ of equation (8). The exogeneity indicates that the instrumental variable cannot be correlated with the error term conditionally to other controls and thus, does not directly affect the outcome variable. Although exogeneity is an untestable assumption, we believe that the local share of homeowners is an exogenous instrument in our setting. The spatial size of cities fundamentally depends on the dynamics associated with structural factors (such as income, population size, and local geography), which interest groups do not easily drive. Therefore, we expected that the proportion of homeowners does not directly affect the outcome variable and does not correlate with the error term. The only way homevoters can (indirectly) affect the size of urban areas is their role in driving urban planning policies. Therefore, a clear threat to the validity of the exogeneity condition is the possibility that homeowners guide the adoption of land-use restriction policies that are alternative to the maximum-allowed FAR with the potential to affect the urban spatial area. We addressed this concern by including a set of alternative land-use instruments as control variables in our second-stage regression.

# 5 Results

## 5.1 Main Results

**OLS Results.** Table 2 presents the results of the estimation of equation (7) by OLS. In column (1), we report a simple linear regression with only the inclusion of the maximum-allowed FAR variable, in column (2) with the addition of the theoretical determinants of the spatial size of cities, in column (3) with the addition of state fixed effects and, finally, in column (4) with the addition of geographic control variables. We used heteroscedasticity-robust standard errors.

[*Table 2 here*]

From Table 2, it is observed that the coefficient of interest is negative and statistically significant in all specifications, indicating that cities that implement a lower maximum-allowed FAR have larger urban areas. This result is consistent with the theoretical model of Bertaud and Brueckner (2005), which predicts that the imposition of stricter height restrictions in the city center encourages the occupation of distant areas and, consequently, increases the use of urban land. In the complete specification (column (4)), we note that the reduction of one standard deviation in the maximum allowed FAR (equivalent to 2.89) generates an increase of 1.427% in the size of the urban area. In addition, we note that the average household income (elasticity of 0.34%) and the size of the population (elasticity 0.78%) generate the expected effects on the size of the city's spatial area.

As discussed earlier, urban growth boundaries (UGBs) are the land-use instruments most used by policymakers to contain the growth of urban sprawl. We checked the effect of UGBs on the spatial size of Brazilian cities using equation (7). We adopted as a variable of interest a dummy that assumes 1 for cities that implement a UGB and 0 otherwise. Table A1 in the appendix reports the results of the OLS estimation. The coefficient associated with UGB is not statistically different from zero in most specifications, suggesting that this policy is ineffective in containing urban sprawl.

**OLS Sensitivity Analysis.** To investigate whether the results of Table 2 are sensitive to the presence of unobservable confounders, Panel A of Table 3 presents the summary statistics proposed by Cinelli and Hazlett (2020). The partial $R^2$ of the treatment with outcome shows that in an extreme scenario in which unobservable confounders explain all the residual variation in the spatial size of cities, these confounders must explain at least 1.48% of the residual variation of the maximum allowed FAR to bring the coefficient of interest to zero. The robustness value indicates that the unobservable confounders that explain 11.53% of the residual variance of both the maximum allowed FAR and the size of the urban area are strong enough to fully explain the results obtained in Table 3.

[*Table 3 here*]

To determine if these values are reasonably high in our research context, we build relative claims associated with the average household income covariate. As discussed in subsection 4.2, it is difficult to imagine an unobservable variable as strong as income in its ability to explain variations in maximum-allowed FAR and in the spatial size of cities. Panel B of Table 3 shows the bounds constructed based on the average income's strength (1x, 2x, and 3x). Both $R^2_{D\sim Z|X}$ and $R^2_{Y\sim Z|X,D}$ are simultaneously lower than the Robustness Value. This indicates that even confounders up to three times stronger than the average income cannot bring our OLS coefficient of interest to zero. Furthermore, it is noted that the value of $R^2_{D\sim Z|X}$ of an unobservable variable as strong as average income is equal to 1.14%. This value is lower than the partial $R^2$ of the treatment with the outcome (1.48%). In this way, we can conclude that the existence of an unobservable variable that explains 100% of the residual variation in the urban area and is strongly associated with the FAR regulation as average income would not overturn the results of Table 2.

Figure 3 shows contour plots that are useful for checking the coefficient of interest and the t-value when we modify the confounder strength concerning the treatment variable (horizontal axis) and the outcome variable (vertical axis). We note that only in the hypothetical case of the existence of a confounder that is three times stronger than the average household income, our coefficient of interest is no longer statistically significant at a level of 10%.

[*Figure 3 here*]

In summary, this set of sensitivity results shows that simply confounders do not drive the OLS estimated coefficient. Although sensitivity analysis cannot demonstrate a causal relationship, it gives us some degree of confidence that the relationship between maximum allowed FAR and urban area size is not a simple spurious correlation.

**2SLS Estimates.** Table 4 reports the results of the 2SLS estimations in which we used the local proportion of homeowners among the high-income households as an instrument for maximum-allowed FAR regulation. To measure high-income group, we adopted three concepts existing in the 2010 Brazilian Demographic Census: the share (%) of homeowners among households with a monthly income of more than ten minimum wages (columns (1) and (2)), with more than 20 minimum wages (column (3) and (4)) and with more than 30 minimum wages (column (5) and (6)). Panel A of Table 4 presents the results of the second stage equation considering specifications containing the complete control set and alternative specifications, including the land-use regulatory index. Panel B of Table 4 shows the respective first-stage equations.

[*Table 4 here*]

Initially, the second-stage estimations of Panel A in Table 4 indicate that in all 2SLS specifications, there is a negative and statistically significant relationship between the maximum-allowed FAR regulation and the spatial size of cities. Estimates do not change when we include the regulation index as an additional control. This set of results shows that the relationship between FAR regulation and the spatial size of urban areas is robust to different identification strategies. Note that the 2SLS coefficients are higher than those obtained by OLS: the reduction of one standard deviation in the maximum-allowed FAR generates an average increase of about 12.71% in the urban area of the cities (column (6)). However, the comparison of OLS and 2SLS results is not straightforward. The OLS estimate captures the average effect of the maximum-allowed FAR for the entire set of cities (average treatment effect). Conversely, the IV/2SLS estimate captures the average effect of the FAR regulation for just the subset of cities where the local share of homeowners effectively shifts the value of the maximum-allowed FAR (*local* average treatment effect). Additionally, the OLS coefficient may be biased downward because it does not consider unobserved variables.

In terms of magnitude, the coefficient estimated in Panel A of Table 4 is lower than that found in previous studies. For example, Brueckner and Sridhar (2012) evaluated the effects of maximum-allowed FAR on the spatial size of Indian cities. It showed that reducing the maximum FAR by one unit increases the average spatial area by about 19%. For the United States, Geshkov and DeSalvo (2012) show that counties that adopt the maximum-allowed FAR as a regulatory instrument have an average increase in urban areas of about 24.65%. Behind the methodological differences, this comparison highlights the importance of evaluating the economic consequences of land use regulations in different contexts and suggests that the effects of maximum FAR on the spatial size of cities are relatively modest in Brazil.

The first stage estimates from Panel B of Table 4 show a positive and statistically significant correlation between the local share of homeowners and the maximum allowed FAR value in all specifications, suggesting that the relevance condition of IV holds. Furthermore, this evidence indicates that cities with a more substantial home voter base are more likely to adopt looser land-use regulations, meaning that YIMBY is the prevailing behavior in Brazilian cities. The regulatory index coefficient has the expected sign: cities with stricter regulatory environments implement a lower maximum-allowed FAR.

The first stage Kleibergen-Paap F statistic is lower than the usual rule of thumb (equal to 10, as shown by Stock and Yogo (2002)) in the specifications of columns (1) to (4), suggesting the possibility of weak instruments. To check if the problem of weak instruments is a concern in our setting, we followed the recommendation of Andrews, Stock and Sun (2019) and calculated the Anderson-Rubin (AR) test that evaluated the null hypothesis that the coefficient of the endogenous variable is equal to zero. We also report the identification-robust AR confidence intervals for the coefficient of interest by inverting the AR test statistic. Both procedures are fully robust to weak instruments. The results at the bottom of Table 4 show that the AR test null hypothesis is rejected in all specifications and that all confidence intervals exclude zero. Anyway, to avoid the weak instrument problem, we will adopt the specification in column (6) as our preferred 2SLS estimation in the remainder of the paper.

## 5.2 Robustness Checks

In this subsection, we will check if our main results (OLS and IV/2SLS) are maintained when we change the empirical specification, remove outlier cities from the sample, or modify the way of performing inference. Additionally, we also evaluate how 2SLS estimates vary when considering alternative estimation strategies or adopting an internal instrumental variables approach. Table 5 reports the coefficient of interest for each robustness exercise common to OLS and 2SLS For comparison purposes, we also present the baseline coefficients.

[*Table 5 here*]

**Different Empirical Specifications.** As discussed in subsection 3.2, we deal with cities with unrestricted FAR regulation through the imputation of values. An alternative strategy is to remove these cities from the sample. Estimation A.1 of Table 5 shows OLS and IV/2SLS results removing the set of cities without FAR regulation. We note that, although the effects are robust, the magnitude and imprecision of the coefficients increase considerably. We also evaluated the robustness of the results relative to alternative specifications to equation (7). Estimation A.2 of Table 5 presents a specification in which the dependent variable is measured in level instead of logarithm. Estimation A.3 shows a specification with the inverse hyperbolic sine of the dependent variable. Estimation A.4 uses the log of the maximum-allowed FAR as our variable of interest. Finally, in estimation A.5, we estimate equation (7) through weighted least squares using the population city size as observation weight. Overall, we observed that the OLS and 2SLS coefficients are robust to alternative specifications.

**Dropping outlier cities.** To check if our results are driven by a specific group of cities (such as large metropolises), we also perform the estimations by dropping outlier observations. Estimation B.1 of Table 5 presents the main results of OLS and 2SLS dropping the state capitals, the estimation B.2 dropping cities with more than one million inhabitants, estimation B.3 removing the cities with the 10% largest urban areas, and B.4 removing the cities with the 10% smallest urban areas from the sample. The main results are robust to the existence of outlier observations.

**Alternative ways of inference.** We perform inference using heteroscedasticity-robust standard errors in our main specifications. However, there are no clear and consensual guidelines regarding practice for performing inference in applied econometrics. Thus, Panel C of Table 5 presents alternative ways of inference. The estimates of C.1 and C.2 present clustered standard errors at the state and macroregion[8] levels, respectively. Estimations C.3 and C.4 present the Conley (1999) standard errors, which consider the spatial correlation of the data. In these specifications, we used two cut-off distances for the cities: 50km and 100km. It is possible to notice in Panel C of Table 5 that, regardless of the form of inference, the standard errors of the OLS and 2SLS estimates do not vary significantly.

**Internal Instrumental Variable Approach.** Considering that the validity of the exogeneity assumption can be easily challenged in empirical research, we evaluate whether our IV/2SLS results are robust when exploiting an internal instrumental variable. For this purpose, we adopt Lewbel (2012) approach, which involves exploring the potential heteroscedasticity of first-stage regression errors to build an internal instrument. The Lewbel (2012) method identifies structural parameters through control variables that do not correlate with the product of heteroscedastic errors. Therefore, the first-stage errors must be heteroscedastic for the model to be identified. Column (1) of Table 6 presents the results of the second stage of the 2SLS estimation using the internal instrument generated by heteroscedasticity, and column (2) presents the results of the 2SLS estimation considering both the internal instrument and the local share of homeowners among high-income households (exogenous instrument). The Breush-Pagan test at the bottom of Table 6 shows that the null hypothesis of homoscedasticity is rejected in both specifications. Furthermore, we note that our coefficient of interest remains negative and statistically significant.

**Alternative IV Estimators.** We also check whether our IV/2SLS results are robust against alternative 2SLS estimators. In this sense, we used two estimators of the k-estimator class: the Fuller (1977) estimator and the limited information maximum likelihood (LIML) estimator. It is noted that LIML is more robust compared to 2SLS when there is a weak instrument concern (Stock, Wright and Yogo, 2002). Columns (3) and (4) show the results of estimating IV using these alternative estimators. Note that the coefficients of interest are very similar to those observed in Table 4.

[*Table 6 here*]

[8]Brazil has five macro-regions: North, Northeast, Midwest, Southeast, and South.

# 6 Estimating the welfare costs of FAR

Our estimation closely follows the framework devised by Brueckner and Sridhar (2012) for India but is adjusted for the parameters of the current Brazilian economy. Reducing the maximum FAR by one standard deviation would result in an expansion of the area of the city by 12.4%. Taking the average area of the sample cities (64.5 km$^2$), this would increase of 8 km$^2$ (a $\approx$ 300 meter extension of the radius of a circular city. This increase in the area reduces the welfare of passengers by making travel longer and more expensive for residents of the urban periphery.

The estimation of FAR costs is focused on three components: the monetary cost related to transportation, the increase in CO2 emissions, and the health cost of additional pollution. Table 7 summarizes our estimates and Table A2 in the Appendix presents the parameters and their sources.

**Transport costs.** Based on parameters based on literature and data, we estimate that the annual transportation cost for each household is around US$ 81. This value has two components: one related to the bus fare and the other from the loss of welfare from commuting. Data from official guidelines suggest that urban fares are around US$0,0054 per passenger kilometer (see Table A2)[9]. Brueckner and Sridhar (2012) assume that commuting costs are 60% of the hour wage. Using Brazilian data, results in US$ 2.75 per hour or US$81 per year per household (See panel A of Table 7) .

The annual welfare loss for a household living in the outer ring of the city (US$23.67) equals the increase of the radius of the city (0.29 km) times the annual commuting costs per household kilometer (US$81.24). Therefore, the total annual welfare loss for an average city of our sample is US$23.7 times 70,637 households, i.e. US$ 1.9 million.

[*Table 7 here*]

**Health costs** Transport is behind almost all fine particulate air pollution in contemporary cities. The Global Model of Ambient Particulates (GMAPS) (Cohen et al., 2005) enabled Miraglia and Gouveia (2014) estimating air pollution in several Brazilian cities. These values were used as input to calculate years of life lost (YLL), years lived with disability (YLD), and the value of life-year losses. They came up with an estimated value of US$ 1.7 billion/year of losses due to urban air pollution for the cities in their sample. This corresponds to an annual cost of US$21.75 per capita and a total of US$ 6.9 million for an average city of our sample.

In the literature on market potential, the 'own distance' is the average distance from a random point to the CBD, assuming a circular city (Keeble, Owens and Thompson, 1982; Overman, Redding and Venables, 2003). In its simplest form, it is equal to half of the radius. Therefore, we will assume that an increase in the radius would increase commuting trips in the same proportion, as well as air pollution and health costs.

Adapting these estimates to the average city of our sample, we conclude that the health costs of reducing the maximum FAR by one standard deviation for the typical city of our sample are in the order of roughly US$ 449 thousand per year. The details are on panel B of Table 7).

[9]This value is probably underestimated because it takes oil prices as the only variable cost.

**CO2 emissions costs** In addition to particulate air pollution, CO2 emissions is a major negative externality generated by urban transport. We chose the price of 60 euros per ton of CO2 based on the guidelines of the most recent OECD report on carbon pricing (OECD, 2021). The CO2 emissions of Brazilian urban transport were based on estimates of Carvalho (2011). Although his calculations are more than a decade old, we judge them appropriately since there has been no significant technological or structural change in transportation in Brazil.

Carvalho (2011) estimates that Brazil's CO2 emissions per passenger per kilometer in Brazil are around 0.0609 kg or US$ 0.004. Annually, this corresponds to US$0.80. Assuming that each household has around 1.6 workers and following the same logic as the estimates of the "own distance", this means that the annual cost of CO2 emissions for the average city is US$463,298 (US$0.80 * 4.5 km * 79,637 households * 1.6). So our estimates of the impact of raising the FAR by one sd on the city radius would result in additional emissions of 0,5 tons of CO2 per year or US$ 29,774. Panel C of 7 shows the parameters and the calculations.

Adding up the costs of transportation, health care, and CO2 emissions, we estimate that a one standard deviation reduction of the maximum FAR would result in annual welfare losses of US$ 2.36 million for an average city.

# 7 Conclusion

In this paper, we analyzed the impact of FAR regulations on the urban area of Brazilian cities. We developed a novel database for the 325 larger Brazilian cities with information on building-height restrictions and combined it with satellite data that captured the local urban areas. Exploiting the local share of homeowners among the high-income households group as a source of exogenous variation for FAR stringency, our IV/2SLS estimates suggest that the decrease of one standard deviation in the maximum-allowed FAR increases the spatial area of a city by 12.4%. Thus, our results indicate that the theoretical implications of Bertaud and Brueckner (2005) are valid in Brazil, a middle-income country where zoning is not strictly enforced.

The relationship between FAR regulations and the spatial size of cities is robust to OLS estimates, the sensitivity framework proposed by Cinelli and Hazlett (2020), different empirical specifications, the existence of outlier cities, and alternative ways of performing inference. Additionally, the IV/2SLS estimates are robust to alternative estimators (LIML and Fuller) and the identification proposed by Lewbel (2012) that uses the heteroscedasticity of errors to build an internal instrument.

Bertaud and Brueckner (2005) considered the commuting cost of urban sprawl caused by FAR regulation, but ignored the effects of additional pollution. We overcame this limitation by including the cost of additional CO2 emissions and health costs. Our estimations relied on international criteria and Brazilian data. We estimate that one standard deviation reduction of the maximum FAR would result in welfare losses equivalent to US$ 2.36 million per year in a average city of our sample.

70 million Brazilians live in cities with more than 300 thousand inhabitants. Therefore, although the estimated value of losses is somewhat low for a typical municipality in our sample, extrapolating the estimated losses for the rest of the population would lead to much more considerable values.

We should be cautious about drawing specific policy recommendations from our estimates. Urban zoning is a complex issue, and changes may have long-term consequences that can hardly be predicted. Furthermore, the parameters we use are not free from criticism. Nevertheless, we present a framework that provides an empirical basis for a modern discussion on zoning. We also show that although the controversy on NIMBYism is centered on cities of developed countries, such a theme is important for Brazil and, most likely, other developing economies.

## References


**ABRATI** (2015). "Resolução n. 4768," URL: http://www.abrati.org.br/resolucao/resolucao-no-4-768-de-25-06-2015-dou-30-06-2015/.

**Andrews, Isaiah, James H Stock, and Liyang Sun** (2019). "Weak instruments in instrumental variables regression: Theory and practice," *Annual Review of Economics*, 11(1): .

**ANP** (2018). "Síntese dos Preços Praticados - Brasil," URL: https://preco.anp.gov.br/include/Resumo_Mensal_Combustiveis.asp.

**Avila, Paulo C** (2006). "Urban land use regulations in Brazilian cities: impacts on urban land markets and access of low income people to land and housing," *Brazil: inputs for a strategy for cities–a contribution with a focus on cities and municipalities. Washington: The World Bank*, 1.

**Been, Vicki, Josiah Madar, and Simon McDonnell** (2014). "Urban Land-Use Regulation: Are Homevoters Overtaking the Growth Machine?" *Journal of Empirical Legal Studies*, 11(2): 227–265.

**Bertaud, Alain and Jan K Brueckner** (2005). "Analyzing building-height restrictions: predicted impacts and welfare costs," *Regional Science and Urban Economics*, 35(2): 109–125.

**Brueckner, Jan K and David A Fansler** (1983). "The economics of urban sprawl: Theory and evidence on the spatial sizes of cities," *The review of Economics and Statistics* 479–482.

**Brueckner, Jan K and Kala Seetharam Sridhar** (2012). "Measuring welfare gains from relaxation of land-use restrictions: The case of India's building-height limits," *Regional Science and Urban Economics*, 42(6): 1061–1067.

**Carvalho, Carlos Henrique Ribeiro de** (2011). "Emissões relativas de poluentes do transporte urbano."

**Cinelli, Carlos and Chad Hazlett** (2020). "Making sense of sensitivity: Extending omitted variable bias," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1): 39–67.

**Cohen, Aaron J, H Ross Anderson, Bart Ostro, Kiran Dev Pandey, Michal Krzyzanowski, Nino Künzli, Kersten Gutschmidt, Arden Pope, Isabelle Romieu, Jonathan M Samet et al.** (2005). "The global burden of disease due to outdoor air pollution," *Journal of Toxicology and Environmental Health, Part A*, 68(13-14): 1301–1307.

**Conley, Timothy G** (1999). "GMM estimation with cross sectional dependence," *Journal of econometrics*, 92(1): 1–45.

**Dempsey, Judith A and Andrew J Plantinga** (2013). "How well do urban growth boundaries contain development? Results for Oregon using a difference-in-difference estimator," *Regional Science and Urban Economics*, 43(6): 996–1007.

**Deng, Xiangzheng, Jikun Huang, Scott Rozelle, and Emi Uchida** (2008). "Growth, population and industrialization, and urban land expansion of China," *Journal of Urban Economics*, 63(1): 96–115.

**Fischel, William A** (2005). *The homevoter hypothesis: How home values influence local government taxation, school finance, and land-use policies*, Harvard University Press.

**Fuller, Wayne A** (1977). "Some properties of a modification of the limited information estimator," *Econometrica: Journal of the Econometric Society* 939–953.

**Geshkov, Marin V and Joseph S DeSalvo** (2012). "The effect of land-use controls on the spatial size of US urbanized areas," *Journal of Regional Science*, 52(4): 648–675.

**Hortas-Rico, Miriam** (2015). "Sprawl, blight, and the role of urban containment policies: Evidence from US cities," *Journal of Regional Science*, 55(2): 298–323.

**Howell-Moroney, Michael** (2007). "Studying the effects of the intensity of US state growth management approaches on land development outcomes," *Urban Studies*, 44(11): 2163–2178.

**IBGE** (2018). "Censo 2010 - Resultados," URL: `https://censo2010.ibge.gov.br/resultados.html`.

**IPEA** (2022). "IpeaData," URL: `http://www.ipeadata.gov.br/`.

**Keeble, David, Peter L Owens, and Chris Thompson** (1982). "Regional accessibility and economic potential in the European Community," *Regional Studies*, 16(6): 419–432.

**Lewbel, Arthur** (2012). "Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models," *Journal of Business & Economic Statistics*, 30(1): 67–80.

**Lima, Ricardo Carvalho de Andrade and Raul da Mota Silveira-Neto** (2019). "Zoning ordinances and the housing market in developing countries: Evidence from Brazilian municipalities," *Journal of Housing Economics*, 46 p. 101653.

**McGrath, Daniel T** (2005). "More evidence on the spatial scale of cities," *Journal of Urban Economics*, 58(1): 1–10.

**MDR** (2018). "Pesquisa Nacional de Mobilidade Urbana," URL: `https://antigo.mdr.gov.br/images/SEMOB/Relatorio_Pemob_2018.pdf`.

**Miraglia, Simone Georges El Khouri and Nelson Gouveia** (2014). "Custos da poluição atmosférica nas regiões metropolitanas brasileiras," *Ciência & Saúde Coletiva*, 19 4141–4147.

**Moon, Byunggeor and Sungin Ahn** (2022). "The effects of a FAR regulation in a model of durable building with redevelopment: The case of New York City," *Regional Science and Urban Economics*, 95 p. 103775.

**MTE** (2019). "Relatório Anual de Informações Sociais," URL: `http://pdet.mte.gov.br/rais/rais-anos-anteriores`.

**Nechyba, Thomas J and Randall P Walsh** (2004). "Urban sprawl," *Journal of economic perspectives*, 18(4): 177–200.

**Nery Júnior, José Marinho and Flávio José Magalhães Villaça** (2002). "Um século de política para poucos: zoneamento paulistano 1886-1986."

**OECD** (2021). *Effective Carbon Rates 2021*, 40.

**Overman, Henry G., Stephen Redding, and Anthony J. Venables** (2003). "The Economic Geography of Trade, Production, and Income: A Survey of Empirics," in *Handbook of International Trade*, John Wiley & Sons, Ltd 350–387, Section: 12 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470756461.ch12.

**Paulsen, Kurt** (2012). "Yet even more evidence on the spatial size of cities: Urban spatial expansion in the US, 1980–2000," *Regional Science and Urban Economics*, 42(4): 561–568.

**Paulsen, Kurt** (2013). "The effects of growth management on the spatial extent of urban development, revisited," *Land Economics*, 89(2): 193–210.

**Quigley, John M and Steven Raphael** (2005). "Regulation and the high cost of housing in California," *American Economic Review*, 95(2): 323–328.

**Santos, Angelo Fernando Gonzaga dos** (2020). "Determinantes da expansão urbana no Brasil," Master's thesis, Universidade Federal de Pernambuco.

**Spivey, Christy** (2008). "The Mills—Muth model of urban spatial structure: Surviving the test of time?" *Urban Studies*, 45(2): 295–312.

**Stock, James H, Jonathan H Wright, and Motohiro Yogo** (2002). "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics*, 20(4): 518–529.

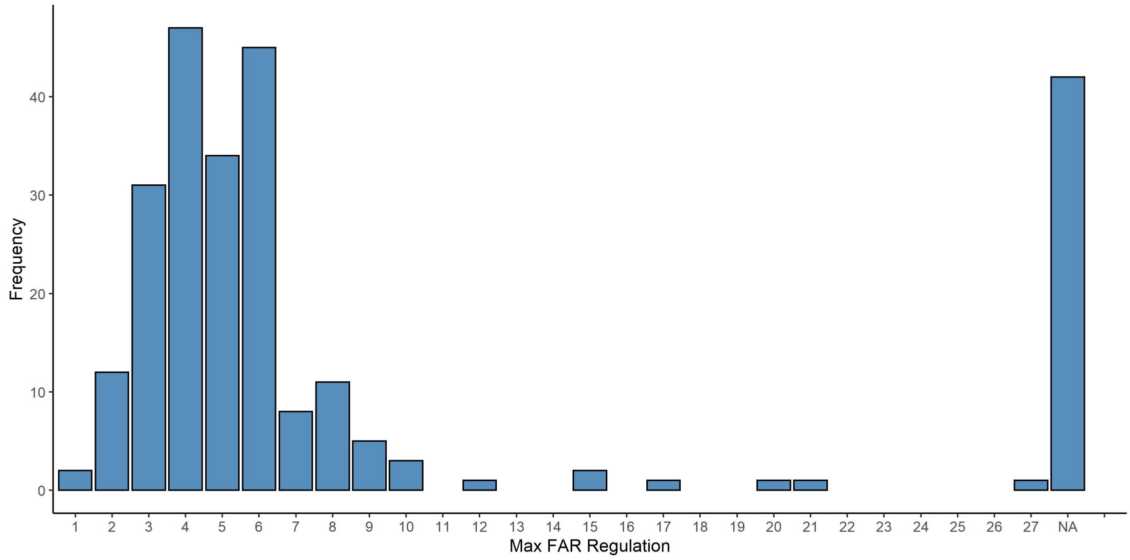
**Stock, James H and Motohiro Yogo** (2002). "Testing for weak instruments in linear IV regression."

**Wassmer, Robert W** (2006). "The influence of local urban containment policies and statewide growth management on the size of United States urban areas," *Journal of Regional Science*, 46(1): 25–65.

**Wheaton, William C** (1974). "A comparative static analysis of urban spatial structure," *Journal of Economic Theory*, 9(2): 223–237.

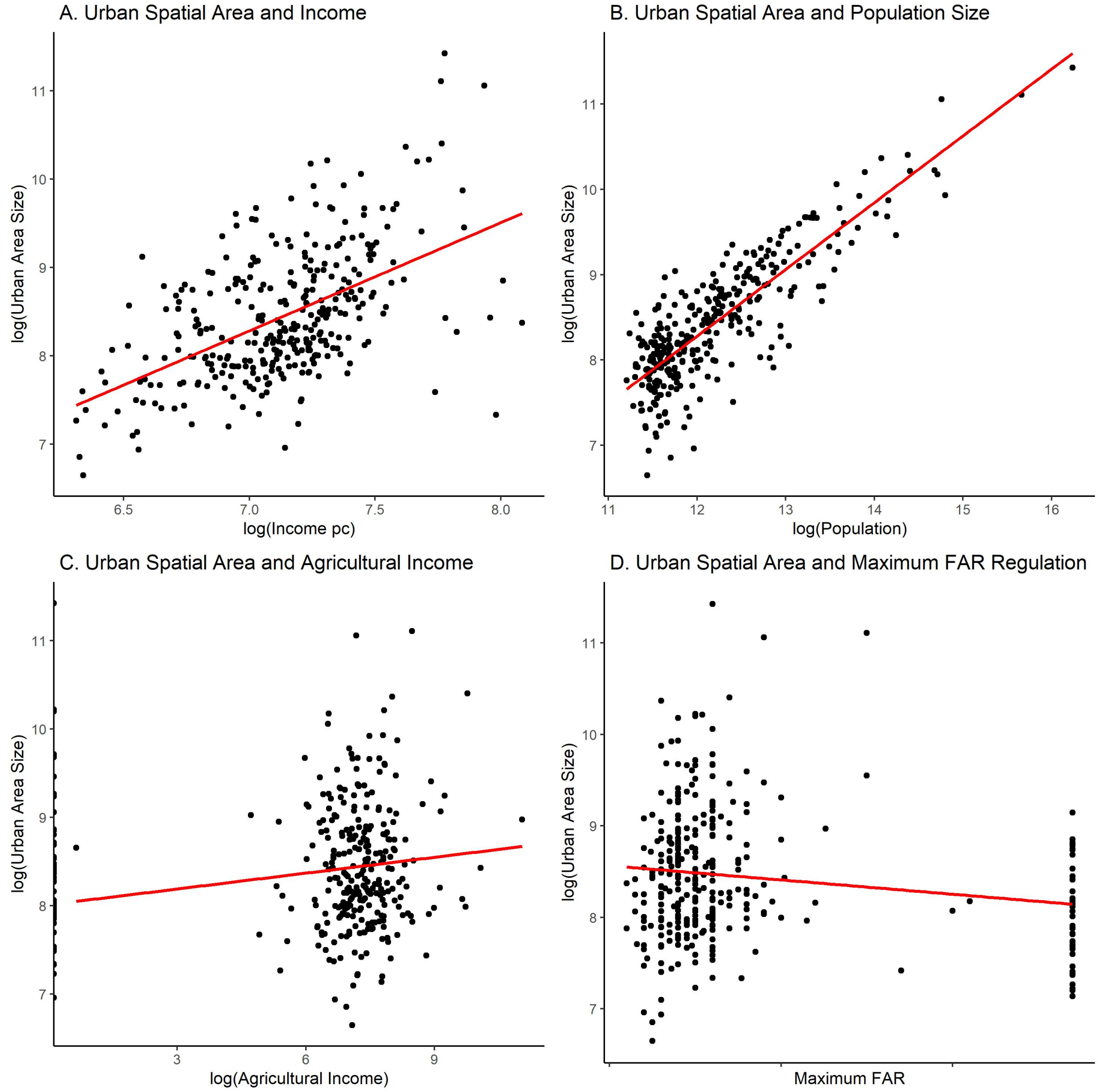
**Zhou, Yan, Xianjin Huang, Yi Chen, Taiyang Zhong, Guoliang Xu, Jinliao He, Yuting Xu, and Hao Meng** (2017). "The effect of land use planning (2006–2020) on construction land growth in China," *Cities*, 68 37–47.

**Figure 1:** Distribution of FAR Regulation in Brazilian Cities.



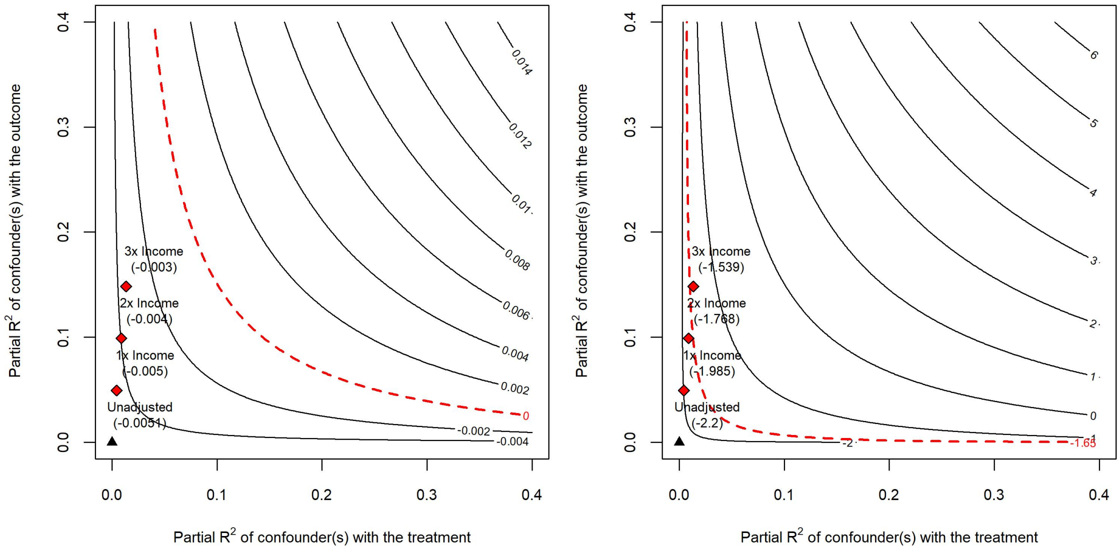

Note: The figure shows the histogram of maximum-allowed FAR values among the 325 largest Brazilian cities.

**Figure 2:** The Determinants of the Spatial Size of Cities.




Note: The upper left plot (A) shows the relationship between the log of urban area size and the log of household income. The upper right-hand plot (B) shows the relationship between the log of urban area size and the log of population size. The plot on the lower left (C) shows the relationship between the log of urban area size and the log of agricultural income. Finally, the plot on the lower right side (D) shows the relationship between the log of urban area size and the maximum FAR value. Each point represents a city, and the red line represents a simple linear regression line.

**Figure 3:** Sensitivity Contour Plots of Coefficient of Interest and t-value.




Notes: These plots shows the sensitivity contour plots for the FAR regulation using the average household income as a reference for bounds of unobserved confounders. The figure on the left shows the contour plot for the point estimate, and the figure on the right shows the contour plot for the t-statistic.

**Table 1:** Summary Statistics

| **A. Main Variables** | **Mean** | **SD** | **Min** | **Max** |
|---|---|---|---|---|
| Maximum-allowed FAR | 5.172 | 2.897 | 1 | 27 |
| Urban Spatial Size (ha) | 6451.296 | 8291.22 | 6992.7 | 91708.73 |
| Population Size (mil, 2010) | 321.22 | 776.91 | 36.30 | 11152.34 |
| Household Income pc (R$, 2010) | 1310.772 | 428.007 | 550.9 | 3246.05 |
| Agricultural Income (R$, 2010) | 1911.446 | 4130.218 | 0 | 62337.608 |
| **B. Geographical Controls** | **Mean** | **SD** | **Min** | **Max** |
| Distance to Capital City (km) | 148 | 160.124 | 0 | 888 |
| Average Altitude (m) | 395.526 | 354.28 | 1 | 1196 |
| Average Terrain Ruggedness Index | 7.504 | 4.544 | 1.313 | 24.032 |
| Number of Conservation Units | 3.471 | 7.422 | 0 | 89 |
| Hydrographic Area (ha) | 8244.886 | 33361.329 | 0 | 366070.60 |
| Coastal City (0/1) | 0.175 | 0.381 | 0 | 1 |
| Metropolitan City (0/1) | 0.4277 | 0.4955 | 0 | 1 |
| **C. Instrumental Variables** | **Mean** | **SD** | **Min** | **Max** |
| % of Homeowners with more than 30 MW | 0.835 | 0.0722 | 0.3 | 1 |
| % of Homeowners with more than 20 MW | 0.825 | 0.0677 | 0.324 | 0.96 |
| % of Homeowners with more than 10 MW | 0.8068 | 0.0654 | 0.399 | 0.9547 |
| **C. Land-Use Regulations** | **Mean** | **SD** | **Min** | **Max** |
| Higher Minimum Lot Size (0/1) | 0.645 | 0.479 | 0 | 1 |
| Urban Growth Boundary (0/1) | 0.846 | 0.361 | 0 | 1 |
| Land Subdivision (0/1) | 0.803 | 0.398 | 0 | 1 |
| Building Code (0/1) | 0.843 | 0.364 | 0 | 1 |
| Regulatory Index | 3.135 | 1.068 | 0 | 4 |

**Table 2:** The Effect of Maximum FAR Regulation on the Spatial Size of Cities

| | log (Spatial City Size) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Max FAR | -0.0160*** | -0.0050** | -0.0053** | -0.0049** |
| | (0.0050) | (0.0024) | (0.0024) | (0.0024) |
| log (Income pc) | - | 0.5004*** | 0.4548*** | 0.3438** |
| | | (0.0834) | (0.0913) | (0.1229) |
| log(Population) | - | 0.7271*** | 0.7417*** | 0.7797*** |
| | | (0.0258) | (0.0312) | (0.0350) |
| log(Agricultural Income) | - | 0.0555*** | 0.0543*** | 0.0577*** |
| | | (0.0086) | (0.0091) | (0.0093) |
| Geographical Controls | No | No | No | Yes |
| State Fixed-Effects | No | No | Yes | Yes |
| Adjusted $R^2$ | 0.0249 | 0.7947 | 0.7962 | 0.8271 |
| Number of Observations | 325 | 325 | 325 | 325 |

Notes: *** represents $p < 0.01$,** represents $p < 0.05$,* represents $p < 0.1$. We use heteroscedasticity-robust standard errors. The standard deviations are presented in parentheses. The dependent variable is the log of urban area size. Geographical controls includes the following variables: distance to capital city, average altitude, average terrain ruggdness index, number of conservation units, hydrographic area, dummy for coastal cities and a dummy for metropolitan cities.

**Table 3:** The Sensitivity of OLS Results to Unobservable Confounders

| **A. Sensivity Statistics** | | |
|---|---|---|
| Partial $R^2$ of Treatment with Outcome | 1.48% | |
| Robustness Value | 11.53% | |
| Robustness Value, $\alpha = 0.1$ | 2.46% | |
| **B. Bounds** | $R^2_{D \sim Z\|X}$ | $R^2_{Y \sim Z\|X,D}$ |
| 1x log(Income pc) | 1.14% | 5.13% |
| 2x log(Income pc) | 2.28% | 10.26% |
| 3x log(Income pc) | 3.43% | 15.40% |

Notes: This sensitivity analysis is based on the coefficient of interest associated with column (4) of Table 2. The Robustness Value measures the minimum explanatory power (in %) that a confounder needs to have simultaneously with $Y$ and with $D$ to bring the OLS coefficient of interest to zero.

**Table 4:** The Effect of Maximum FAR Regulation on the Spatial Size of Cities: IV Estimates

| **Panel A. Second-Stage IV** | log (Spatial City Size) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Max FAR | -0.0675** | -0.0678** | -0.0681** | -0.0684** | -0.0440** | -0.0440** |
| | (0.0311) | (0.0316) | (0.0314) | (0.0318) | (0.0209) | (0.0206) |
| Regulatory Index | - | -0.0339 | - | -0.0347 | - | 0.0000 |
| | | (0.0614) | | (0.0627) | | (0.0434) |
| **Panel B. First-Stage IV** | Max FAR | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| % Homeowners | 0.2294** | 0.2264** | 0.2282** | 0.2259** | 0.2272*** | 0.2307*** |
| | (0.0829) | (0.0812) | (0.0806) | (0.0787) | (0.0741) | (0.0728) |
| Regulatory Index | - | -1.4104** | - | -1.4147** | - | -1.4550** |
| | | (0.5933) | | (0.5888) | | (0.5662) |
| Income Category IV | >10 MW | >10 MW | >20 MW | >20 MW | >30 MW | >30 MW |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| State Fixed-Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| KP F-Statistic - First Stage | 8.1017 | 8.0937 | 8.0588 | 8.1027 | 10.4186 | 11.0328 |
| Anderson-Rubin Test | 21.060*** | 21.2939*** | 21.3756*** | 21.7081*** | 11.0681*** | 11.7443*** |
| Weak IV Robust 95% CI | [-0.214,-0.033] | [-0.218,-0.033] | [-0.217,-0.034] | [-0.211,-0.033] | [-0.115,-0.017] | [-0.112,-0.018] |
| Number of Observations | 325 | 325 | 325 | 325 | 325 | 325 |

Notes: *** represents $p < 0.01$,** represents $p < 0.05$,* represents $p < 0.1$. We use heteroscedasticity-robust standard errors. The standard deviations are presented in parentheses. The 2SLS regressions includes the following control variables: log of population, log of income pc, log of agricultural income, distance to capital city, average altitude, average terrain ruggdness index, number of conservation units, hydrographic area, dummy for coastal cities and a dummy for metropolitan cities.

**Table 5:** Robustness Checks: alternative empirical specifications, outlier cities and alternative inference.

| | OLS | IV/2SLS | N. Observations |
|---|---|---|---|
| | (1) | (2) | (3) |
| Baseline Max FAR coefficient | -0.0051** | -0.0440** | 325 |
| | (0.002) | (0.0206) | |
| **A. Robustness to different specifications** | | | |
| A.1 - Droping cities without FAR | -0.0097* | -0.2371 | 283 |
| | (0.0058) | (0.1862) | |
| A.2 - Dependent variable in level | -52.4544** | -834.5396** | 325 |
| | (26.0562) | (301.6552) | |
| A.3 - Dependent variable in Asinh | -0.0049** | -0.0412** | 325 |
| | (0.0024) | (0.0203) | |
| A.4 - FAR variable in log | -0.0583** | -0.4334** | 325 |
| | (0.0262) | (0.2047) | |
| A.5 - Weighted Least Squares Estimator | -0.0063** | -0.0590*** | 325 |
| | (0.0025) | (0.0220) | |
| **B. Robustness to outliers** | | | |
| B.1 - Droping capital cities | -0.0053** | -0.0408* | 298 |
| | (0.0025) | (0.0219) | |
| B.2 - Droping cities > 1 million inhab. | -0.0054** | -0.0386* | 308 |
| | (0.0024) | (0.0205) | |
| B.3 - Trimming the 10% lower Urban Areas | -0.0048** | -0.0417** | 292 |
| | (0.0021) | (0.0170) | |
| B.4 - Trimming the 10% higher Urban Areas | -0.0049** | -0.0361* | 292 |
| | (0.0024) | (0.0217) | |
| **C. Alternative ways of inference** | | | |
| C.1 - Clustered standard errors by State | -0.0049** | -0.0440*** | 325 |
| | (0.0020) | (0.0146) | |
| C.2 - Clustered standard errors by Macrorregion | -0.0049*** | -0.0440*** | 325 |
| | (0.0011) | (0.0113) | |
| C.3 - Conley (1999) standard errors - 50 km | -0.0049** | -0.0440** | 325 |
| | (0.0023) | (0.0196) | |
| C.4 - Conley (1999) standard errors - 100 km | -0.0049* | -0.0440** | 325 |
| | (0.0026) | (0.0182) | |

Notes: *** represents $p < 0.01$,** represents $p < 0.05$,* represents $p < 0.1$. Only the coefficients associated with the effect of maximum-allowed FAR on the spatial size of cities are shown. The standard deviations are presented in parentheses. All the regressions includes the following control variables: log of population, log of income pc, log of agricultural income, distance to capital city, average altitude, average terrain ruggdness index, number of conservation units, hydrographic area, dummy for coastal cities and a dummy for metropolitan cities.

**Table 6:** Robustness Checks in IV/2SLS Estimations: alternative estimators and internal instrumental variable.

| | log (Spatial City Size) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Lewbel (2012) 2SLS | Lewbel (2012) and Exogenous Instrument 2SLS | Fuller IV Estimator | LIML IV Estimator |
| Max FAR | -0.0075** | -0.0076** | -0.0408** | -0.0440** |
| | (0.0038) | (0.0038) | (0.0170) | (0.0194) |
| Controls | Yes | Yes | Yes | Yes |
| State Fixed-Effects | Yes | Yes | Yes | Yes |
| Breush-Pagan Test | 16.533*** | 15.38*** | - | - |
| Number of Observations | 325 | 325 | 325 | 325 |

Notes: *** represents $p < 0.01$,** represents $p < 0.05$,* represents $p < 0.1$. The standard deviations are presented in parentheses. All the regressions includes the following control variables: log of population, log of income pc, log of agricultural income, distance to capital city, average altitude, average terrain ruggdness index, number of conservation units, hydrographic area, dummy for coastal cities and a dummy for metropolitan cities.

**Table 7:** Estimation losses from a one sd decrease in FAR

| | |
|---|---|
| **A. Area of the city** | |
| Percentage increase in city area | 12 % |
| Increase in city area | |
| 12% * 64.51 km$^2$ | 8.02 km$^2$ |
| Increase in the radius of the city | |
| $\sqrt{(64.5km + 8.02) * \pi^{-1}} - \sqrt{64.5km * \pi^{-1}}$ | 0.29 km |
| **B. Commuting costs** | |
| Hour wage US$734.17/ (40 hours*4 weeks) | US$4.59 |
| Commuting costs per hour | |
| US$4.59* 0.6 | US$ 2.75 |
| Round trip total cost per km | |
| 2 *( US$ 0.0054 + (US$ 2.75/ 25 km/h)) | US$ 0.23 |
| Annual commuting costs per household km | |
| US$0.23 * 220 days * 1.59 workers | US$ 81.24 |
| Increase in edge household's commuting cost | |
| 0.29 km * US$ 81.24 | US$ 23.67 |
| Annual welfare loss | |
| US$ 23.7 * 79,637 households | US$ 1.9 million |
| **C. Health costs** | |
| Health costs per capita | US$ 21.75 |
| Health costs of the average city | US$ 6.9 million |
| Increase in health costs | |
| US$ 21.75 * 321000 people * (0.29 km/4.53 km) | US$ 448.694 |
| **D. CO2 costs** | |
| Number of passengers | 127,183 |
| CO2 emissions per passenger/km | 0.0609 kg |
| Cost CO2 | US$ 60/ton |
| Cost CO2 per passenger km year | |
| 0.0609 kg * (US$ 60/1000) * 220 days | US$ 0.803 |
| Annual cost CO2 of emissions average city | |
| US$ 0.80 * 4.53 km * 127,173 | US$ 463,298 |
| Change CO2 cost | |
| US$ 463,298 * (0.29 km /4.53 km ) | US$ 29,774 |

# A Appendix

## A.1 Urban Growth Boundaries and Spatial Size of Cities.

Table A1 presents the results of the estimation of equation (7) using the urban growth boundary (UGB) as the variable of interest instead of the maximum-allowed FAR. Similar to Table 2, we report the results of four different specifications: column (1) shows the specification without controls, column (2) shows the specification with the inclusion of theoretical determinants of the spatial size of cities, and column (3) with the addition of state fixed effects and, finally, column (4) with the addition of geographic controls. In the specification of column (4), it is possible to see that UGBs do not exert statistically significant effects on the size of the urban area.

**Table A1:** The Effect of UGB Regulation on the Spatial Size of Cities.

| | log (Spatial City Size) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Urban Growth Boundary | 0.2735** | 0.0679 | 0.0519 | 0.0598 |
| | (0.1183) | (0.0585) | (0.0594) | (0.0602) |
| log (Income pc) | | 0.5067*** | 0.4774*** | 0.3599*** |
| | | (0.0833) | (0.0896) | (0.1234) |
| log(Population) | | 0.7287*** | 0.7416*** | 0.7790*** |
| | | (0.0256) | (0.0314) | (0.0359) |
| log(Agricultural Income) | | 0.0550*** | 0.0554*** | 0.0593*** |
| | | (0.0087) | (0.0093) | (0.0094) |
| Geographical Controls | No | No | Yes | Yes |
| State Fixed-Effects | No | No | No | Yes |
| Adjusted $R^2$ | 0.0153 | 0.7930 | 0.7938 | 0.8254 |
| Number of Observations | 325 | 325 | 325 | 325 |

Notes: *** represents $p < 0.01$,** represents $p < 0.05$,* represents $p < 0.1$. We use heteroscedasticity-robust standard errors. The standard deviations are presented in parentheses. The dependent variable is the log of urban area size. Geographical controls includes the following variables: distance to capital city, average altitude, average terrain ruggdness index, number of conservation units, hydrographic area, dummy for coastal cities and a dummy for metropolitan cities.

## A.2 Parameters for calculating the economic costs associated with stringent FAR regulation.

**Table A2:** Parameters for impact estimation

| Variable | Value | Source |
|---|---|---|
| Average area | 64.51 $km^2$ | Database |
| Average radius | 4.53 km | Database |
| Average population | 321,000 | Database |
| Average number of households | 79,637 | Database |
| Exchange rate 2019 R$/US$ | 3.95 | IPEA (2022) |
| Monthly Wage 2019 US$ (R$) | US$ 734.17 (R$ 2900) | MTE (2019) |
| Traffic speed | 25 km/h | MDR (2018) |
| Workers per household | 1.59 | IBGE (2018) |
| Price diesel US$ (R$) | US$ 0.82 (R$ 3.25) | ANP (2018) |
| Fuel consumption | 3 km/l | ABRATI (2015) |
| Bus fare per passenger km<br>(US$ 0.82/3 km/l)/50 passengers | US$ 0.0054 | ABRATI (2015) |