Estimating the Economic Value of Zoning Reform*

Santosh Anagol Fernando Ferreira Jonah Rexer

October 26, 2021

Abstract

We develop a framework to estimate the economic value of a recent zoning reform in the city of São Paulo, which altered maximum permitted construction at the city-block level. Using a spatial regression discontinuity design, we find that developers file for more multi-family construction permits in blocks with higher allowable densities. We incorporate these micro-estimates into an equilibrium model of housing supply and demand to estimate the long term impact of zoning changes. Supply responses from the reform produce a 2.2 percent increase in the total housing stock, leading to a 0.5% reduction in prices on average, with substantial heterogeneity across neighborhoods. Consumer welfare gains due to price reductions are small, but increase 4-fold once we account for changes in the built environment, with more gains accruing to college-educated and higher income households. However, nominal house price losses faced by existing homeowners and landlords overshadow all consumer welfare gains.

Keywords: housing supply, building restrictions, zoning reform, welfare gains.

*Anagol: Wharton School, University of Pennsylvania. Ferreira: Wharton School, University of Pennsylvania and NBER. Rexer: School of Public and International Affairs, Princeton University. We are grateful for support from the Wharton Dean's Research Fund, and the Research Sponsors Program of the Zell/Lurie Real Estate Center. We thank Tom Cui, Anna Gao, Renan Muta, Sophia Winston, Alexandru Zanca, and Holly Zhang for excellent research assistance. We also thank Rohan Ganduri and seminar participants at Imperial College Business School, Wharton Urban Lunch, European Urban Economics Association meeting, American Urban Economics Association meeting, NBER Summer Institute Real Estate, FGV-SP, Brazilian Society of Econometrics, LACEA-LAMES, NBER Public Economics, and the SITE conference at Stanford for helpful comments and suggestions.

1 Introduction

The global urban population grew from 33% in 1960 to 56 percent in 2019, and is predicted to reach 68 percent by 2050.[1] While these predictions are consistent with past trends in urban growth, there is considerable uncertainty and debate about the quantity and quality of housing the world's cities will be able to supply for their residents. Satisfying the predicted increase in demand to live and work in cities will only be possible with a strong supply response from developers who build residential structures. In most cities, however, the supply of buildings is highly regulated by local authorities - the so called "not in my back yard" NIMBYism - who can legally determine if and how developers can construct real estate structures (Glaeser, Gyourko and Saks, 2005; Gyourko and Molloy, 2015; Gyourko, Hartley and Krimmel, 2021).

Zoning restrictions, such as limits on the density of buildings, are generally associated with increases in the cost of living (Glaeser and Gyourko, 2018; Brueckner and Sridhar, 2012; Ding, 2013), greater segregation and reduction in economic convergence (Trounstine, 2018; Ganong and Shoag, 2017), and result in the loss of economic output (Hsieh and Moretti, 2019). The literature has primarily focused on cross-metropolitan area studies because of the variation in zoning rules across those geographies. However, estimation of causal effects of zoning changes may have three major problems. First, zoning policies are often themselves determined by price levels (e.g. high price areas choosing to restrict building to maintain high prices) – so comparisons of places with different zoning policies may conflate price effects with zoning reform effects. Second, zoning policies are multi-dimensional and difficult to compare across cities, leading to less clarity in the relevant composition of the zoning changes across areas. Third, even when these two challenges are solved, it is difficult to trace out the impacts of actual zoning reforms on built environments, household sorting, and welfare given the many aspects of supply and demand of housing impacted by changes in regulation. For example, new housing construction may improve the quality of the housing stock relative to the old and depreciated stock of homes, reduce housing prices, and allow households to move from the suburbs to the city and closer to their workplaces. But densification may also lead to increases in congestion, construction of less desirable apartment units (as opposed to single family housing), and the sorting of households with different levels of income and education with difficult to predict demographic patterns, which is sometimes feared by proponents of NIMBYism.

[1] United Nations Population Division. World Urbanization Prospects: 2018 Revision.

Our paper contributes to the literature by tackling the three challenges above. We study the impact of a major zoning reform on the built environment in a city, and estimate a quantitative model of housing supply and neighborhood demand in order to assess how changes in housing regulations affect the welfare of city residents. We focus on São Paulo, Brazil, which is the 4th largest metropolitan area on earth with a population of 21 million residents. The city of São Paulo implemented a detailed, block-by-block, change in real estate regulation in 2016. The reform centralized the ability to set density parameters previously under control of neighborhoods, and had the general goal of providing more dignified housing for its residents and allowing more densification along some transportation corridors.[2] In practice each city block was assigned a maximum built-area-ratio (BAR) - the ratio of constructed square meters per square meter of land area - which fundamentally defines the density of units that developers could develop on a given land parcel.[3] On average, the maximum BAR in the city's approximate 45,000 blocks increased from 1.54 to 2.09, allowing 36% more construction for a given lot size, and 45% of the city blocks had a maximum BAR increase of 1 or more.[4] This reform gives us unique variation in zoning restrictions over time – so we can observe how developers respond to zoning changes – and variation over space (city blocks) – which allows us to control more precisely for the endogenous determinants of zoning policies such as prices. Moreover, São Paulo has an unusually rich set of administrative data on building permits, the built environment, demographics, commuting patterns and wages to allow for detailed tracking of the zoning reform's impacts and estimation of its welfare consequences.

We begin our analysis by leveraging these block-by-block changes in BAR policy to estimate the supply response of developers. We implement a boundary discontinuity design where each block is categorized as either a treatment block, defined as a block where BAR increased in the 2016 reform, or a control block, a block where BAR stayed constant or decreased after the reform. For treatment blocks, our running variable is defined as the distance to the nearest control block, and

[2] The São Paulo reform's focus on increasing housing density near transportation corridors is similar to a recently failed reform aiming to increase housing density along California's public transit system (Dougherty, 2020). Mumbai's 2034 development plan also originally included, but then removed, plans for allowing high density construction near metro and commuter rail stations (Ashar, 2018).

[3] We use the term BAR as it corresponds directly with term used in Sao Paulo for this concept. It is closely related to the "FAR" ("floor-to-area"), "FSI" ("floor-space index") and "FSR" ("floor-space-ratio") metrics used in other parts of the world. Sao Paulo has a unique set of rules defining what counts as constructed area; it includes livable space but excludes hallways, elevator shafts and in some cases mezzanine lobbies and swimming pools. Given these idiosyncracies we prefer to use the "BAR" terminology.

[4] See Tabarrok and Cowen (2018) for a discussion of the floor space index parameter in India, similar to the BAR parameter, and its importance in determining housing availability in Mumbai.

for control blocks our running variable is defined as the distance to the nearest treatment block. The boundary discontinuity design focuses on the variation in BAR between nearby treatment and control blocks, and balance tests suggest that the typical endogenous determinants of zoning policy, such as prices and density, are indeed balanced at the cut-off.

We find, in our preferred specification, that the 1.4 BAR point increase at the cut-off between nearby treatment and control blocks causes an increase of .003 multi-family dwelling permits filed per block per quarter, which is a sixty six percent increase in permits per unit of BAR relative to nearby control blocks. The differences between treatment and control blocks emerge approximately one year after the zoning reform was passed, which is plausible given the time it likely takes for developers to create project plans, aquire land, etc.. We find no effect of a BAR increase on single-family home permits, which is consistent with the BAR change allowing greater development of larger structures. We also estimate a spatial spillover model (following a strategy similar to Diamond and McQuade (2019)). We find little evidence of treatment effects on control blocks near the boundary discontinuity, suggesting little substitution of projects from control to treatment blocks. We also find similar sized effects for treatment blocks near and far from the boundary, suggesting muted agglomeration effects so far.

Our next step is to set up a micro-level model of supply and demand to estimate the welfare effects of the 2016 zoning reform, and to evaluate other more aggressive counter-factual zoning reforms. In our supply model, developers apply for permits as a function of expected profits, which are empirically proxied with prices, neighborhood features, and regulatory constraints. We estimate this model at the commuting-zone level and use a Poisson model in order to avoid the problem introduced by log-linear models when many blocks have zero permits. We instrument for BAR constraints using the boundary based variation, while at the same time controlling for other regulations, such as rules limiting the size of a new building's footprint (i.e. the "shadow ratio"). In this specification we estimate that a max BAR increase of 1 changes the number of permits issued by 75 percent, similar to our boundary discontinuity model.

On the demand side, individuals maximize utility when choosing city neighborhoods as a function of prices, location features, access to employment opportunities, and commuting, allowing for residents to value both positive and negative aspects of more densification. To estimate the model, we use survey data on 24,800 São Paulo metropolitan statistical area households which includes information on demographics, place of residence and place of work. We combine this information with neighborhood-level listing price data from a multiple listing service (MLS). To es-

timate the elasticity of location choice to price we instrument listing prices by geographic features in a ring outside the commuting zone; this strategy follows Berry, Levinsohn and Pakes (1995) and rests on the assumption that these outside features impact local prices through competition but do not directly impact the utility of individuals living in a given zone. To estimate consumers' preference for being close to high paying jobs, we follow Tsivanidis (2019) and define an index of residential commuter access (RCMA) which measures a location's average travel time to zones with high paying jobs. Finally, we include other important neighborhood features, such as travel time to the household head's current place of work, age of buildings, number of units per building, density of neighborhood, share of households with a paved road, average zone income, and share of zone adults with a college degree. Preferences for all of these features are allowed to vary according to an individual's demographic characteristics, and we estimate preference parameters using the multinomial logit model as in Bayer, Ferreira and McMillan (2007).

With supply and demand parameters in hand, we next estimate counterfactual welfare and distributional consequences from the 2016 reform. The baseline zoning parameters come from the pre-2016 period. Imposing the new 2016 zoning map on our supply equation leads to new construction in different areas throughout the city.[5] We then apply the equilibrium assumption that supply must equal demand. This market clearing condition delivers us a new price vector, which in turn leads to changes in the neighborhood demographic composition and built environment variables via the demand and supply parameters. Average travel time to work by neighborhood may rise due to the congestion costs of increased density. We analyze the new sorting of residents across the city to calculate the impact of the zoning reform, and also estimate the welfare effects, i.e., consumer surplus coming from the expected utility of having more access to neighborhoods with different characteristics.[6]

We find that the 2016 reform extra flow of construction represents only a 2.2% percent net increase of the total housing stock in the city, leading to 0.5% reductions in prices on average, and consequently small gains in welfare. But this average result masks some heterogeneity; neighborhoods with the largest BAR shocks see more construction and lower prices. The largest construction increases and price decreases are 37.5% and 6.2%, respectively. We also find that college-educated and higher income households gained the most from the reform, especially be-

[5] The 2016 zoning reform produces short term changes in permits, and we use longer term data to estimate a function that translates permits to construction changes.

[6] Our preferred model allows for the city population to grow as more housing is built, but we also test a model that assumes a closed economy. Details are shown in Section V.

cause more of those families can now move from the suburbs to the more central parts of the city. As such, demographic change in the city is limited even under more aggressive reform counterfactuals, since those who migrate are positively selected and therefore resemble incumbent residents.

Interestingly, welfare gains from price changes increase more than 4-fold once we account for changes in the built environment, as density increases and the age of housing units falls. This suggests that a fair amount of the economic value of zoning reforms comes through the presence of a newer housing stock and individual preferences for densification. Congestion costs, in contrast, only reduce aggregate welfare gains by 1.3%. Counterfactual simulations of more aggressive zoning reforms - doubling allowed densities in upzoned neighborhoods - produce much larger gains in welfare (2.98% of city GDP). Since the BAR averages implemented by São Paulo in 2016 are a third or less of what is allowed in denser cities, such as Singapore and Hong Kong, local policymakers still have substantial room to implement more aggressive densification strategies that improve resident welfare.

Finally, we compare the consumer surplus with the nominal house price losses suffered by both homeowners and landlords. Those losses are about 16 times larger than consumer gains, which may explain the lack of homeowner support for more dense construction. We also measure other effects of the reform, such as producer surplus, welfare for new incoming residents, and potential changes in productivity due to more agglomeration in Sao Paulo. The producer surplus is the largest among those factors, but even the aggregation of all potential gains still do not match the nominal house price losses faced by existing real estate owners.

Our work is at the intersection of many literatures. In addition to the research on zoning and housing regulation cited above, other papers have focused on the impact of geographic constraints on housing supply (Saiz, 2010; Baum-Snow and Han, 2021). There is also a growing literature on structural models of housing supply. Murphy (2018) estimates a model of housing supply with a focus on the role of construction costs, while Paciorek (2013) studies the relationship between supply constraints and price volatility. Calder-Wang (2021) estimates the welfare of New York City residents given changes in availability of rental units due to the expansion of Airbnb.

Another related literature focuses on understanding the internal structure of cities using detailed micro data on the built environment and sorting. Harari (2020) investigates how city shapes in India affect transit accessibility, land use regulations, and city growth. Ahlfeldt et al. (2015) estimate a model of internal city structure to quantify the effect of densification. A few recent

papers study policy reforms within cities to estimate the value of urban policies and amenities. Allen, Arkolakis and Li (2016) studies optimal city structure, and apply their theoretical results to evaluate Chicago's existing zoning system. Tsivanidis (2019) estimates how new transit lines impact worker sorting and welfare in Bogota, and Balboni et al. (2020) estimates the impact of a new bus rapid transit system in Dar Es Salaam.

This paper also fits in to three other broad literatures. First, there is a large literature on the economics of urban density and agglomeration effects recently summarized by Duranton and Puga (2020). Those authors note that future progress in this literature could encompass the dynamics of building construction and raise empirical standards in the identification of causal effects, both accomplished in our work. Second, there is a growing literature studying how housing supply will respond in the face of larger demand for cities, recently summarized in Brueckner and Lall (2015); our paper provides a first complete evaluation of a zoning reform in a developing country city context. Finally, Epple, Gordon and Sieg (2010) and Combes, Duranton and Gobillon (2021) directly estimate production functions for housing. Our paper contributes to that body of work by estimating in detail how developers respond to zoning reforms at very granular geographies.

2 History of Zoning in São Paulo

São Paulo has had three major zoning reforms in the past fifty years: 1972, 2004 and 2016. All of these reforms created "zone types" such as "Mixed Residential Use," with each zone type assigned a set of building parameters. Each of the city's blocks are assigned to a zone type, and the zone type designation determines the block's building parameters. The primary building parameters set in each reform are:

- Built-area-ratio (BAR): The built-to-area ratio (BAR) is the ratio between the computable area of the building and the lot size.
- Shadow ratio (SR): the ratio between the projected area of the building and the lot size; setting this determines how much of a given lot can be covered by a structure. Although both BAR and SR are density indices, the SR is independent of the building's number of floors, since it only depends on the projected area, while the BAR increases with the insertion of additional built area.
- Building usage: defines whether the building can be used for residential, commercial or

industrial purposes and was set at the time a building permit was issued. Most zone types allow multiple building usages, but there are some that require only certain building usages.

In the main text of this paper we focus on the built-area-ratio (BAR), as variation in this parameter can most greatly affect the built environment. In the Appendix we discuss how the other zoning parameters vary with BAR, and report specifications that control for other zoning parameter changes when we estimate BAR impacts.

São Paulo's first city wide zoning regime was established in 1972 in response to rapid, haphazard, urban growth. Poor areas lacked enforcement of building rules allowing developers to build at their own will (Saconi and Entini, 2013). Wealthy neighborhoods, such as Jardim América, had rigorous regulation implemented by private developers.[7] The 1972 zoning law primarily aimed to preserve the architecture of richer neighborhoods, while guiding the city's growth towards the periphery (Giaquinto et al., 2010); this law also established that no new building could have a BAR above 4. However, the majority of urban land resided in zones with a maximum BAR of one.[8]

São Paulo enacted a Master Plan in 2002 and a new Zoning Law in 2004, primarily in response to new federal laws mandating cities to have urban master plans. The 2004 Zoning Law had two major features relevant for our analysis. First, the determination of specific building parameters, such as BAR, was decentralized to the "subprefeitura" or neighborhood level (São Paulo has 32 subprefeituras). In particular, the city would determine the zone type of each block, but the same zone type could have different building parameters based on the block's subprefeitura.

Second, the city expanded the BAR parameter scheme to include minimum BAR, basic BAR and maximum BAR levels, each of which were chosen by the subprefeitura government for each zone-type within the subprefeitura. A building developed with a BAR between the basic and maximum BAR was required to pay the so called "onerous grant" fee, charged per square meter built above the basic BAR level. Buildings developed between the minimum and basic BAR did not have to pay any extra fees. The 2004 zoning reform also legislated, at the city level, that each district – subsets of subprefeituras - would have a limited stock of square meters above the basic BAR that could be constructed.[9] Once a given district exhausted its available capacity above the

[7] For example, the Jardim America neighborhood was administered by a British developer called Companhia City.

[8] There were exceptions to maximum BAR rules. Hospitals, hotels and schools, for instance, had their maximum BAR levels determined by separate legislation. Similarly, regions under Joint Urban Operations – partnerships between the local government and the private sector to improve certain areas – had their BAR decided case by case. Another mechanism created to exceed the BAR cap was the Adiron Formula, which allowed developers to build above the zone legislation BAR - but no more than 4 - if the building decreased its shadow ratio.

[9] These limited stock levels were set separately for residential and non-residential structures.

basic BAR, no further construction could occur above the basic BAR level.[10] [11]

Overall, the 2004 reform kept the maximum allowable BAR similar to what was established in the 1972 reform. The main change was to require developers to pay extra fees to build above the newly established basic BAR levels.

The 2016 zoning reform was initiated by Workers Party mayor Fernando Haddad in 2014 with the creation of a new city Master Plan. The stated goals of the reform were to provide dignified housing, guide urban growth, improve urban mobility, improve life in the neighborhoods, promote economic development, incorporate an environmental agenda, and preserve cultural heritage. A key feature of this reform was to standardize building parameters across the whole city by assigning a fixed set of building parameters to each zone type. Under the new reform blocks were assigned to a zone type, and then the city-wide building parameters associated with that zone type would be applied consistently across the city. This reform removed the power of local subprefeituira governments to set BAR and other building parameters within their jurisdictions. In addition to that, it simplified and centralized the BAR regulation, by implementing a Basic BAR of 1.0 in the entire city. The 2016 reform also eliminated the district-level maximum amounts that could be built above the basic BAR level.

A main idea in the 2016 law was to group zone types in to three major groups corresponding to a particular development strategy. Every zone type was labeled as one of the following categories: transformation, qualification, or preservation. The goal in transformation zones was to promote higher urban density, in terms of both residential and non-residential structures, near the city's main transportation corridors. The aim was to reduce the city's traffic and bring people closer to their jobs by improving land use in areas closer to medium and high public transportation networks, such as train, subway, monorail, and bus corridors. The objective in qualification zones was to improve life in residential neighborhoods by favoring moderate urban density; the standard maximum BAR for these zones was set at 2.0 and they implemented a building height restriction of 28m. Zones were designated "preservation" status with the purpose of preserving the environment and cultural heritage of the city. Figure A1 shows a color map of the 2016 reform, with the transformation zones in maroon, qualification zones in gray and yellow, and preservation

[10] Resources from the onerous grant, however, only appeared in 2005, since developers got permits before the implementation of the 2002 Master Plan, which granted them a three-year exception period for the reduction in BAR (Sandroni, 2010).

[11] In addition to decentralizing the determination of building parameters, the 2004 Zoning Law also created the Social Interest Special zone type (ZEIS) which aimed to create new low-income housing and regularize existing informal housing.

zones in green. Figure A2 displays the striking detail of the block-by-block land use regulation in the middle class neighborhood of Jabaquara.

The 2014 Master Plan and associated 2016 zoning reform are valid for 16 years with a revision scheduled in 2021. Given that much of the predicted worldwide urban growth will occur in developing countries, it is interesting to test whether a zoning reform of this type, in a developing country city, actually leads to changes in the built environment and improvements in resident welfare.[12]

3 Data

3.1 Zoning data

Zoning data comes from the Cidade de São Paulo Desenvolvimento Urbano. We geo-reference and digitize maps of zoning boundaries at the block level for 2004 and 2016; there are 22 zone types in 2004, and 38 in 2016. In total, we are able to match 45,082 of São Paulo's city blocks to a zone-type, or 96% of the city's 46,987 blocks. We then match these zone-types to the relevant minimum and maximum allowable density parameters at the neighborhood and zone level. From this, we calculate the maximum allowable BAR. In some cases, these parameters vary within a zone-type depending on the size of the lot. In these cases, we define the maximum allowable density parameter as the maximum of all possible values.[13] For now we focus on the BAR parameter; see the Appendix for discussion of variation and analyses using the other zoning parameters.

We obtain the maximum BAR values in both periods for 43,250 city blocks.[14] From this data, we define treatment blocks, control blocks, and the running variable for our boundary discontinuity design. A treatment block is defined as a block whose maximum allowable BAR increased from the 2004 zoning regime to the 2016 zoning regime. A block is designated as control if its BAR declined or stayed the same between the 2004 and 2016 reform. More than 50% of all blocks experienced an increase in maximum BAR.

[12] The reform may be toothless if, for example, developers can evade zoning rules via paying bribes – in such an environment zoning rule changes would have little impact as even existing rules are not enforced. Also, it is possible that zoning rule changes translate in to only minor built environment changes because of other market frictions, such as problems in land acquisition (Bryan et al., 2017) or developer credit constraints.

[13] In defining the maximum BAR before 2016, we account for the fact that some blocks are in districts where the allowed capacity for building above the basic BAR level has been exhausted. In these blocks, the "effective" BAR prior to the 2016 reform is the basic BAR level.

[14] Note that this is less than the 45,082 city blocks for which zoning information is available. These missing blocks are primarily parks, municipal areas, and bodies of water. The remainder are cases in which zoning information was available, but BAR parameters were missing or not relevant for the particular category.

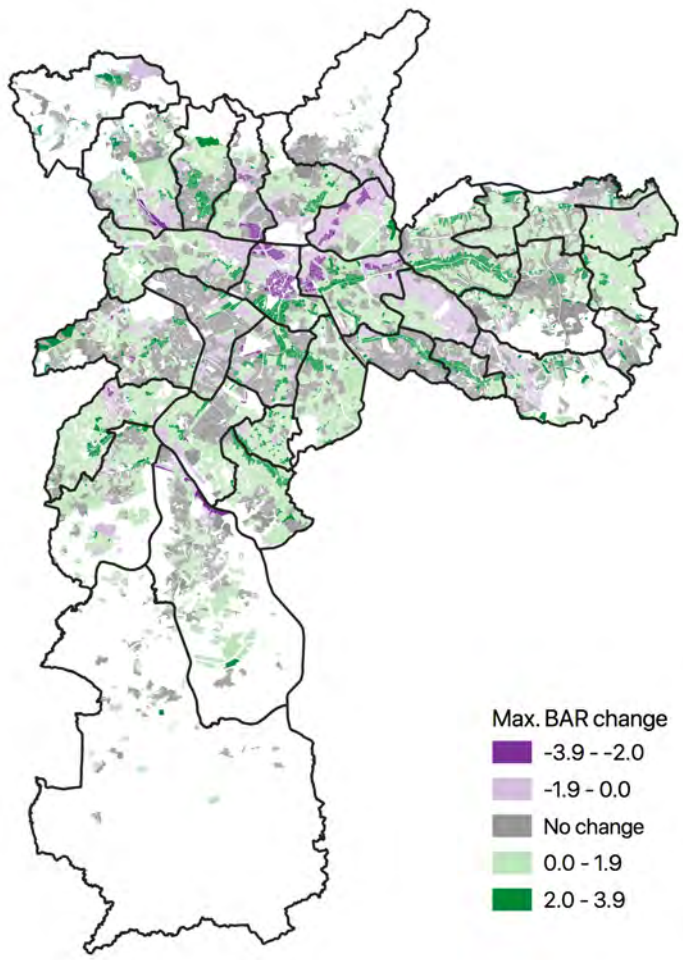


Figure 1: Map of Change in Maximum Allowed Built Area Ratio (BAR) from 2004 to 2016 Zoning Regime

The underlying BAR variation that determines the treatment/control status for each block is mapped in Figure 1, along with the borders of the city's subprefeituras. There is a somewhat general pattern of blocks in the outskirts of the city experiencing positive changes in maximum allowable BAR, and blocks within the central regimes experiencing negative changes or no change in maximum allowable BAR. However, there are many blocks in the central area that received positive (green) changes in their maximum allowable BAR - typically along major transportation corridors. Figure A5 presents binned scatter plots on the relationship between block level characteristics (x-axis) and the change in maximum allowed BAR from the 2004 to 2016 zoning regime. On average, blocks with lower levels of BAR in the 2004 regime received greater increases in allowable BAR. We also see that blocks with higher residential shares of constructed area received greater BAR increases. Blocks with greater density (as measured by the log constructed area per square kilometer) received on average lower changes in BAR, and blocks with greater average land values also received lower changes in BAR.

3.2 Permitting data

Data on building permits were scraped from the Urbanism and Licensing Center of the City of São Paulo, or SMUL.[15] Building Permit data is useful in tracking where and at what volume development is taking place in the city of São Paulo. Since the year 2000 SMUL has been publishing all the filed permits that go through their office, this includes permits for new buildings, demolition, installation of security systems, and certificates of conclusion, among many other types of building related permits. There are over 50 different types of permits issued through SMUL, the majority of which are related to the construction of new buildings. The data is organized by quarter and by year of its filing and includes information about which region of the city the permit is in, the address, the zoning and land use classification, as well as the engineers, architects, and owners leading the project.

Residential building permit filing data from 1997-2020 forms the basis for our key outcome variables. [16] We obtain data on approximately 30 thousand total residential permits, of which approximately 85% are multifamily buildings and 15% are single family dwellings. We then aggregate the number of permits and the number of units at the quarter-block level, yielding a panel

[15] https://www3.prefeitura.sp.gov.br/deolhonaobra/Forms/frmConsultaSlc.aspx. SMUL also handles issues related to zoning and land use for the city of São Paulo and is part of the development of the comprehensive plans that have influenced zoning changes in 2004 and 2016.

[16] Projects that were approved before 2000 are not listed in this data set although if a request for a permit was issued prior to the year 2000 and approved afterwards, it does appear in the data set.

of 3,195,116 quarter-block observations. The main outcome variables are the quarterly count of total new building permits for single and multifamily buildings, plotted in Figure A3. Permitting activity was much larger in the mid-2000s, and declined since 2014 due to the national economic and political crisis. Interestingly, multifamily permits filings more than doubled since the approval of the 2016 reform.

3.3 IPTU data

Information on the stock of buildings in São Paulo, including constructed area, lot sizes, assessed construction value, number of units per building, and assessed land value comes from the IPTU property tax data, which we obtain annually from 1995 to 2019. In 2016, the year of the zoning reform, this data covers 3,316,608 individual tax paying units in 1,582,532 unique buildings. The construction and land values are assessed values produced by the São Paulo property tax assessor office, and form the basis for annual property tax payments; these values are only indirectly based on market transactions. We collapse this data to the block-level to obtain block-level average lot area and constructed area in m^2, as well as mean land value and construction value per m^2. We also obtain the share of lot and constructed area with residential vs. commercial designation by the property tax authority, as well as the total number of units and buildings in each block. This information is available in the pre-2016 period for a total of 43,990 São Paulo city blocks, or 94%.

3.4 Listing Price Data

Given that the IPTU value data is based on assessments as opposed to market values, we also collect apartment listing price data from the online marketplace Properati. The data has approximately 200,000 buildings listed for sale and for rent in 2016. Each listing in the system contains the price, the type of transaction (rent or sale), the location, the type of unit (i.e. apartment, house, office, etc.), and a general description of usage (residential, commercial, bare-land or non-specified). The IPH (Hiperdados-Properati Index), which used data from Properati, was the most complete real estate pricing indicator in Brazil until it was sold and renamed to Casafy (who no longer published the data).

3.5 RAIS Data

We measure economic activity at the block level using the total formal sector wages paid to workers whose firm address is within a block. We obtain this variable from the RAIS data, which is individual level monthly wages data for all formal sector workers in São Paulo. We aggregate the individual monthly wage data up by year, and then further aggregate at the block level to obtain total annual wages paid per block. This data is available from 2000-2019.

3.6 Commuting Zone Survey Data

We use commuting survey data from the "Pesquisa Origem e Destino 2017" survey.[17] The commuting survey covered São Paulo metropolitan statistical area households and includes information on demographics, place of residence and place of work. We take as our sample the 24,800 households for which the household head is working. The survey was stratified by commuting zones, and for this employed sub-sample we obtain 492 commuting zones within the São Paulo metropolitan area (i.e. including both the São Paulo municipality that we study as well as the surrounding suburbs). 329 of these zones are within the municipality itself. From this data, we take individual home and work locations, which we use to calculate commuting distances, household head education and age, household size and total monthly income, and dwelling ownership status. We also use this data to estimate several commuting zone-level characteristics, including average income, education, and the share of paved roads.

4 Border Discontinuity Design and Results

Our primary identification strategy to estimate the impact of zoning reform is to compare the evolution of permitting activity in geographically close blocks with different treatment status. To construct regression discontinuity plots we begin by defining as "treated" all blocks in zones that experienced an increase in maximum allowable BAR as a result of the 2016 reform. Control blocks are those that fall in zones which experienced either no change or a reduction in maximum BAR. Then, for each treatment block, we calculate the distance in kilometers to the nearest control block. Distance is calculated from the block's centroid to the edge of the nearest block in the other group. This distance defines the running variable in our border discontinuity design for treatment blocks. The running variable for control blocks is the distance to the nearest treatment block; we define

[17] The survey data is available at `http://www.metro.sp.gov.br/pesquisa-od/`.

this distance as negative. Our boundary discontinuity design focuses on outcome comparisons between control blocks with small absolute values of the running variable (i.e. control blocks near to treatment blocks) and treatment blocks with small values of the running variable (i.e. treatment blocks near control blocks). [18] In the Appendix we report the number of blocks in each of these 50 meter bins, and find no discrete jump around the cut-off (Figure A4).

Our key identification assumption is that, in the absence of the 2016 zoning reform, outcomes would have evolved similarly across the zoning borders at which the change in maximum BAR switches from positive to non-positive. To corroborate this assumption Figures A6 and A7 assess the covariate balance of existing structures and economic activity across treatment and control blocks in the year prior to the zoning reform. The x-axis groups blocks within 50 meter bins away from the cut-off, and the y-axis plots the binned-average change in covariates. We find relatively small jumps at the zoning cut-off for total land and constructed area, and no difference in average assessed value of land and constructed are. The share of buildings categorized as residential and commercial also does not appear to change discretely at the zoning border.

The core treatment variation we wish to focus on is the change in maximum allowable BAR in nearby treatment and control blocks. Figure 2 shows how the change in BAR from the 2004 to 2016 zoning regime varies as we move towards the 2016 zoning borders from control blocks to treatment blocks. City blocks just to the left of the cut-off experienced an approximate .1 decrease in their maximum allowable BAR, while blocks just to the right of the cut-off experienced an approximate 1.3 increase in their maximum allowable BAR. [19]

To formally estimate the treatment effect of the higher BAR levels we estimate the following regression discontinuity model:

$$y_{ij} = \beta 1\{x_{ij} > 0\} + f(x_{ij}) + \delta_j + \epsilon_{ij} \tag{1}$$

[18] We investigated a "bunching" strategy where we estimated how the fraction of new buildings built at exactly the maximum allowable BAR changes with the reform; in principle such a strategy could be used to estimate the marginal increase in constructed area resulting from an increase maximum BAR. The primary challenge is that neither the permit nor IPTU data report the 'computable' built area measure required to exactly calculate a building's BAR. Instead these data just report constructed area. Given the bunching strategy relies on exact measures of BAR, we chose not pursue this approach. We do have a strong sense that maximum BAR limits were binding throughout the city, from the fact that almost all of the allowed building up to maximum BAR levels was utilized prior to the reform.

[19] Figure A8 reports the average maximum BAR values in our treatment and control blocks before and after the 2016 reform, not just the change in max BAR shown in Figure 2. Treatment blocks had lower maximum BAR values prior to the reform relative to control blocks, and have higher maximum BAR values after the reform. The pre-existing differences prior to the reform are to some extent mechanical, in that treatment blocks are defined as those which experienced an increase in BAR in the 2016 reform. But the fact that max BAR did differ prior to the reform, even in a narrow bandwidth around the border, strongly suggests that we should focus on how outcomes change before and after the reform, as opposed to just analyzing a cross-sectional border discontinuity design after the reform.

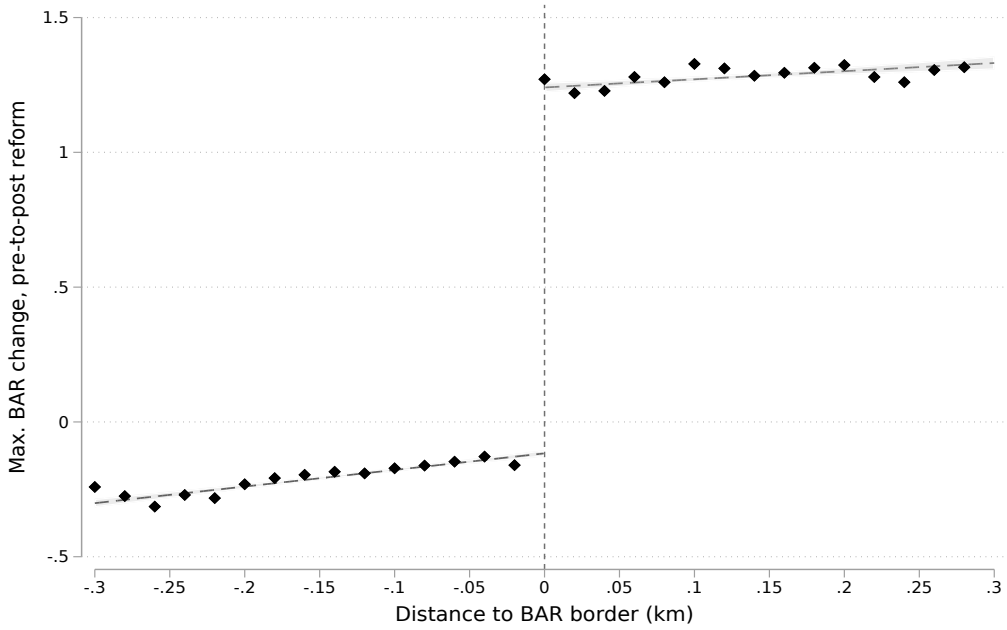


Figure 2: Built Area Ratio Change, Pre-to-Post 2016 Reform

where y_{ij} is an outcome in block i which is located in subprefeitura j, x_{ij} is the value of the running variable for block i in subprefeitura j, δ_j is a subprefeitura fixed effect, and ϵ_{ij} is an error term.[20] The indicator function $1\{x_{ij} > 0\}$, "Treat BAR", is our main independent variable of interest, and β is our estimate of the treatment effect. The function $f(\cdot)$ is a polynomial fully interacted with $1\{x_{ij} > 0\}$.

Table 1 reports changes in max BAR RD point estimates and standard errors using four versions of this specification: Column 1 compares all treatment versus control blocks (i.e. only includes the indicator $1\{x_{ij} > 0\}$ in the model), column 2 adds a linear control of the running variable (and interacts it with a treatment indicator), columns 3 adds quadratic controls, and column 4 add cubic controls. RD estimates show a remarkably stable 1.4 point estimate. In Panel B of table 1 we add subprefeitura fixed effects, so that all variation comes from changes within subprefeitura. Estimates remain practically unchanged, suggesting that the reform treatment is not driven solely by pre-reform differences in how different neighborhoods controlled zoning parameters.

[20] Figure 1 shows São Paulo's subprefeituras outlined in black.

Table 1: RD First stage

Outcome	Maximum BAR change			
	(1)	(2)	(3)	(4)
Panel A: No sub-prefeitura FE				
Treat BAR	1.519***	1.426***	1.365***	1.354***
	(0.00515)	(0.00678)	(0.00775)	(0.00950)
Specification	Base	Linear	Quadratic	Cubic
Observations	43231	43231	43231	43231
Mean of Dep. Variable	0.597	0.597	0.597	0.597
Panel B: With sub-prefeitura FE				
Treat BAR	1.516***	1.453***	1.386***	1.355***
	(0.00513)	(0.00654)	(0.00768)	(0.00968)
Specification	Base	Linear	Quadratic	Cubic
Observations	43225	43225	43225	43225
Mean of Dep. Variable	0.597	0.597	0.597	0.597

Robust standard errors in parentheses. Specification refers to the order of the polynomial for the running variable, which is distance to the RD boundary. The polynomial is always interacted with the treatment indicator. Sample is all city blocks with zoning information. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.1 Reform Effect on Building Permits

We now use the same set of treatment and control blocks to estimate the causal effect of the greater allowable BAR on building permits. Figure 3 splits our permit outcome variable into multifamily permits (top two figures) and single family permits (bottom two) figures. The panels show averages of the outcome variable by bins of .02 km distance to the closest border, i.e., the closest block from opposite treatment status. The outcomes represent average quarterly building permits in a block, and the left panels show pre-reform data since 2012. Focusing on multi-family permits first, the pre-reform panel shows a small difference in permits at the discontinuity. In contrast, the post-reform period show that treatment blocks just on the higher BAR side of the zoning border experienced approximately 0.003 more permits issued relative to the control side of the border (0.007 versus 0.004). The higher BAR allowance causes a doubling of average multi-family permits filed per quarter relative to control blocks.

Interestingly, the zoning treatment effect on permits is concentrated in multi-family units. The bottom two panels of figure 3 show that not only are total number of single-family permits smaller in both pre- and post-reform periods, but also that there are no differences around the spatial

discontinuity. Overall, the causal effects of the reform are concentrated in multifamily permits, consistent with taller buildings being more sensitive to BAR constraints.

Both pre- and post-reform figures show a downward slope in average quarterly permitting activity as we move from deeper in the control area towards the cut-off, and then a slightly positive slope from the cut-off towards deeper in to the treatment area. We find a similar pattern in the total number of buildings per block in the 2015 IPTU data (see figure A6).

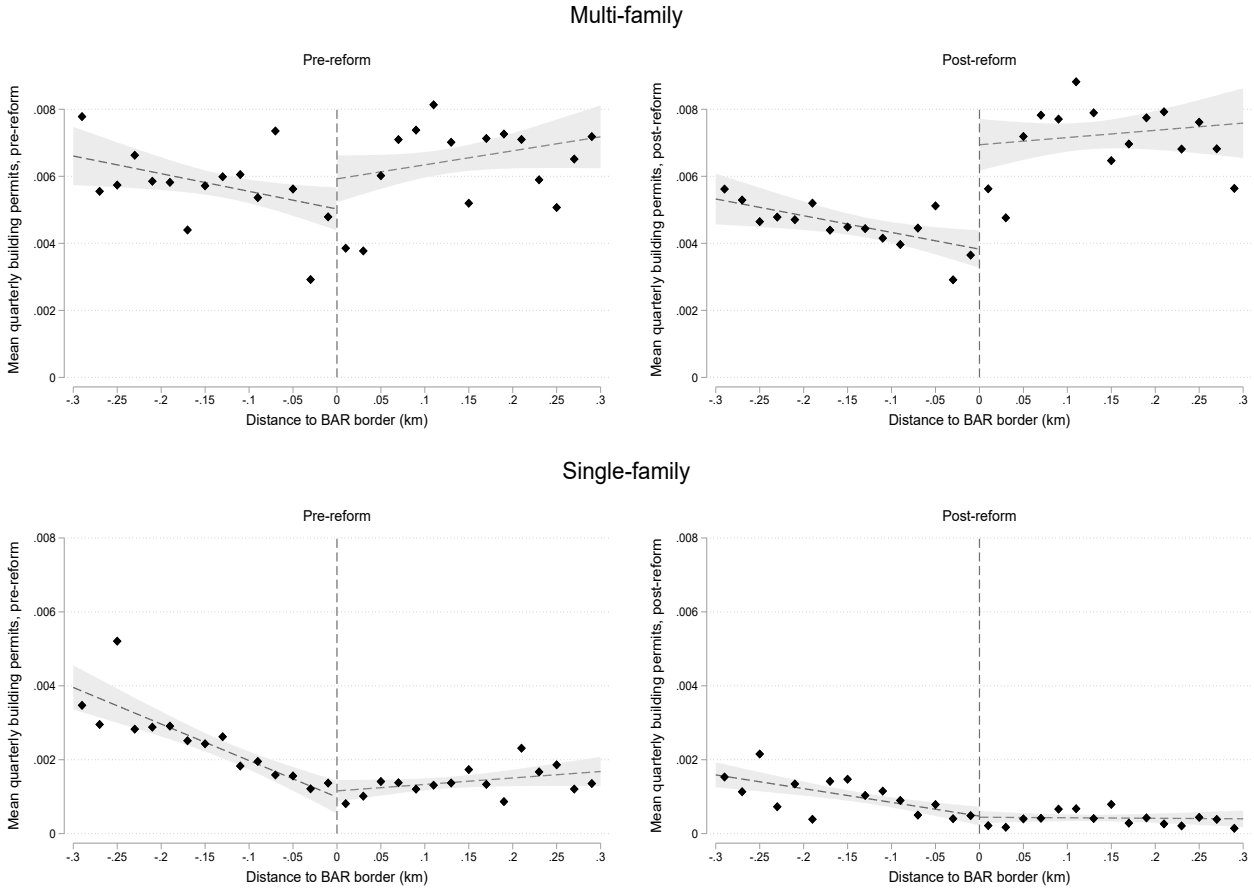


Figure 3: Multi-Family vs. Single-Family Permit Fillings

Table 2 reports RD estimates for multi-family permits, in analogous form to the presentation in Table 1. Panel A reports RD estimates ranging from 0.0036 to 0.0043 (excluding the Column (1) estimate which does not control for the running variable). The inclusion of subprefeitura fixed effects in Panel B has the impact of reducing the magnitude of causal estimates to a range of 0.0022 to 0.0031. This means that an important part of the permitting effects are explained by differences in subprefeitura boundaries. Combining these estimates with the maximum BAR treatment magnitudes estimated in table 1 reveals that increasing maximum BAR by 1 leads to an increase in

multi-family permits between 32% - 67%, relative to nearby control blocks. Appendix Table A4 applies the same RD strategy in a Poisson estimation, finding that increasing maximum BAR by 1 lead to an increase in multi-family permits between 40% - 70%. We also re-estimate this supply model in the next section, after incorporating prices and other local features, and find estimates similar to the upper end of this range.

Table 2: RD reduced form

Outcome	New multi-family building permits			
	(1)	(2)	(3)	(4)
Panel A: No sub-prefeitura FE				
Treat BAR	0.00145***	0.00361***	0.00427***	0.00405***
	(0.000213)	(0.000289)	(0.000342)	(0.000423)
Specification	Base	Linear	Quadratic	Cubic
Observations	43231	43231	43231	43231
Mean of Dep. Variable	0.00470	0.00470	0.00470	0.00470
Panel B: With sub-prefeitura FE				
Treat BAR	0.00166***	0.00216***	0.00257***	0.00309***
	(0.000232)	(0.000284)	(0.000337)	(0.000413)
Specification	Base	Linear	Quadratic	Cubic
Observations	43225	43225	43225	43225
Mean of Dep. Variable	0.00470	0.00470	0.00470	0.00470

Robust standard errors in parentheses. Specification refers to the order of the polynomial for the running variable, which is distance to the RD boundary. The polynomial is always interacted with the treatment indicator. Sample is all city blocks with zoning information. Mean of dependent variable calculated for control blocks within 0.1 km of the BAR boundary. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A9 presents regression discontinuity estimates on the treatment effect of allowing greater BAR ratios separately for blocks with below and above median land values. The binned averages in these plots are produced by first splitting the sample in to below and above median groups, and then using the distance to the nearest zoning border as the running variable.[21] Nonetheless, the figure suggests that the zoning treatment effects are largest in areas with higher pre-existing land values. The zoning reform appears to have spurred greater construction activity where developers expect greater demand and higher profits.

Figure 4 tests if the multi-family permits point estimates, pre and post reform, change accord-

[21] Note that this procedure does not necessarily include treatment and control units on each side of every border (as it is possible that treatment units on one side of a border could have above median land values but the corresponding control units had below median values).

ing to how much of the sample we use away from the cut-off (i.e. considering larger bandwidths). The x-axis shows bandwidths in kilometers around the cut-off (i.e. a bandwidth of .1 includes control and treatment blocks within .5 km of the cut-off). The y-axis shows the regression discontinuity estimate. The top panels show post reform effects with and without subprefeitura fixed effects. Estimates without subprefeitura fixed effects become larger as the bandwidth gets bigger. Post-reform estimates with fixed effects are more consistently around .002 and .003, independent of the bandwidth size. The bottom panels show similar specifications for the pre-reform period of 2012 to February 2016. While estimates without fixed effects show statistically significant effects for bandwidths above .15, those estimates become smaller in magnitude and not statistically different from zero upon inclusion of subprefeitura fixed effects.

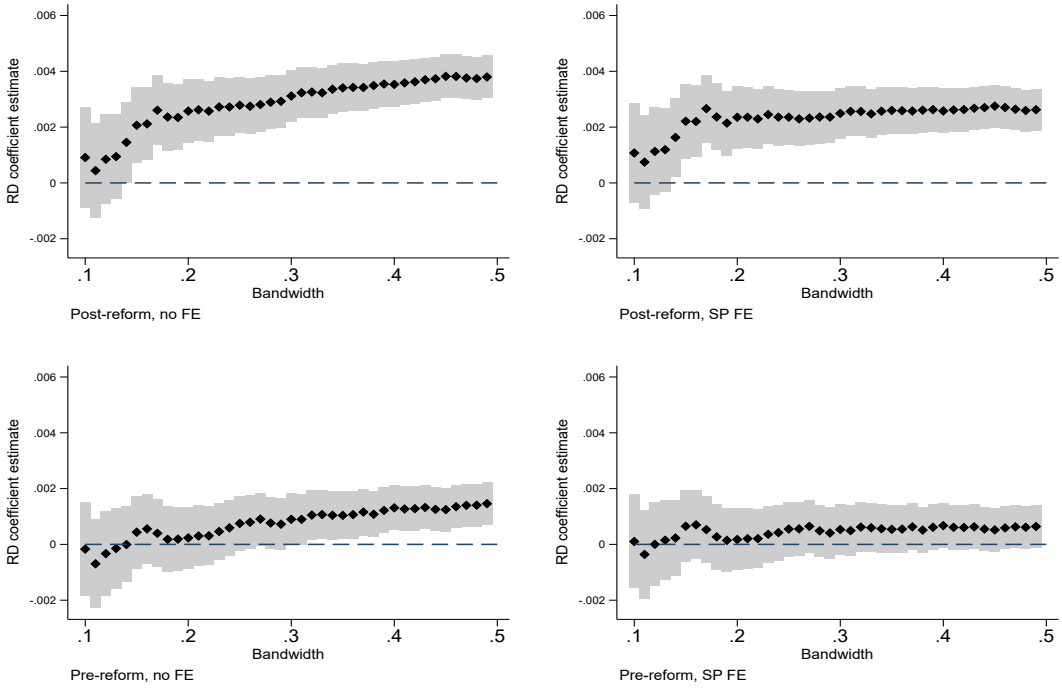


Figure 4: Multi-Family Permit Fillings by Bandwidth of Running Variable

To get a sense of the time-series of treatment effects, Figure A10 presents separate regression discontinuity estimates for each quarter, including both pre- and post-reform quarters. The RD estimate is generally not statistically significant prior to the reform. We see the treatment blocks experiencing greater permitting activity approximately four quarters after the reform, and the point estimates more than double two to three years after the reform.

4.2 Other Zoning Parameters

Our analysis has focused on changes in BAR in nearby blocks; given that the reform included both changes in BAR and zone type designations by block, it is useful to characterize how nearby treatment and control blocks differ in zone type designations as well. Figure A11 presents RD coefficients on the probability that a block falls in to a 2004 zone type (left-panel) and 2016 zone type (right-panel). Both panels only include the top 10 most common zone types under each regime. Focusing on the left-panel, zone types with the largest coefficients are those that were more "targeted" to receive BAR increases in the 2016 reform; the results suggest that mixed used medium density, mixed use low density and environmental blocks were all positively selected to have BAR increases. The effects are relatively small for zone types that were negatively selected for BAR increases, but we note that residential low density blocks were actually less likely to receive BAR increases (at least at the cut-off). Turning to the right panel, we see that treatment blocks are also more likely to be commercial/residential A and zones of social interest, and less likely to be commercial/residential high density B zones.[22]

Tables A1 and A2 present our first stage and reduced form results including fixed effects for the 2004 zoning type of the block (Panels A and B) and 2016 zoning type of the block (Panels C and D). The purpose is to determine to what extent the BAR effects are picking up other permissions that, explicitly or implicitly, go along with certain zoning types, in addition to only using the BAR comparisons that occur within zoning types. Overall, controlling for the 2004 zoning type of the block leads to very small changes for the first stage and reduced form results (compared to Tables 1 and 2). Controlling for the 2016 zoning type leads to a generally smaller first stage effect on BAR (.982 with the subprefeitura fixed effects and the cubic polynomial in Panel D, Table A1 versus 1.355 in Panel B, Column 4, Table 1). Dividing the reduced form effects by this smaller first stage, we see that controlling for the 2016 zone fixed effects leads to a treatment effect size of .00144 extra permits per quarter per BAR point, versus the .0023 estimate without the 2016 zone controls.[23] This suggests approximately 1/3 of the BAR effect could be associated with other zoning permissions implicitly associated with zone types, although it is also possible that these differences are driven by different local average treatment effects induced by the inclusion of the 2016 zone type controls.

[22] The "A" and "B" designations are used to indicate qualitatively more dense areas in the government's description of zone types.

[23] .00144 = .00142/.982 from Tables A1 and Table A2 versus .0023 = .00309/1.355 from Table 2 and Table 1.

4.3 Approvals and Construction

So far we have analyzed the reaction of developers to zoning reform based on the fastest outcome we can observe, i.e., the filing of zoning permits. It turns out that more than 75 percent of permits are approved within 3 to 4 years of their filings, a number that only slightly increased after the 2016 reform. In Table A3 we estimate a similar RD for approved permits, and find estimates in the range of 47-88 percent increase in approved permits for each 1 unit of max BAR increase.

Given the lags involved in filling and approval of permits, and the additional years necessary to observe any type of multifamily projects - which generally have extra construction delays - using constructed buildings as an outcome would give a biased picture of the zoning reform's full effects, at least in the short term. However, we can use data prior to the reform to estimate the conversion rate of permits to constructed buildings, in order to help quantify how our results on permits are likely to convert to constructed buildings. Figure A13 estimates a block-level event-study model on the impact of a permit being issued on the density of new construction, measured as new constructed area divided by total land area in the IPTU data, in the period prior to the reform (2004-2016). The figure suggests that a new permit issued in a block is correlated with increases in density up to 15 years after the issuance. These calculations are only useful to the extent that past relationships between permits and finished construction hold, which may or may not be a good assumption, especially given the recent Covid-19 pandemic. Nonetheless, we will use this relationship to translate permit effects in to construction effects when we estimate our structural model to evaluate welfare.

A second consequence of the limited time that has passed since the 2016 reform is that we do not yet have price for the upcoming construction projects and demographic information that reveals the types of households that will move in to newly constructed housing in response to the reform, primarily because most of the new housing has not been built yet. However, we will readily address this challenge when estimating counterfactual simulations based on the equilibrium model developed in the next section.

4.4 Spillover Effects

We identify the treatment effect of relaxing zoning rules by comparing areas that received a relaxation of BAR requirements to nearby areas that did not receive a relaxation. In this section we analyze the extent to which increases in allowable BAR levels might affect nearby blocks (i.e.

spillover effects). On the one hand, developers may act independently, and spillover effects might be small during the post reform period when developers are simply filing for permits. But there are at least three reasons to estimate spillover effects in our context. First, projects could move from nearby control blocks to treatment blocks, leading us to over-estimate the effect of the BAR reform. Second, buildings in nearby control blocks could *increase* if the greater expected density in the nearby treatment blocks make nearby control blocks more attractive areas to develop as well. Finally, part of the treatment effect in treatment blocks could reflect agglomeration benefits of other nearby treatment blocks - although these benefits would likely appear with large lags.

We follow the Diamond and McQuade (2019) methodology to assess the importance of spillovers in the context of a highly localized zoning reform. The basic strategy is to treat both treatment and control blocks near a zoning boundary as "treated," in the sense that they are nearby to an area where a major zoning change occurred. We compare outcomes for these blocks near boundaries to a "pure control" set of blocks that are farther (i.e., greater than half kilometer) away from the boundary, before and after the reform. To operationalise this spatial difference-in-differences strategy we estimate the following regression model:

$$p_{it} = \sum_{j=1}^{12} I(t > 2016Q2) * I(dc_{j*-.04} = 1) + \sum_{j=1}^{13} I(t > 2016Q2) * I(dt_{j*.04} = 1) + b_i + q_t + \epsilon_{ij} \quad (2)$$

where p_{it} is the number of permits issued in block i in a quarter t, $I(t > 2016Q2)$ is an indicator for post reform, $I(dt_{j.04} = 1)$ is an indicator for a treatment block that is $j * .04$ km away from the nearest control block, and $I(dc_{j*-.04} = 1)$ is an indicator for a control block that is $j * .04$ km away from the nearest treatment block. The equation also includes block and quarter fixed effects. The omitted category is control blocks more than .48 km away from the boundary, and all treatment blocks with distance greater than .5 km are bunched in the 0.52 bin. For visual clarity the graph only shows bins from -.3 to .3 km away from the cut-off.

Figure A12 presents the estimated coefficients. On average permits are higher in all treatment blocks (0.00085) relative to all control blocks (0.0001). The difference between treatment and control groups is slightly larger within .12 km to the boundary (0.001 for treatment blocks and -0.00002 for control blocks). But they are not statistically different from each other, in part because the distance bins are quite small. Moreover, we do not observe control blocks near treatment blocks appearing to have particularly low permit averages. Relative to control blocks 0.5 km away from

treatment blocks, the distance of a given treatment (control) block away from a nearby control (treatment) does not appear to be strongly associated with permitting activity.

5 Welfare Evaluation

5.1 Model of Residential Demand

5.1.1 Choice model

Our model of residential housing demand follows a standard discrete choice framework (Berry, Levinsohn and Pakes, 1995; Bayer, Ferreira and McMillan, 2007). Household i chooses between $j = 1, ..., J$ alternatives, where the alternatives are one of 329 commuting zones within the city of Sao Paulo based on our residential commuting zone survey data. The outside option is living outside the city - in the suburbs - taken by roughly 40% of the individuals in the residential commuting data. Individual utility from choosing to live in zone j will be:

$$u_{ij} = \alpha_i^d p_j + \beta_i^d X_j^d + \gamma_i \tau_{ij} + \xi_j + \epsilon_{ij} \tag{3}$$

Where p_j is the price of housing in location j, measured as the average Properati listing price per square meter in commuting zone j. X_j^d is a K-dimensional vector of housing amenities and demographics of location j. This includes an index of residential commuter market access (RCMA) that measures the extent to which j is located near high-paying jobs, using travel times from a zone commuting matrix.[24] X_j^d also includes the average age, in years, of housing units, average number of units per building, the overall zone-level constructed area density – defined as the sum of all constructed area divided by the zone geographic area – the share of households with a paved road, the average zone income, and the share of zone adults with a college degree. The term τ_{ij} measures commuting costs as the predicted travel time in minutes between the zone in which i works, taken as given, and zone j.[25] Finally, ξ_j is an unobserved location-specific "structural"

[24] We define the RCMA for zone j as:

$$RCMA_j = \sum_{i \in I_j} \frac{w_i}{d_{i,j}}$$

where I_j is the set of zones that have at least one worker living in j and w_i is the average wage in i. Iceberg commute costs are modeled as $d_{i,j} = \exp(\kappa \tau_{i,j})$, where $\tau_{i,j}$ is the average reported travel time in minutes between i and j. We set $\kappa = 0.01$, following micro-estimates from Ahlfeldt et al. (2015) and Tsivanidis (2019).

[25] We model the travel time between living zone k and working zone j with a regression of the form

$$\tau_{jk} = \mu + \eta \log(pop_k) + f(d_{jk}, \phi) + v_{jk}$$

utility shock, which may possibly be correlated with price, although we assume the variables in X_j^d are exogenous. The shocks ϵ_{ij} are distributed i.i.d. type I extreme value. The utility of the outside option is normalized to zero, $u_{i0} = 0$.

Let Z_i be a D-dimensional vector of household characteristics, including household size, age of the household head, a rental indicator, household income, and a college indicator for the household head. The heterogeneous demand parameters α_i^d and β_i^d take the form:

$$\begin{pmatrix} \alpha_i^d \\ \beta_i^d \\ \gamma_i \end{pmatrix} = \begin{pmatrix} \alpha^d \\ \beta^d \\ \gamma \end{pmatrix} + \Pi Z_i \tag{4}$$

where Z_i is the $D \times 1$ vector of demographic variables $z_{i,1}, ..., z_{i,D}$ and Π is a $(K+2) \times D$ matrix of coefficients containing: $i)$ $\pi_{\alpha,1}...\pi_{\alpha,D}$, the interactions of price and each of the D demographics in Z_i $ii)$ $\pi_{\gamma,1}...\pi_{\gamma,D}$, the interactions of travel time and each of the D demographics in Z_i, and $iii)$ $\pi_{\beta^k,1}...\pi_{\beta^k,D}$ for $1, ..., K$, the interactions of all of the demographic variables with all of the K commuting zone characteristics.

First, re-write utility separating out the parameters that only vary by neighborhood and the parameters that contain heterogeneity in preferences as follows:

$$u_{ij} = \delta_j + \mu_{ij} + \epsilon_{ij} \tag{5}$$

where $\delta_j = \alpha^d p_j + \beta^d X_j^d + \xi_j$ and $\mu_{ij} = (p_j, X_j^d, \tau_{ij})\Pi Z_i + \gamma\tau_{ij}$.

Since we observe individual demographics, we can calculate the conditional choice probability that each individual i chooses option j. Define y_i as the choice indicator and θ^d as the vector of demand parameters. We have:

$$Pr(y_i = j | Z_i, X, p, \xi, \theta^d) = \frac{\exp\left(\delta_j + (p_j, X_j^d, \tau_{ij})\Pi Z_i + \gamma\tau_{ij}\right)}{1 + \sum_{k \in J} \exp\left(\delta_k + (p_k, X_k^d, \tau_{ik})\Pi Z_i + \gamma\tau_{ik}\right)} \tag{6}$$

Note that the denominator of the conditional choice probability is taken over all locations in

where τ is the average reported travel time in minutes between the zones j and k, f is a flexible polynomial function of distance d and regression parameters ϕ, and pop_k is the number of households in zone k. We estimate this equation on 30,934 route-level observations for which we observe any trip, using a cubic polynomial in distance. We then predict τ for all possible combinations of living zones in our model (329) and working zones reported in the data (517). There are more working zones than living zones because $i)$ living zones include suburban zones, and $ii)$ some survey respondents work in city zones where no-one in the survey lives and are thus dropped from the choice model. This predicted value then enters the utility equation in the estimation based on individual i's working zone, taken as fixed.

the city, implicitly assuming that all consumers choose from all possible neighborhoods in the city as well as the outside option.[26]

For estimation, we follow the standard two-step approach. In the first step, we use maximum likelihood to estimate the nonlinear heterogeneity parameters $\hat{\Pi}$ and the fixed effects $\hat{\delta}_j$. The likelihood function is:

$$L(\Pi, \delta | X, p, Z, \xi) = \prod_i \prod_{j \in J_i} Pr(y_i = j | Z_i, X_j, p_j, \tau_{ij}, \xi, \Pi, \delta) \tag{7}$$

In the second step, given the estimates of the location fixed effects, we can recover the level coefficients by estimating $\hat{\delta}_j = \alpha^d p_j + \beta^d X_j^d + \xi_j$ using an instrumental variables regression. Our demand-side price instruments follow Bayer, Ferreira and McMillan (2007), taking the average housing and spatial characteristics within a geographic "donut" around the neighborhood centroid. We select our IVs from a set of housing characteristics that includes the paved road share, RCMA, housing stock age, average units per building, and density and spatial characteristics that include the favela share of zone area, flood-zone share of zone area, average slope, and metro station presence. These IVs follow the logic of between-neighborhood competition: if location j is surrounded by higher quality zones, then j must lower its price to attract residents, implying a strong first stage. However, the nearest neighborhoods to j may create direct quality spillovers, violating the exclusion restriction. As such they are excluded from the calculation, creating a "donut" around j of neighborhoods that only influence the choice problem through their indirect impact on price in j.[27]

We estimate the model on a sample of 24,800 households with at least one employed member using the 2017-18 City of Sao Paulo Origin-Destination commuting survey, which contains information on housing and employment location choice, as well as household demographics. To measure commuting zone-level amenities we use data on the universe of residential housing units in Sao Paulo obtained from the IPTU tax records, and prices are measured according to the Properati listing price data. We collapse these variables – measured as of 2017 – into averages for each of the 329 commuting zones in our data. To improve computational performance in the

[26] In another version of the model, we restrict individual-specific choice sets choice set to the outside option and all of the locations in the consideration set J_i. To construct this consideration set at the individual-level, we first assume that the location of individual employment is exogenous. We then identify, for each individual, all of the locations j that are connected to i's location of work in the commuting matrix. This then becomes the consideration set J_i. In other words, we assume that, having decided where to work, individuals then make their housing location choice among the options allowed by their place of work. The results are similar.

[27] We include neighborhoods from 5-20 miles from j in the average characteristics of competitors.

estimation routine, all variables are standardized by subtracting the mean value across commuting zones and dividing by the across commuting zone standard deviation. We estimate standard errors on all parameters by double-bootstrapping both the first and second stage of the estimation with 500 replications.[28]

5.1.2 Demand estimation results

Table 3 presents our estimated demand parameters. The columns indicate the nine commuting zone characteristics (price, travel time, RCMA, age, units, density, paved, income and education), for which we allow demand to vary by demographics. The household size, age, renter, income, and college degree rows correspond to characteristics of the household. Bootstrapped standard errors are in parentheses. The base coefficient indicates the utility change from a one standard deviation increase in the corresponding column variable at the mean level of the demographic variables in the rows (because the household demographic characteristic variables are standardized to be centered around zero as well).[29] The base coefficients are taken from column (9) in Table A6, which compares several different specifications of the demand-side IVs.[30]

Focusing on these base coefficients, we find a negative elasticity of demand with respect to price and travel times. We find a positive elasticity with respect to RCMA, which implies that conditional on price and travel time to a given workplace, households prefer to live near areas with many high paying jobs. This could be driven by households with multiple working members (our travel time variable is only defined for the head of household), or due to other benefits of neighborhoods with high paying jobs. Regarding neighborhood characteristics, we find consumers dislike old housing stock and also dislike greater density within a given building (units). However, they like denser neighborhoods (density) that have better infrastructure (paved). The base coefficient on income is negative, suggesting there is an average preference for living around

[28] Bootstrap results for the price coefficients can be found in Figure A15. Distributions of estimates from the bootstrap procedure for other parameters look similar.

[29] For context in interpreting magnitudes, summary statistics of the main demand variables are in Table A5.

[30] Table A6 shows results from different instrumental variable specifications. Column (1) presents the OLS estimate, column (2) includes only the average X characteristics of competitors, column (3) includes only spatial characteristics, and column (4) includes both sets of IVs. Columns (5)-(9) contain different subsets of the most powerful IVs, as indicated in the Table footer. The price coefficient is smallest in the OLS regression, at 0.79, and increases in magnitude to roughly 1.5-3.9 for the IV the different specifications. This downward bias in the OLS estimate of the price elasticity is consistent with the standard simultaneous equations bias in supply and demand systems. The IV specifications vary substantially in strength, with the strongest being the parsimonious single-IV specifications using RCMA and density in columns (5) and (6). Still, the price coefficients remain relatively stable across IV models. Since using multiple IVs increases the amount of information used for identification – although at the cost of first-stage power – our preferred specification in (9) uses the subset of four jointly strongest instruments. These are the favela share, slope, RCMA, and age.

Table 3: Estimated demand parameters

Demographic	Price (1)	Travel time (2)	RCMA (3)	Age (4)	Units (5)	Density (6)	Paved (7)	Income (8)	Education (9)
Household size	-0.087	0.127	0.026	-0.064	-0.071	-0.050	0.002	-0.042	-0.041
	(0.027)	(0.012)	(0.023)	(0.018)	(0.040)	(0.022)	(0.013)	(0.031)	(0.038)
Age	-0.064	-0.153	-0.044	0.056	-0.151	0.005	0.035	0.060	0.123
	(0.026)	(0.011)	(0.024)	(0.019)	(0.029)	(0.021)	(0.017)	(0.033)	(0.039)
Renter	-0.060	-0.122	0.004	0.198	-0.002	0.034	0.037	-0.003	-0.015
	(0.026)	(0.011)	(0.022)	(0.019)	(0.027)	(0.018)	(0.017)	(0.032)	(0.036)
Income	-0.077	-0.053	0.039	0.110	-0.145	0.081	0.115	0.467	-0.005
	(0.034)	(0.023)	(0.033)	(0.036)	(0.060)	(0.024)	(0.046)	(0.041)	(0.058)
College degree	-0.185	-0.047	-0.017	-0.011	-0.163	0.068	0.058	-0.314	0.880
	(0.036)	(0.014)	(0.031)	(0.026)	(0.050)	(0.025)	(0.030)	(0.037)	(0.046)
Base coefficients	-1.975	-2.413	0.749	-0.980	-0.647	0.765	0.090	-0.420	0.572
	(0.454)	(0.021)	(0.225)	(0.124)	(0.432)	(0.159)	(0.104)	(0.250)	(0.294)

Results are from the estimation of demand-side preference parameters using two-step maximum likelihood and 2SLS. Top row gives variable names, while leftmost column gives the demographic variables. Estimation sample is 329 commuting zones and 24,800 individual households. All location characteristics including price are standardized relative to the zone-level sample mean and standard deviation. Travel time is normalized across all individual-zone combinations. Base coefficients are from column (9) of Table A6, which instruments for housing prices using the average spatial and housing characteristics of zones 5-20 miles from a zone centroid. These characteristics are favela share of zone area, slope, RCMA, and housing stock age. Bootstrapped standard errors with 500 replications in parentheses.

lower income people - the majority of residents in Sao Paulo are lower income. [31]. The base coefficient on education is positive, suggesting a preference for living near higher education people. All demographics and neighborhood characteristics are standardized, so base coefficients are interpreted as the preference for a one standard deviation increase in the characteristic for the demographically "average" household.

The other rows in Table 3 present the coefficient estimate on the interaction term between the neighborhood characteristic in the column and the household characteristic in the row. For example, the negative coefficient -.185 on College degree in the first column indicates that the price elasticity of college graduates is more negative than that of the average household. The most important interactions between neighborhood characteristics and demographics are as follows. Higher income and college educated households appear to be more sensitive to price changes in this sample, and have somewhat greater taste for local density. As far as taste for number of units in the building, older, richer, and more educated households have greater dis-utility from

[31] Even though we assume that education and income are exogenous, these variables are demographic characteristics that could be correlated with the average unobserved quality of the neighborhood, potentially biasing the coefficients on other characteristics. Instead of instrumenting for demographics, in Appendix Figure A14 we plot price estimates from the second-stage 2SLS estimation, constraining the income and education taste parameters at various multiples of their estimated values, from 0 to 3 (as seen on the x-axis). We find this sensitivity test leaves our main price estimate essentially unchanged.

living in buildings with greater units. Column (8) shows that higher income households have a preference for living near other high income households. Column (9) suggests that high education households also prefer to live near higher education households. These patterns suggest strong income and education-based geographical sorting.

5.2 Model of Residential Supply

We model the construction of new residential housing as an exponential function of building density restrictions, housing prices, and location characteristics.[32] In particular, we estimate the following supply equation for location j:

$$E[s_j|p_j, X_j^s, M_j] = \exp(\alpha^s p_j + \beta^s X_j^s + \psi M_j) \tag{8}$$

where s_j is the total number of building permits in location j and p_j is the average residential listing price in j. X_j^s is a vector of other housing and regulatory characteristics of location j that affect developer profits. This includes construction density in j, the average building age, the average number of units per building, and the average value of the pre-2016 zoning parameters and the 2016 non-BAR zoning parameters, across all blocks in j.[33] M_j is the maximum allowable BAR in neighborhood j. We model the conditional expectation of permits as an exponential function (Poisson regression) because of the discrete count nature of this outcome variable.

We model the supply location choice at the subprefeitura-quantile-level. We do this in order to exploit exogenous variation from the regression discontinuity of the 2016 reform to identify the supply-side parameters. First, we place all of the city blocks into 40 quantiles of our regression discontinuity running variable (i.e. distance to the 2016 zoning change boundary). We then collapse the variables X, p, M in to a subprefeitura by quantile level dataset. This formulation allows us to preserve the reduced-form relationship between (binned) distance, BAR, and permitting activity identified in Section 4; our estimation procedure will focus on comparisons between permit outcomes for the quantiles just above and below the BAR increase cut-off within subprefeituras.

The Poisson regression in equation 8 contains two endogenous variables, p_j and M_j, while other physical location characteristics are assumed exogenous. In order to overcome this endogeneity problem, we estimate the supply model using the GMM estimator of Mullahy (1997) with

[32] This can be considered a simplified version of Calder-Wang (2021), who applies the BLP discrete choice framework to the supply side of a housing market.

[33] We weight all blocks equally. Summary statistics for our supply variables are in Table A7.

additive errors to form moment conditions.[34] The instrument set is $W_j^s = [T_j, v_j, X_j^s]$. The first instrument T_j is used for maximum BAR, and it is an indicator for whether the subprefeitura-quantile j is treated by the 2016 reform. To leverage the fuzzy regression discontinuity design for identification of ψ, we also idnclude D_j and $D_j \times T_j$ as control variables in the vector X_j^s. We instrument for market listing prices p_j in the subprefeitura-quantile using average assessed land values from the IPTU dataset, v_j. The dependent variable is the number of new building permits in location j that are filed in the post-reform period (2016-2019). Since the 2016 reform also affected parameters aside from maximum BAR, we control for the pre-and-post reform levels of all other zoning parameters in X_j^s to isolate the causal effect of changing BAR restrictions on supply behavior.[35]

We estimate the supply model using data on 3,611 new building permits filed with the Sao Paulo city government between 2016-2019. The right-hand-side variables are taken from either IPTU or the block-level zoning maps and collapsed to the subprefeitura-quantile-level. In total, we obtain 1182 subprefeitura by quantile observations, of which 809 have any permitting activity over this period.

Table 4 presents the estimated supply coefficients. Column (1) estimates the model using only the instrument for M_j, while column (2), our preferred specification, instruments for both M_j and p_j. In column (2), the estimates imply that a one-unit increase in maximum allowable BAR at the mean leads to approximately 5.2 additional new building permits for the average unit.[36] This estimate is in line with the RD estimates from Table 2. The price coefficient in column (2) implies a 1,000 reais increase in price (18.7% of the mean) is associated with a 10.5 percent increase in permits, or 0.48 new permits at the mean, though the price term is not always significant. Column (3) uses a multiplicative error structure to form the moment condition, yielding similar results, with a slightly larger BAR effect and price effect. In columns (4)-(7), we show that BAR effects are large for multifamily permits, and negative and statistically insignificant for single family permits, consistent with the reduced form results in Section 4 (although they are noisily estimated here and so we cannot reject moderate sized effects for single-family homes).

[34] We consider the robustness of the estimates to a simple log-linear specification - which drops locations with zero permits - and find qualitatively similar results. We prefer the Poisson method to a simple log-transformation of the outcome, as it preserves all observations in the data. Results are also similar using multiplicative errors, see column (3) of Table 4.

[35] These other parameters are the maximum shadow ratio and the minimum and basic BAR of 2004 and 2016, as well as the max BAR of 2004. All of these parameters are averaged across blocks within the subprefeitura-quantile.

[36] $5.2 = 4.55 * (\exp(.757) - 1)$, where 4.55 is the average of the outcome variable.

Table 4: Supply estimates: Poisson IV regressions

Outcome	All new buildings			Multi-family		Single-family	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Max BAR	0.752***	0.753***	0.847***	0.903***	0.906***	-0.519	-0.499
	(0.175)	(0.180)	(0.257)	(0.200)	(0.206)	(0.366)	(0.356)
Price	0.103***	0.100	0.115*	0.018	0.008	0.074	0.117*
	(0.034)	(0.062)	(0.069)	(0.042)	(0.086)	(0.055)	(0.064)
Density	0.087	0.093	-0.075	0.178	0.193	0.470**	0.384*
	(0.134)	(0.139)	(0.193)	(0.162)	(0.174)	(0.208)	(0.216)
Age	0.023***	0.023**	0.038***	0.027***	0.028**	0.010	0.004
	(0.008)	(0.010)	(0.012)	(0.009)	(0.013)	(0.014)	(0.015)
Units per building	-0.005	-0.005	0.007	-0.015*	-0.015	-0.007	-0.008
	(0.006)	(0.006)	(0.005)	(0.009)	(0.010)	(0.009)	(0.009)
Historical preservation	-0.530	-0.535	-1.070**	-0.831**	-0.843**	-0.197	-0.170
	(0.352)	(0.358)	(0.450)	(0.418)	(0.423)	(0.661)	(0.661)
Q	4.416e-31	1.307e-31	4.935e-28	2.010e-31	4.829e-31	4.497e-32	1.666e-28
Observations	1182	1182	1182	1182	1182	1182	1182
IVs	BAR	BAR, Price	BAR, Price	BAR	BAR, Price	BAR	BAR, Price

Robust standard errors in parentheses. Results are from the estimation of a supply-side fuzzy regression discontinuity exponential (Poisson) model, estimated with GMM-IV, on the sample of subprefeitura-quantiles. All models use an additive error specification to form moment conditions, except column (3) which uses a multiplicative error. All specifications include controls for the running variable interacted with the treatment, maximum shadow ratio, minimum and basic BAR of 2004 and 2016, as well as the max BAR of 2004 (zoning variables averaged within subprefeitura-quantile). Q-statistic refers to the value of the GMM criterion function at the optimal parameters. Market prices are instrumented with land values. The outcome variable is the number of total new building permit, single-family, or multi-family applications between 2016-2019, as indicated in table header. Estimation sample is all subprefeitura-quantiles. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5.3 Equilibrium

With the estimated supply side parameters $\hat{\theta}^s$ and demand side parameters $\hat{\theta}^d$ in hand, we can solve for equilibrium in the residential housing market. Our counterfactual exercises will consist of imposing an exogenous zoning map M (the policy experiment), obtaining $p(M)$, a J-vector of counterfactual prices $p = [p_1, ..., p_J]$ in each location j such that supply and demand are equated under M. We conduct the equilibrium analysis at the commuting zone level. We analyze the following map M scenarios: 1) A baseline equilibrium that takes observed zone-level market shares in 2016 (just prior to the 2016 reform), and estimated demand parameters and calculates the price vector necessary to equate supply and demand.[37] 2) A counterfactual where we simulate the model for ten years given the 2004 zoning map. 3) A counterfactual where we simulate the model for ten years given the 2016 reform zoning map. 4) A "double BAR" counterfactual where we keep BAR at 2004 levels for blocks that in reality received lower BARs in the 2016 reform, and double the ultimate 2016 BAR for blocks that received an increase in BAR in the 2016 reform. In simulations (2)-(4) we run the supply model for 10 years to estimate impacts of new permits on

[37] As a validation exercise we will compare these model computed prices with the observed 2016 prices.

new construction, prices, residential sorting, and welfare.

To calculate the equilibrium prices for a given zoning map, we first take the estimated supply parameters $\hat{\theta}^s$ and use them to calculate $S_j(p; M, X^s, \hat{\theta}^s)$, the market share of total housing supply in location j for a given price vector, supply characteristics X^s, and zoning map M. Then, using the demand parameters $\hat{\theta}^d$, we calculate $D_j(p; X^d, \hat{\theta}^d)$, the predicted market share of location j given prices and the demand structure. The equilibrium condition is that supply equal demand in each commuting zone J:

$$S_j(p; M, X^s, \hat{\theta}^s) = D_j(p; X^d, \hat{\theta}^d) \; \forall j \in [1, ..., J] \tag{9}$$

So we obtain a system of J-equations in J unknowns and search for the price vector p that solves the equilibrium system.

This equilibrium condition implicitly assumes that households do not anticipate demographic changes that will occur as households respond to prices; they simply respond to price changes and then subsequently observe demographic changes, which affect their ex-post welfare. Given the 2016 reform ultimately represents a small change in the housing stock, and therefore small demographic changes, it is unlikely allowing households to forecast demographic changes would greatly change our results. Disallowing this demographic forecasting behavior also removes the problem of equilibrium selection. Bayer, McMillan and Reuben (2004) and Bayer and Timmins (2005) provide detailed analyses on the issue of multiple equilibria in residential sorting models, and in Appendix B we discuss a modified equilibrium calculation procedure in which we allow multiple rounds of demographic change.

To calculate the commuting zone-level demand shares $D_j(p; X^d, \hat{\theta}^d)$, we must aggregate the individual conditional choice probabilities to the commuting zone-level by integrating over the empirical distribution of demographics:

$$D_j(p, X^d, \hat{\theta}^d) = \int Pr(y_{it} = j | Z_i, X^d, p, \hat{\theta}^d) dF_Z \tag{10}$$

Calculating the zone-level supply shares $S_j(p; M, X^s, \hat{\theta}^s)$ is more complicated, since it requires solving two aggregation problems. First, our choice equation refers to new permits rather than the stock of buildings, so we must translate new building permits into market shares of total housing units. Second, our supply side equations are subprefeitura-quantile level and must be aggregated to the commuting zone level. For details on this aggregation procedure, see Appendix A.

Our equilibrium condition equates the *market shares* of each location as predicted by our estimated demand and supply models. Implicitly, this assumes that the population of the MSA will grow to meet the new housing stock built after a given shock to BAR. If this were not the case, then there would have to be real vacancies somewhere in the city after a positive housing supply shock in the model, since the total number of housing units would exceed the number of possible residents. In this sense, the equilibrium in shares is an "open city" model where new migrants are assumed to enter and fill in new vacancies. We believe that this is a more reasonable assumption for the city of Sao Paulo, which is the wealthiest city in Brazil and plays a somewhat similar role as that of New York City in the United States.

One potential criticism is that such "equilibrium in shares" assumption may be restrictive and lead us to underestimate price effects. An alternative assumption is that the Sao Paulo MSA is a closed economy, such that any counterfactual increases in housing supply in the city must be populated by corresponding vacancies in the suburbs. This implies an equilibrium condition in *levels*, assuming no vacancies at $t = 0$. For each counterfactual, we calculate equilibria under both the shares and levels assumptions and interpret these as upper and lower bounds, respectively, on the true counterfactual prices.

To get a sense of the importance of migration in to Sao Paulo municipality from the suburbs versus the rest of Brazil, we present summary statistics on the fraction 2010 census respondents who lived outside of the municipality as of 2005. Only 3.56 % of Sao Paulo city residents report having lived outside of Sao Paulo metropolitan area as of 2005, which translates to an average annual in-migration rate of .71 %. Among those recent migrants, 17.6% have a college degree, which is approximately 4% higher than the average college degree holding rate (from the same 2010 census data). The average income of these migrants is R$900 reals per month, which is 10% less than the Sao Paulo city average. An even smaller share, .35% of Sao Paulo city residents, report living in the suburbs as of 2005. These migrants from the suburbs are 17 percentage points more likely to have gone to college and earn approximately R$600 more per month. To summarize, the 2005 to 2010 migration in to Sao Paulo is small, and mostly comes from outside of the metropolitan area. The average characteristics are quite similar to the Sao Paulo city as a whole. Given those numbers, we also assume in our simulations that the new out of MSA migration from new construction will mechanically reflect the demographics of the city as a whole.

5.4 Counterfactual Results

As a model validation exercise, we begin by calculating the model implied prices that equate the observed commuting-zone market shares (supply) and estimated demand (based on our demand model) just prior to the 2016 reform (the "baseline scenario"). This validation does not use our estimated supply model at all; the purpose here is to get a sense of how well the model implied prices can replicate observed listing prices when using our equilibrium calculation procedure. Figure A16 shows the correlation between model-predicted prices for the baseline scenario and the observed listing price data from the Properati multiple listing price service. Our model prices do a good job of replicating the observed market prices, with an $R^2 = 0.75$. Figure A17 similarly plots the observed demographic composition of zones (log of average income and share of household heads with college education) in the data against the demographics that would be predicted by the individual-level choice probabilities of the model at baseline.[38]

Table 5 presents the main zone-level results on prices and quantities from our simulations. All of the results show model simulated outcomes ten years after 2016 (i.e. predictions for 2026). Column 1 presents model simulated results assuming the 2004 zoning reform zoning stayed in place from 2016 to 2026, with an average maximum BAR of 1.55. Column 2 presents results for a simulation where zoning is changed according to the 2016 reform map in 2016, with a higher maximum BAR of 2.09. Column 3, the "Double BAR scenario" keeps BAR at the 2004 level for those blocks that had a BAR decrease in the 2016 reform, and doubles post-2016 reform BAR for all blocks that received a BAR increase in the 2016 reform. Even under the Double BAR scenario, the average maximum BAR in the city is still just 3.49, which is substantially lower than the BAR levels observed in the most permissive zoning regimes in the world.[39]

Row 1 of each panel gives the total new units that are created within the city in 10 years as a result of the corresponding zoning policy; the totals exclude units created as a result of the city-wide growth trend. Row 2 shows the share of the total within city stock that the new units in row 1 represent. In Panel A, relative to a 10 year continuation of the 2004 zoning rules (Column 1), the model predicts that the 2016 zoning reform will produce approximately 54,123 net new housing units, or an approximate 2.2 percent increase in the housing stock of the city. The relatively small aggregate effect on the supply of housing is consistent with the fact the average BAR in the city only increased by 0.54, and did not change or was reduced in 48% of city blocks. In Column 3

[38] The R^2 are 0.89 and 0.96 for log income and college-educated share, respectively.
[39] For example, Singapore has many plots allowing BAR levels in the 8 to 10 range.

Table 5: Simulation results: prices and quantities

Scenario	2004 zoning	2016 zoning	Double BAR
Max BAR	1.55	2.09	3.49
Panel A: Equating shares			
New units (ths)	184.277	238.400	1170.403
New units (share of stock)	0.075	0.097	0.476
Avg. price (ths of reales)	6.104	6.070	5.471
Share Living in City	0.588	0.593	0.657
Avg. zone income	4.822	4.812	4.596
Avg. zone education	0.307	0.306	0.279
Panel B: Equating levels			
New units (ths)	181.500	233.691	1032.682
New units (share of stock)	0.074	0.095	0.420
Avg. price (ths of reales)	5.951	5.869	4.216
Share Living in City	0.611	0.622	0.790
Avg. zone income	4.803	4.787	4.493
Avg. zone education	0.306	0.304	0.268

Table shows zone-level results from equilibrium simulations under three different zoning scenarios, as indicated in table header. Double BAR scenario holds BAR constant at 2004 levels for all locations where BAR was reduced in 2016, and doubles the post-2016 BAR value in all locations where BAR was increased in 2016. First row shows average block-level maximum allowable BAR under each scenario. Panel A equates market shares in the equilibrium condition, implicitly assuming that all new construction within the city is occupied by new migrants from outside the MSA. Panel B equates levels in the equilibrium condition, implicitly assuming that all new construction in the city is occupied by migrants from the outside the city but within the MSA. Share Living in City is fraction of households living within the municipality. Avg. zone income is in thousands of reales per household per year. Avg. zone education is fraction of households with a college degree.

we see that in the Double BAR reform scenario the housing stock in São Paulo increases by 40.1% relative to the 2004 zoning reform.

Row 3 gives the model-predicted average zone-level price in thousands of reais per square meter. Prices are similar in Columns 1 and 2, falling by only 0.55%, indicating that the 2016 reform has small impacts on average housing prices in São Paulo. Even under the Double BAR scenario we estimate only a 10.4% decrease in prices on average. One reason is that the suburbs are assumed to grow at an annual rate (1.6%) that is faster than the city (1%); so although the increase in housing units in the city is large relative to the stock of city housing, the supply increase relative to the total metropolitan area is smaller and therefore the price response is also commensurately smaller. Another reason is that there is latent demand to move from the suburbs to the city as shown in Row 4, with "Share living in city" increasing as we look at larger zoning reforms. But if we assume instead that the MSA is a closed economy, as in Panel B, we see substantially larger price reductions from the 2016 and Double BAR scenarios, at 1.4% and 29.2%, respectively.

Rows 5-6 show that average zone-level demographics do not change much in the city, even with a very large shock. To explore the muted change in average city demographics, Figure A18 plots the average demographics for city and suburbs under the different scenarios considered so far.[40] The "Double BAR" scenario provides some insight in to why average city demographics are not changing in response to the zoning reform. Relative to the other scenarios, the "Double BAR" scenario shows that the major zoning reform in the municipality has the largest ramifications for the demographics of households living in the suburbs. At baseline, the suburbs are lower income and less educated. The large zoning shock leads to lower prices in the city, and the demand model suggests that higher income and more educated individuals will be more responsive to these lower prices, and therefore more likely to occupy the new housing. Because of this positive selection effect, these new residents are similar to the original city population. As a result, the city demographics do not change much in response to the major zoning reform, while the suburban demographics change in response to out-migration of higher income and higher education individuals to the city. These simulations therefore suggest that the new housing built in São Paulo in response to even a major zoning reform are unlikely to house low income residents.

Figures 5 and 6 show that while there is only a small aggregate effect of the 2016 reform – consistent with the small overall change in the housing shock – there is substantial heterogeneity

[40] The baseline scenario reflects our implied demographics from the demand side model prior to the 2016 reform.

across the city, mostly predicted by where BAR changed most.[41] Figure 5 plots the distribution of zone-level changes in prices and units, and then correlates them with the average BAR change within a commuting zone. As expected, places with larger BAR shocks see more construction and lower prices. Quantitatively, the largest price reductions are roughly $322 reais per meter squared, or around 6.2% of that zone's market price under the 2004 counterfactual. The largest supply shocks are approximately 1995 additional units, or around 20.9% of that zone's housing stock under the 2004 counterfactual.[42] Figure 6 visualizes, geographically, the BAR change in the 2016 reform, as well as the model simulated predictions for changes in the number of housing units, prices, and market shares. BAR changes were larger in more outlying areas of the city, rather than in the denser core where BAR actually fell in many cases. As such, these more outlying areas saw lower prices and gained market share in a way that maps directly on to the BAR change.

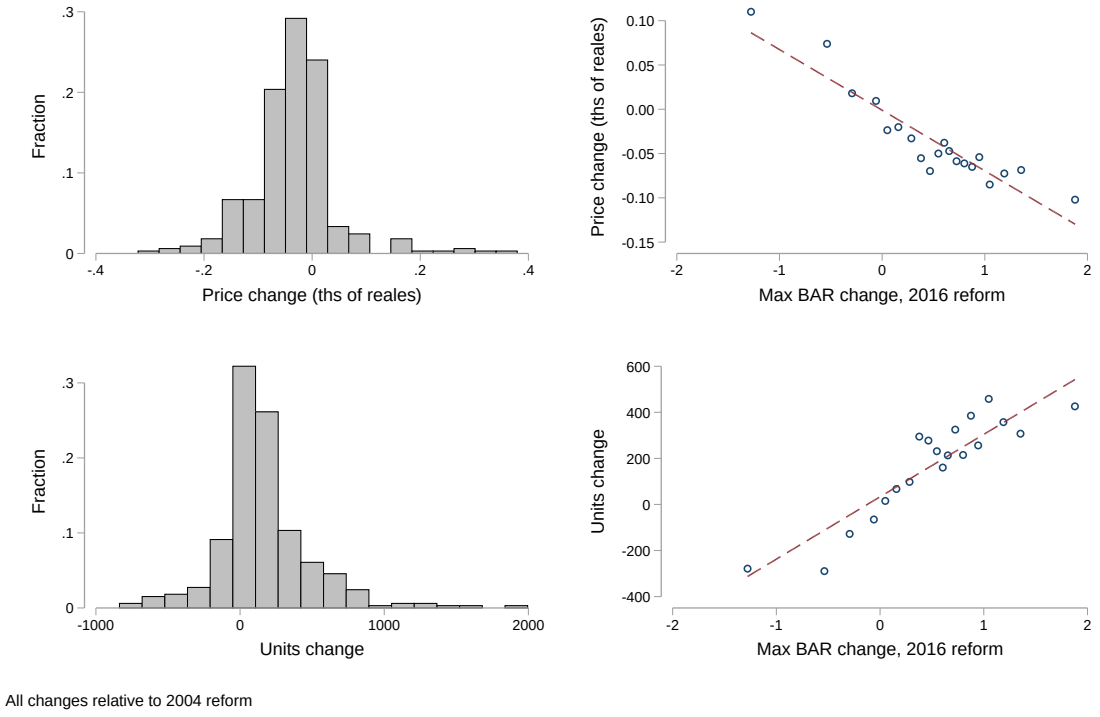


Figure 5: Equilibrium changes under the 2016 reform

[41] Note that all these results reflect changes in outcomes relative to the scenario where the 2004 zoning reform persisted for 10 years and reflect the equilibrium in market shares.

[42] If instead we solve for equilibrium in levels, we find that the largest price effects are around 7.2% of the zone's price under the 2004 counterfactual.

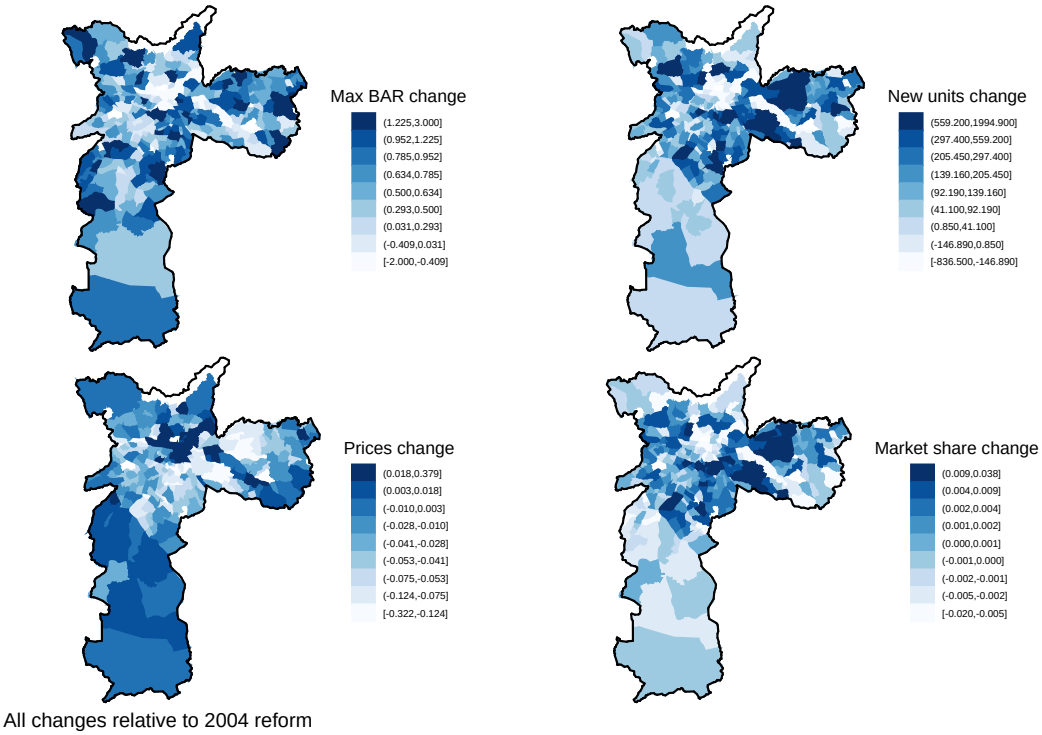


Figure 6: Equilibrium changes under the 2016 reform: map

5.4.1 Welfare analysis

Table 6 presents our estimates of the 10 year welfare impacts of the 2016 reform and Double BAR scenarios relative to a continuation of the 2004 zoning policy.[43] The estimates in the table give per-household average consumer surplus gains in reais, except the bottom row which gives MSA-wide aggregate gains in millions of reais.[44] In columns (1) and (4) we update only the prices in calculating welfare in the new equilibrium, while in columns (2) and (5) we update prices and demographic and neighborhood characteristics. Focusing first on column (1), our model predicts that there is an unequal distribution of surplus across demographic groups. The average college educated household has an estimated gain relative to the 2004 zoning regime of 44.71 reais over ten years (approximately $13.89 USD in total at 2016 exchange rates), and the highest income quintile has a similar $40.46 reais gain after ten years. High income and college groups are more price sensitive in our demand model, and therefore obtain the largest gains from the new housing supply. The aggregate welfare gains over 10 years are $156.03 million reales, or about $48.5 million

[43] Table A8 presents corresponding results for the equilibrium in levels.

[44] Consumer surplus is calculated as the "inclusive value," also know as the logsum, from Small and Rosen (1981), which gives the ex-ante expected value of a utility-maximizing choice for consumer i, normalized by the individual-specific price coefficient.

USD.

Table 6: Simulation results: individual-level consumer surplus

Scenario	2016 zoning			Double BAR		
Update	P	X	τ	P	X	τ
	(1)	(2)	(3)	(4)	(5)	(6)
By demographic group						
Owner	28.35	133.23	131.74	466.15	4114.54	4003.49
Renter	31.36	107.01	105.10	569.30	3567.34	3455.95
Non-college	24.83	111.55	110.29	406.29	3534.46	3426.26
College	44.71	182.01	179.18	805.01	5598.43	5476.56
By income quintile						
1	22.78	103.33	102.22	370.46	3344.02	3238.10
2	24.38	111.68	110.39	393.54	3474.29	3367.57
3	26.97	122.19	120.71	447.63	3796.57	3686.69
4	31.63	140.61	138.76	538.81	4288.04	4172.89
5	40.46	158.24	155.94	723.10	5058.55	4939.82
Totals						
Full sample	29.10	126.68	125.08	491.93	3977.77	3866.64
Aggregate consumer surplus (mm reales)	156.03	679.23	670.67	2637.63	21327.95	20732.05

Table shows per-household expected change in consumer surplus from equilibrium simulation of the 2016 zoning reform and Double BAR reform for different subgroups, measured in Brazilian reais. Bottom row shows the total consumer surplus aggregating across all households in millions of reais. Columns (1) and (3) update only equilibrium prices from the 2016 reform scenario, while columns (2) and (4) update both prices and the housing and neighborhood attributes included in X_j. Columns (3) and (6) update all variables, including travel time τ. All changes are evaluated relative to 2004 (status quo) zoning.

Column (2) shows there is approximately 4.4 times larger consumer surplus gains when we allow welfare to respond to changes in the demographic and neighborhood characteristics induced by our zoning reform. The aggregate welfare gain rises to \$679.2 million reais, or about \$210.9 million USD. This is primarily because density increases and the age of housing units go down.[45] These results suggest that most of the value of zoning comes through the presence of a newer housing stock and greater density, as opposed to price reductions.[46] Interestingly, the unequal distribution of surplus across income quintiles actually falls slightly once we incorporate changes

[45] As discussed earlier, our model predicts that demographic change in the city in response to the 2016 zoning reform is minimal.

[46] In Table A9 we decompose the effect of each housing characteristic and confirm that age and density are the largest contributors to these welfare gains.

in neighborhood characteristics, because poorer people have a greater taste for newer units. Inequality also emerges between renters and owners after updating X because owners prefer newer buildings. Finally, after accounting for effects on characteristics, the Double BAR scenario in column (5) produces much larger consumer surplus gains, at $3,977.8 reais ($1.24 billion USD) per household, or $21.3 billion ($6.62 billion USD) in aggregate. Lastly in columns (3) and (6), we account for the role of increased congestion arising from greater densification, given households' disutility of commuting time. These congestion effects erode the aggregate welfare gains by 1.3% and 2.8% for the 2016 reform and the Double BAR scenario, respectively.[47]

Table 7: Comparison of welfare estimates and other policy effects

Scenario	2016 Reform		Double BAR	
Units	R$ bi	% of GDP	R$ bi	% of GDP
Change in welfare	0.67	0.10	20.73	2.96
Change in value - homeowners	-11.11	-1.59	-170.74	-24.39
Change in value - landlords	-3.89	-0.56	-63.92	-9.13
Change in developer profits	4.47	0.64	74.01	10.57
Change in productivity	0.92	0.13	16.89	2.408
New resident welfare gains	0.0068	0.001	3.81	0.54

Table shows calculations for policy effects on consumer welfare, housing wealth, and developers' profits, from equilibrium simulations under two different zoning scenarios, as indicated in table header. Double BAR scenario holds BAR constant at 2004 levels for all locations where BAR was reduced in 2016, and doubles the post-2016 BAR value in all locations where BAR was increased in 2016. All gains and losses are evaluated relative to the 2004 status quo simulation. Developers are assumed a 10% profit margin on total housing revenues. Homeowners refers to value of all owned-occupied units, landlords refers to value of all rental units. GDP refers to 2017 total output of Sao Paulo city. All simulations equate market shares in the equilibrium condition, implicitly assuming that all new construction within the city is occupied by new migrants from outside the MSA.

The results above described changes in welfare for households in the Sao Paulo metropolitan area, according to their expected value of all neighborhood options. However, zoning reforms present other costs and benefits that may influence the political passage of such densification measures. In Table 7 we show aggregate estimates for five additional measures of costs and benefits: changes in nominal housing wealth for existing homeowners and landlords, changes in developer profits, welfare gains for new residents, and changes in city productivity. The first row of Table 7 reprints the final aggregate welfare gain estimates from Table 6 for the equalizing shares equilib-

[47] Appendix Figure A19 displays individual-level changes in expected commuting time and decomposes these changes into location and population effects.

rium. The second and third rows show how housing wealth varies by type of reform relative to the 2004 policy status quo. Our simulation predicts that homeowners and landlords will face nominal house prices losses of R$11.11 and R$3.89 billions due to the 2016 reform, respectively. Together, these losses are much larger than the welfare gain of $0.67 billion due to more housing options, and can partially explain why existing real estate owners fear zoning reforms that promote more densification.[48] The contrast between changes in house values and welfare is even higher for the Double BAR scenario, given the larger price effects.

On the other hand, row 4 shows that developer profits increase by approximately R$4.47 billion.[49] The aggregate losses to local homeowners are larger than the gains to developers, suggesting that if lobbying efforts for zoning reform are proportional to the potential gains/losses from the zoning, cities around the globe would be unlikely to approve a zoning reform similar to the one implemented in Sao Paulo, much less a Double BAR style reform. Row 5 shows a back-of-the envelope calculation of the potential effect of zoning reform on productivity. For this exercise we heavily rely on the estimates of Glaeser and Gyourko (2018).[50] We find that future gains in productivity could be R$0.92 billion, and are another potential justification for increasing densification. Finally, the last row shows potential welfare gains for new residents, given that more housing allows the migration of new households.[51] Given the relatively small number of extra new houses built because of the 2016 reform, the aggregate consumption gains from new residents is also small. Gains for new residents become sizeable only in the Double BAR scenario, as a much larger number of people can now move to the city.

[48] One important caveat with such housing wealth losses is that housing consumption does not change as house prices fall. Moreover, homeowners are always able to sell their existing homes and buy other houses of similar quality given that price drops occur across the city. Landlords, on the other hand, mostly care about returns on their investment, so the nominal losses represent real reductions in profits.

[49] Predicted developer profits are calculated as the total value of new developed housing units times a mark up of 10% which is a conservative estimate of mark ups for real estate development.

[50] Glaeser and Gyourko (2018) find that a dramatic zoning reform across all cities in the United States - that allow the movement of enough workers to equalize wages across all cities - would generate a GDP gain of 2%. Assuming Sao Paulo in Brazil plays an analogous role in the Brazilian economy as that of New York city in the United States economy, and also accounting for the relative magnitude of new housing developed due to the 2016 reform in Sao Paulo, we can back out potential productivity gains. See Online Appendix for details of the calculation. Perfectly replicating the Glaeser and Gyourko (2018) analysis to Brazil would require data collection for all Brazilian cities and is out of the scope of this paper.

[51] For this calculation we use the conservative assumption that new residents, on average, would have the same welfare benefit as that of existing residents.

6 Conclusion

The impact of zoning on housing markets was one of the most important and hotly debated subjects prior to the Covid-19 pandemic, and is likely to regain its status as the pandemic ebbs and increases in cost of living return to the forefront of household worries. In this paper we contribute to the literature by studying the impact of a 2016 zoning reform in São Paulo that increased the ability of developers to supply housing units by lifting limitations on permitted densification on a block-by-block basis. Using a spatial discontinuity design and the timing of the reform, we find that developers responded swiftly to obtain approximately 65 percent more permits in blocks that relaxed zoning rules.

More importantly, we develop a framework to estimate welfare of local residents by integrating the spatial RD design in to a supply and demand model of residential housing, where we can estimate structural demand and supply parameters, and then simulate ten year forward outcomes of prices, built environment, household sorting, and welfare. Our framework accounts for both costs and benefits of densification, which allows for a more complex picture of the effects of zoning reforms.

We find small aggregate price and supply changes due to the fact the aggregate supply response induced by the 2016 reform is small relative to the São Paulo housing stock. Our welfare analysis suggests that higher income and education groups benefit the most from the reform, due to their greater price sensitivity and the ability to move from the suburbs to areas in the city that are closer to workplaces. Moreover, welfare gains are five times larger when accounting for changes in the built environment of the city, especially with respect to density and age of buildings. This suggests a fair amount of the economic value of zoning reforms comes through the presence of newer housing stock, as opposed to lower prices.

We also show that more aggressive zoning reforms produce much larger welfare gains. In fact, the framework developed here can produce results for any desired reform, including estimating heterogeneity in effects by neighborhood. Finally, we show that such reforms negatively impact the housing wealth of existing homeowners and landlords, which may generate political backlash in the form of NIMBYism - even though those same homeowners should not fear dramatic changes in the socio-economic composition of neighbors.

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm and Nikolaus Wolf. 2015. "The Economics of Density: Evidence from the Berlin Wall." *Econometrica* 83(6):2127–2189.

Allen, Treb, Costas Arkolakis and Xiangliang Li. 2016. "Optimal City Structure." *Dartmouth College, mimeograph* .

Ashar, Sandeep. 2018. "Mumbai: Plan for Taller Buildings, Higher Densities Around Metro, Rail Corridors on Hold." *Indian Express, https://indianexpress.com/article/cities/mumbai/mumbai-plan-for-taller-buildings-higher-densities-around-metro-rail-corridors-on-hold-5333536/* .

Balboni, Clare, Gharad Bryan, Melanie Morten and Bilal Siddiqi. 2020. "Transportation, Gentrification, and Urban Mobility: The Inequality Effects of Place-Based Policies." *Stanford University, mimeograph* .

Baum-Snow, Nathaniel and Lu Han. 2021. "The Microgeography of Housing Supply." *Submitted, Journal of Political Economy* .

Bayer, Patrick and Christopher Timmins. 2005. "On the Equilibrium Properties of Locational Sorting Models." *Journal of Urban Economics* 57(3):462–477.

Bayer, Patrick, Fernando Ferreira and Robert McMillan. 2007. "A Unified Framework for Measuring preferences for Schools and Neighborhoods." *Journal of Political Economy* 115(4):588–638.

Bayer, Patrick, Robert McMillan and Kim Reuben. 2004. An Equilibrium Model of Sorting in an Urban Housing Market. Technical report NBER Working Paper 10865.

Berry, Steven, James Levinsohn and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63(4):841–890.

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.

Brueckner, Jan K and Somik V Lall. 2015. "Cities in Developing Countries: Fueled by Rural–Urban Migration, Lacking in Tenure Security, and Short of Affordable Housing." *Handbook of Regional and Urban Economics* 5:1399–1455.

Bryan, Gharad, Jonathan De Quidt, Tom Wilkening and Nitin Yadav. 2017. Land Trade and Development: A Market Design Approach. Technical report CEPR Discussion Paper No. DP12136.

Calder-Wang, Sophie. 2021. "The Distributional Impact of the Sharing Economy on the Housing Market." *University of Pennsylvania, mimeograph* .

Combes, Pierre-Philippe, Gilles Duranton and Laurent Gobillon. 2021. "The Production Function for Housing: Evidence from France." *Journal of Political Economy* 129(10):2766–2816.

Diamond, Rebecca and Tim McQuade. 2019. "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development." *Journal of Political Economy* 127(3):1063–1117.

Ding, Chengri. 2013. "Building Height Restrictions, Land Development and Economic Costs." *Land Use Policy* 30(1):485–495.

Dougherty, Conor. 2020. "California, Mired in a Housing Crisis, Rejects an Effort to Ease It." *New York Times, https://www.nytimes.com/2020/01/30/business/economy/sb50-california-housing.html* .

Duranton, Gilles and Diego Puga. 2020. "The Economics of Urban Density." *Journal of Economic Perspectives* 34(3):3–26.

Epple, Dennis, Brett Gordon and Holger Sieg. 2010. "A New Approach to Estimating the Production Function for Housing." *American Economic Review* 100(3):905–24.

Ganong, Peter and Daniel Shoag. 2017. "Why Has Regional Income Convergence in the U.S. Declined?" *Journal of Urban Economics* 102:76–90.

Giaquinto, Paulo Ricardo et al. 2010. "Planos diretores estratégicos de São Paulo, nova roupagem velhos modelos." *Universidade Presbiteriana Mackenzie* .

Glaeser, Edward and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives* 32(1):3–30.

Glaeser, Edward L, Joseph Gyourko and Raven E Saks. 2005. "Why Have Housing Prices Gone Up?" *American Economic Review* 95(2):329–333.

Gyourko, Joseph, Jonathan Hartley and Jacob Krimmel. 2021. "The Local Residential Land Use Regulatory Environment Across US Housing Markets: Evidence from a New Wharton Index." *Journal of Urban Economics* 124:103337.

Gyourko, Joseph and Raven Molloy. 2015. "Regulation and Housing Supply." *Handbook of Regional and Urban Economics* 5:1289–1337.

Harari, Mariaflavia. 2020. "Cities in Bad Shape: Urban Geometry in India." *American Economic Review* 110(8):2377–2421.

Hsieh, Chang-Tai and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11(2):1–39.

Mullahy, John. 1997. "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior." *Review of Economics and Statistics* 79(4):586–593.

Murphy, Alvin. 2018. "A Dynamic Model of Housing Supply." *American Economic Journal: Economic Policy* 10(4):243–67.

Paciorek, Andrew. 2013. "Supply Constraints and Housing Market Dynamics." *Journal of Urban Economics* 77:11–26.

Saconi, Rose and Carlos Entini. 2013. Como era São Paulo sem Plano Diretor. Technical report. Accessed: 2021-04-07.

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

Sandroni, Paulo. 2010. Captura de mais valias urbanas em São Paulo através do binômio solo criado/outorga onerosa: análise do impacto do coeficiente de aproveitamento único como instrumento do plano diretor de 2002. Technical report.

Small, Kenneth A. and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49(1):105–130.

Tabarrok, Alex and Tyler Cowen. 2018. Skyscrapers and Slums: What's Driving Mumbai's Housing Crisis? Technical report Marginal Revolution University.

Trounstine, Jessica. 2018. *Segregation by Design: Local Politics and Inequality in American Cities*. Cambridge University Press.

Tsivanidis, Nick. 2019. "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio." *Working Paper* .

Online Appendix: Estimating the Economic Value of Zoning Reform

Santosh Anagol Fernando Ferreira Jonah Rexer

*Anagol: Wharton School, University of Pennsylvania. Ferreira: Wharton School, University of Pennsylvania and NBER. Rexer: School of Public and International Affairs, Princeton University. We are grateful for support from the Wharton Dean's Research Fund, and the Research Sponsors Program of the Zell/Lurie Real Estate Center. We thank Tom Cui, Anna Gao, Renan Muta, Sophia Winston, Alexandru Zanca, and Holly Zhang for excellent research assistance. We also thank Rohan Ganduri and seminar participants at Imperial College Business School, Wharton Urban Lunch, European Urban Economics Association meeting, American Urban Economics Association meeting, NBER Summer Institute Real Estate, FGV-SP, Brazilian Society of Econometrics, LACEA-LAMES, NBER Public Economics, and the SITE conference at Stanford for helpful comments and suggestions.

A Supply side aggregation

Let q be the subprefeitura-quantile $\in [1, ..., Q]$. Then the predicted annual number of new building permits for q is:[1]

$$\hat{s}_q = \frac{1}{4} \exp(\hat{\alpha}^s p_q + \hat{\beta}^s X_q^s + \hat{\psi} M_q) \tag{11}$$

Each new permit is associated with a time-path of new housing units. To obtain this, we take a sample of permits which can be matched to our IPTU data and calculate the cumulative expected number of residential units $\hat{n}_t$ that will be constructed from the average permit, for each year t over a ten year horizon.[2] So each permit is associated with $\hat{s}_q \hat{n}_t$ units. The model-predicted number of new units for location q by year τ, then, is:

$$N_{q,\tau} = \sum_{t=0}^{\tau} \hat{n}_t \hat{s}_q \tag{12}$$

This formula accounts for the fact that, each year into our simulation, new permits are being filed at a constant rate implied by the predicted values of the supply equation. Finally, to obtain the market share of total units for q after 10 years, we add the new units to the existing stock, allowing for differential secular growth rates between the city, r_1, and the outside option, r_0.[3]

$$S_{q,\tau} = \frac{N_{q,\tau} + N_q^0 (1 + r_1)^\tau}{N_0^0 (1 + r_0)^\tau + \sum_{k=1}^{Q} N_{k,\tau} + N_k^0 (1 + r_1)^\tau} \tag{13}$$

S_q is defined at the subprefeitura-quantile-level but our equilibrium prices and quantities must be returned at the commuting zone-level. However, neighborhood-quantiles are not nested in commuting zones. As such, we construct the following mapping between the two. First, we overlay the maps of 1182 neighborhood-quantiles on to the 329 commuting zones and calculate the area of intersection between every q, j pair. Define weights $\omega_{qj} = \frac{km_{qj}^2}{km_q^2}$ as the share of the area in neighborhood-quantile q that falls into commuting zone j. Then, to translate a price vector p_j into p_q to be plugged into the supply equation, we calculate the weighted average of prices in all the zones that overlap with q:

[1] Note that we measure the outcome as the total number of permits for the four years from 2016-2019. So in order to annualize the predicted number of permits, we divide the fitted values by 4.

[2] In our matched sample, by year 10 the average new building permit will create 19 new residential units.

[3] We obtain these growth rates from census data on aggregate housing unit growth from 2000-2010, and estimate $r_1 = 0.01$ and $r_0 = 0.017$; over this period the suburbs have grown .7 percent per year faster in terms of housing units.

$$p_q = \sum_{j=1}^{J} \omega_{qj} p_j \tag{14}$$

Similarly, to translate a set of shares S_q into S_j for the equilibrium calculation, we apportion each neighborhood-quantile share to each of its constituent zones in proportion to their area share and then aggregate up to the commuting zone level:

$$S_j = \sum_{q=1}^{Q} \omega_{qj} S_q \tag{15}$$

B Multiple equilibria and demographic sorting

When neighbors' demographics enter the utility function, then households' decisions may depend on the decisions of others. This strategic complementarity introduces the possibility of multiple equilibria in residential sorting models. Bayer, McMillan and Reuben (2004) defines a "sorting equilibrium" as a set of prices, choice probabilities, and demographics satisfying two conditions: i) supply and demand are equalized *and* ii) households make optimal location decisions accounting for the choices of all other households. Our equilibrium condition in Section 5.3, which follows Calder-Wang (2021), does not account for the external effects of individual choices in ii). As such, there may be other equilibrium prices at which supply and demand are equated, but which obtain distributions of individual demographics and choice probabilities that vary substantially.

We can think of equilibrium selection as imposing one of two opposing assumptions: i) that households are completely myopic, and do not anticipate the demographic balance that will result from the aggregation of their individual choices (our assumption), and ii) that households have perfect foresight, that is, that the demographics at which they make their choices are those that are obtained in equilibrium (the sorting equilibrium). In this section, we show that there exist a continuum of equilibria in which supply equals demand, defined by the difference between ex-ante and ex-post demographics.

We investigate the presence of multiple equilibria as follows. We first run the equilibrium simulation using the 2004 status quo zoning map and initial neighborhood demographics Z_0 (from the data), obtaining our price vector, p_0. We then calculate the implied zone-level demographics based on the equilibrium choices at p_0, yielding Z_1. We then evaluate demand $D(p_0, Z_1)$, while holding supply fixed at $\bar{S} = S(p_0, M, X^s, \theta^s)$, where M is the 2004 zoning map. It may well be that

supply no longer equals demand, since the equilibrium was obtained at Z_0. Therefore, we solve for p_1 such that $D(p_1, Z_1) = \bar{S}$. We then update the individual choices and obtain the implied Z_2. We continue this way until convergence, that is, until $p_{n-1} = p_n$ and $Z_{n-1} = Z_n$.[4] The initial equilibrium thus represents the myopic assumption, while the final equilibrium corresponds to perfect foresight, since the demographics at which the individuals make their decisions are fulfilled in equilibrium. The equilibria obtained along this convergence path are each uniquely determined by the gap between ex-ante and ex-post demographics, and can therefore be thought of as imposing different assumptions on the extent of "forecast error" by agents in the model.

Figure A20 shows the convergence path by plotting the average zone-level price. Relative to the initial equilibrium, subsequent iterations are characterized by higher prices on average. These prices stabilize around p_{20}, and are about 7% higher than p_0. Still, prices are relatively highly correlated over iterations, though this diminishes along the convergence path. Figure A21 shows that the R^2 of prices is generally above 0.8. This correlation is, however, greatly reduced by the presence of a single outlier zone that has very large increases in price along the convergence path. Dropping this outlier increases the R^2 substantially, though the price distribution is still shifted to the right.

While prices are relatively highly correlated across equilibria, demographics are not. As shown in Figure A22, demographics shift substantially when we account for foresight. The histograms show a much more skewed income distribution, with a large mass of low-income zones and a long tail of rich neighborhoods by the final iteration. Education now follows a multi-modal distribution, with mass at the extremes and a missing middle. When individuals account for the future demographic composition of neighborhoods ex-ante in their decisionmaking, they sort into highly segregated enclaves, rather than choosing only based on relative prices and built environment attributes, a phenomenon consistent with the intuition of neighborhood tipping models. This dynamic arises from the strong positive assortative preferences we observe in Table 3. However, while the perfect foresight assumption yields internally consistent results, ultimately we do not observe such sorting equilibria in the data. Instead, the naive equilibrium remains highly correlated with the data (see Figures A16 and A17) and produces much more realistic demographic and price distributions. We therefore maintain this myopic assumption for our welfare analysis, while noting that the impacts of reform may differ under alternate assumptions.

[4] We can appeal to Brower's fixed point theorem to guarantee demographic convergence. In practice, we do not impose a formal convergence criterion. Rather, we run 40 iterations of this procedure, which delivers a stable equilibrium.

C Zoning reform effect on productivity

Our estimates of the productivity effects of zoning reform are heavily based on the assumptions and estimates from Glaeser and Gyourko (2018).[5] Let L_i be the quantity of labor in location i, F_L^i be the marginal product of labor in location i and W the average national wage. Their work assumes that differences in payroll per worker can be considered the true differences in marginal product of labor. From that assumption they consider the thought experiment of moving populations from all areas with low initial wages to all areas with high wages until wages equalize to a similar level (W) across all locations. In this context, the gains from relocation can be written as:

$$Gains = \frac{1}{2\alpha} \sum_i L_i(F_L^i(L_i) - W) \tag{16}$$

α is the inverse elasticity of labor demand. In this set up equalizing wages will generate a reduction in the total wage bill and the output gain from reallocation will be proportionate to the total wage bill reduction.[6]

Glaeser and Gyourko (2018) use data across all MSAs in the U.S. and an estimate of α from the literature to calibrate a 2 percentage change in GDP resulting from a radical reallocation of labor that equalizes wages across all locations. If α=1, then a 33.3 percent increase in population will drop wages in the New York MSA to the national norm.

We use that information to estimate a simple back-of-the envelope calculation, only considering the effects of the Sao Paulo zoning reforming, and ignoring a potential equalization of wages across all cities in the country. Our counterfactual simulation estimates that the reform would increase population in Sao Paulo by an extra 2.2 percentage points in 10 years. Assuming that Sao Paulo plays a similar economic role in Brazil as that of of NYC in the US, the increase in population is 0.0661 of the effect required to equalize wages, assuming linearity of effects.

The Sao Paulo share of national GDP is 9.46%, which means that the reform would generate gains through reallocation of 2% GDP*9.46%*0.0661 = 0.0125% of the Brazilian GDP. That in turn corresponds to 0.132% of Sao Paulo GDP. A similar calculation was conducted for the double BAR simulations.

[5] See Appendix 3 of that paper for details on the calculation method and necessary assumptions.

[6] The reduction in the total wage bill comes intuitively from the fact that formerly high wage areas with stringent zoning restrictions now attract a lot of labor leading to large wage declines relative to low-wage unrestricted places. The key assumption is that that curvature of the marginal product curve is stronger in the restricted versus non restricted areas. The output gain is proportional to this because the higher the wages where in the restricted areas, the greater the productivity gains from labor re-allocation.

D Appendix Tables

Table A1: RD First stage, land use fixed effects

Outcome	Maximum BAR change			
	(1)	(2)	(3)	(4)
Panel A: 2004 Zone FE				
Treat BAR	1.498***	1.437***	1.378***	1.355***
	(0.00535)	(0.00669)	(0.00771)	(0.00933)
Panel B: 2004 Zone FE, sub-prefeitura FE				
Treat BAR	1.521***	1.460***	1.391***	1.357***
	(0.00546)	(0.00672)	(0.00776)	(0.00947)
Panel C: 2016 Zone FE				
Treat BAR	1.239***	1.109***	1.036***	1.003***
	(0.00414)	(0.00544)	(0.00630)	(0.00782)
Panel D: 2016 Zone FE, sub-prefeitura FE				
Treat BAR	1.179***	1.086***	1.018***	0.982***
	(0.00442)	(0.00533)	(0.00628)	(0.00784)
Specification	Base	Linear	Quadratic	Cubic
Observations	43231	43231	43231	43231
Mean of Dep. Variable	0.597	0.597	0.597	0.597

Robust standard errors in parentheses. Specification refers to the order of the polynomial for the running variable, which is distance to the RD boundary. The polynomial is always interacted with the treatment indicator. Sample is all city blocks with zoning information. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A2: RD reduced form, land use fixed effects

Outcome	New multi-family building permits			
	(1)	(2)	(3)	(4)
Panel A: 2004 Zone FE				
Treat BAR	0.00159***	0.00313***	0.00384***	0.00396***
	(0.000243)	(0.000303)	(0.000351)	(0.000427)
Panel B: 2004 Zone FE, sub-prefeitura FE				
Treat BAR	0.00217***	0.00259***	0.00306***	0.00365***
	(0.000266)	(0.000309)	(0.000357)	(0.000428)
Panel C: 2016 Zone FE				
Treat BAR	0.0000169	0.00185***	0.00228***	0.00219***
	(0.000273)	(0.000335)	(0.000380)	(0.000452)
Panel D: 2016 Zone FE, sub-prefeitura FE				
Treat BAR	0.000155	0.000620	0.000955*	0.00142**
	(0.000303)	(0.000346)	(0.000386)	(0.000453)
Specification	Base	Linear	Quadratic	Cubic
Observations	43231	43231	43231	43231
Mean of Dep. Variable	0.00470	0.00470	0.00470	0.00470

Robust standard errors in parentheses. Specification refers to the order of the polynomial for the running variable, which is distance to the RD boundary. The polynomial is always interacted with the treatment indicator. Sample is all city blocks with zoning information. Mean of dependent variable calculated for control blocks within 0.1 km of the BAR boundary. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A3: RD reduced form: approved permits

Outcome	New multi-family building permits			
	(1)	(2)	(3)	(4)
Panel A: No sub-prefeitura FE				
Treat BAR	0.000993***	0.00229***	0.00261***	0.00239***
	(0.000155)	(0.000209)	(0.000245)	(0.000305)
Specification	Base	Linear	Quadratic	Cubic
Observations	43231	43231	43231	43231
Mean of Dep. Variable	0.00218	0.00218	0.00218	0.00218
Panel B: With sub-prefeitura FE				
Treat BAR	0.00116***	0.00148***	0.00163***	0.00183***
	(0.000173)	(0.000210)	(0.000245)	(0.000301)
Specification	Base	Linear	Quadratic	Cubic
Observations	43225	43225	43225	43225
Mean of Dep. Variable	0.00218	0.00218	0.00218	0.00218

Robust standard errors in parentheses. Specification refers to the order of the polynomial for the running variable, which is distance to the RD boundary. The polynomial is always interacted with the treatment indicator. Sample is all city blocks with zoning information. Mean of dependent variable calculated for control blocks within 0.1 km of the BAR boundary. Outcome is the average annual new building permits filed after 2016 and approved in 2017 or later. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A4: RD reduced form: Poisson model

Outcome	New multi-family building permits			
	(1)	(2)	(3)	(4)
Panel A: No sub-prefeitura FE				
Treat BAR	0.262***	0.694***	0.704***	0.613***
	(0.0385)	(0.0518)	(0.0747)	(0.0913)
Specification	Base	Linear	Quadratic	Cubic
Observations	43231	43231	43231	43231
Mean of Dep. Variable	0.00470	0.00470	0.00470	0.00470
Panel B: With sub-prefeitura FE				
Treat BAR	0.297***	0.406***	0.512***	0.528***
	(0.0414)	(0.0541)	(0.0753)	(0.0837)
Specification	Base	Linear	Quadratic	Cubic
Observations	43225	43225	43225	43225
Mean of Dep. Variable	0.00470	0.00470	0.00470	0.00470

Robust standard errors in parentheses. Specification refers to the order of the polynomial for the running variable, which is distance to the RD boundary. The polynomial is always interacted with the treatment indicator. Sample is all city blocks with zoning information. Mean of dependent variable calculated for control blocks within 0.1 km of the BAR boundary. All models are poisson regressions estimated with maximum likelihood. *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A5: Summary statistics for demand variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Price	Travel time	RCMA	Age	Units	Density	Paved	Income	College
Mean	6.24	33.36	2.03	39.89	5.40	5.17	0.98	4.65	0.33
SD	2.35	30.61	1.11	10.38	16.89	4.74	0.05	2.28	0.26

Table shows means and standard deviations for all commuting zone-level variables that enter demand equation (3). Price is measured in R$ ths, travel time in minutes, RCMA is an index of market access (see description in-text), average age of building is in years, units is units per building for the average building in the zone, density is constructed area per unit of zone area, paved is the share of paved roads, income is measured in R$ ths, and college is the share of residents with a college degree.

Table A6: Second-stage demand estimation: IVs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Price	-0.793***	-2.274***	-1.383***	-1.521***	-2.690***	-2.491***	-1.222**	-3.884***	-1.975***
	(0.163)	(0.436)	(0.373)	(0.309)	(0.495)	(0.477)	(0.609)	(0.997)	(0.383)
RCMA	0.463***	0.822***	0.606***	0.640***	0.922***	0.874***	0.567***	1.211***	0.749***
	(0.145)	(0.198)	(0.165)	(0.167)	(0.220)	(0.206)	(0.212)	(0.337)	(0.183)
Age	-1.157***	-0.935***	-1.069***	-1.048***	-0.873***	-0.903***	-1.093***	-0.694***	-0.980***
	(0.099)	(0.111)	(0.109)	(0.100)	(0.120)	(0.118)	(0.122)	(0.197)	(0.105)
Units per building	-0.781***	-0.613***	-0.714***	-0.699***	-0.566***	-0.589***	-0.732***	-0.431***	-0.647***
	(0.049)	(0.072)	(0.066)	(0.059)	(0.079)	(0.077)	(0.081)	(0.132)	(0.067)
Density	0.708***	0.779***	0.736***	0.743***	0.799***	0.789***	0.729***	0.856***	0.765***
	(0.107)	(0.136)	(0.116)	(0.117)	(0.149)	(0.143)	(0.117)	(0.193)	(0.127)
Paved roads	0.083	0.092	0.087	0.087	0.094	0.093	0.086	0.101	0.090
	(0.064)	(0.074)	(0.067)	(0.068)	(0.078)	(0.076)	(0.065)	(0.090)	(0.071)
Average income	-0.600***	-0.375*	-0.510***	-0.489***	-0.312	-0.342	-0.535***	-0.130	-0.420**
	(0.174)	(0.220)	(0.185)	(0.187)	(0.240)	(0.229)	(0.205)	(0.322)	(0.206)
College share	0.104	0.691**	0.338	0.392*	0.855***	0.777***	0.274	1.329***	0.572**
	(0.196)	(0.278)	(0.248)	(0.232)	(0.302)	(0.297)	(0.295)	(0.495)	(0.260)
F-statistic		8.560	8.852	7.459	41.411	38.901	24.005	17.174	12.136
Observations	329	329	329	329	329	329	329	329	329
Instruments	None	X	Spatial	All	RCMA	Density	Pave	Favela	Strong

Robust standard errors in parentheses. Results are from the second step of a two-step demand estimation. The outcome variable is the mean location-specific utility term $\hat{\delta}_i$ estimated in the first step maximum likelihood procedure. All location characteristics including price are standardized relative to the zone-level sample mean and standard deviation. Instruments for housing prices are the average spatial and housing characteristics of all zones greater than 3 miles from a zone centroid. *X* instruments (2) are: paved road share, RCMA, housing stock age, average units per building, and density. Spatial instruments (3) are: favela share of zone area, flood-zone share of zone area, average slope, and metro station presence. Strong instruments (9) are the subset of jointly strongest instruments: favelas, slope, RCMA, and age.

Table A7: Summary statistics for supply variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Permits	MF Permits	SF Permits	Max BAR	Price	Density	Age	Units	Historic
Mean	4.55	3.22	0.35	2.11	5.35	0.58	34.62	5.93	0.04
SD	5.91	5.05	0.85	0.66	1.81	0.41	8.41	12.41	0.10

Table shows means and standard deviations for all subprefeitura-bin-level variables that enter supply equation. Permits is the total count of permits (total new building, multi-family, or single-family). Max BAR is the average Max BAR in 2016 in the subprefeitura-bin. Price is measured in R$ ths, average age of building is in years, units is units per building for the average building in the subprefeitura-bin, density is constructed area per unit of subprefeitura-bin area, and historic is the share of subprefeitura-bin area under historic preservation.

Table A8: Simulation results: individual-level consumer surplus, levels

Scenario	2016 zoning			Double BAR		
Update	P	X	τ	P	X	τ
	(1)	(2)	(3)	(4)	(5)	(6)
By demographic group						
Owner	75.06	175.56	174.14	1540.00	4748.49	4649.26
Renter	78.24	151.24	149.40	1657.49	4254.71	4156.38
Non-college	71.26	155.07	153.88	1471.43	4258.76	4162.88
College	92.62	222.15	219.43	1927.39	5964.22	5853.81
By income quintile						
1	69.28	147.11	146.06	1432.75	4083.88	3990.42
2	70.80	155.28	154.06	1456.19	4202.95	4108.41
3	73.50	165.67	164.26	1517.51	4490.82	4393.04
4	78.36	183.62	181.85	1621.31	4914.67	4811.65
5	88.07	198.28	196.07	1834.83	5494.92	5387.97
Totals						
Full sample	75.85	169.48	167.96	1569.36	4625.07	4526.07
Aggregate consumer surplus (mm reales)	406.70	908.71	900.55	8414.59	24798.63	24267.78

Table shows per-household expected change in consumer surplus from equilibrium simulation of the 2016 zoning reform and Double BAR reform for different subgroups, measured in Brazilian reais. Bottom row shows the total consumer surplus aggregating across all households in millions of reais. Columns (1) and (3) update only equilibrium prices from the 2016 reform scenario, while columns (2) and (4) update both prices and the housing and neighborhood attributes included in X_j. Columns (3) and (6) update all variables, including travel time τ. All changes are evaluated relative to 2004 (status quo) zoning.

Table A9: Decomposition of welfare effects

Scenario	2016 zoning (1)	Double BAR (2)
Price only	29.10	491.93
Price and age	102.42	2079.50
Price and units	28.46	470.05
Price and density	47.49	1460.55
Price and income	31.22	525.54
Price and education	26.69	446.07
Price and all X	126.68	3977.77

Table shows average individual-level welfare changes, measured in Brazilian reais from equilibrium simulation of the 2016 zoning reform and Double BAR reform. Each row represents the welfare change, relative to the 2004, from updating the variable indicated in the first row label.

E Appendix Figures

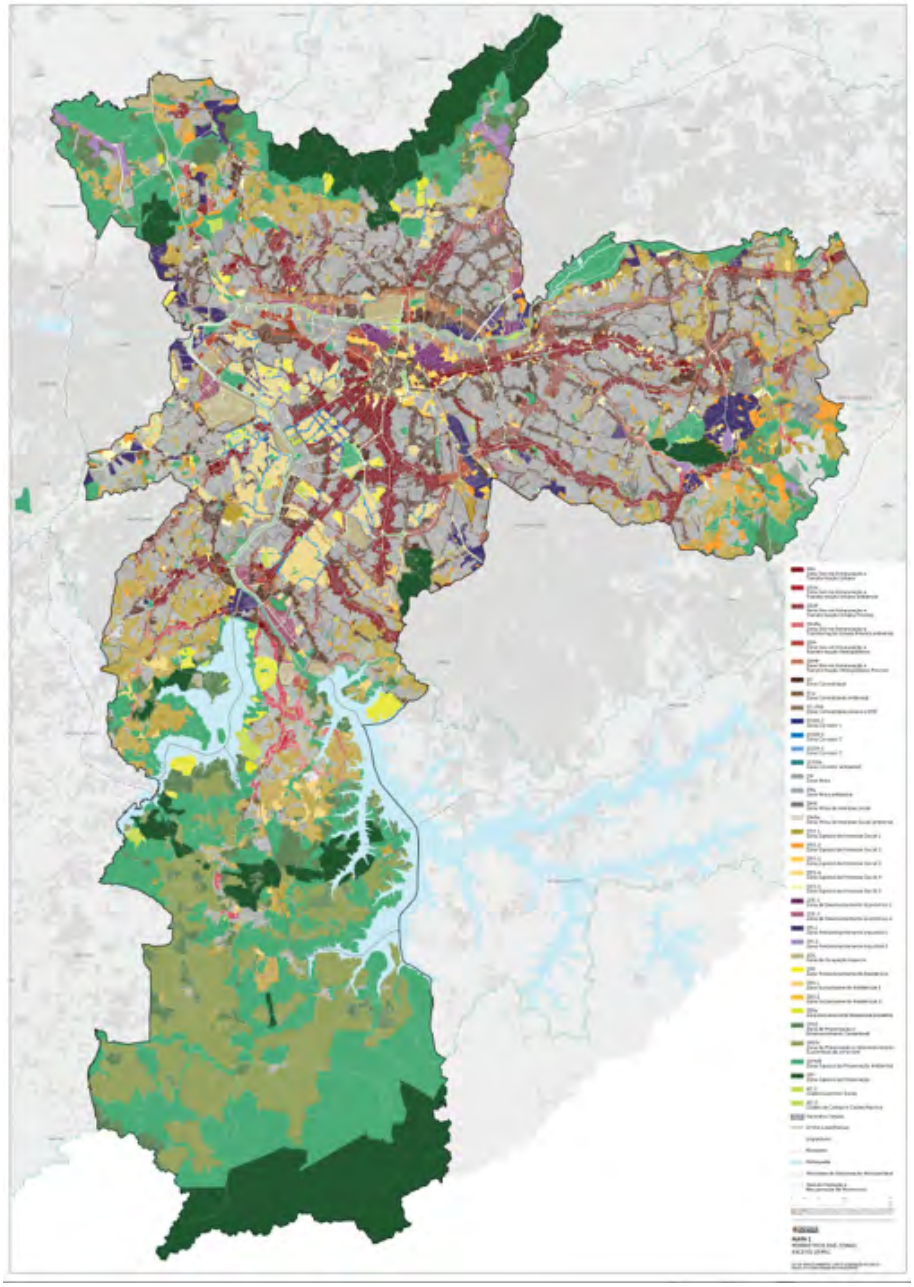


Figure A1: 2016 Zoning Reform Land Use

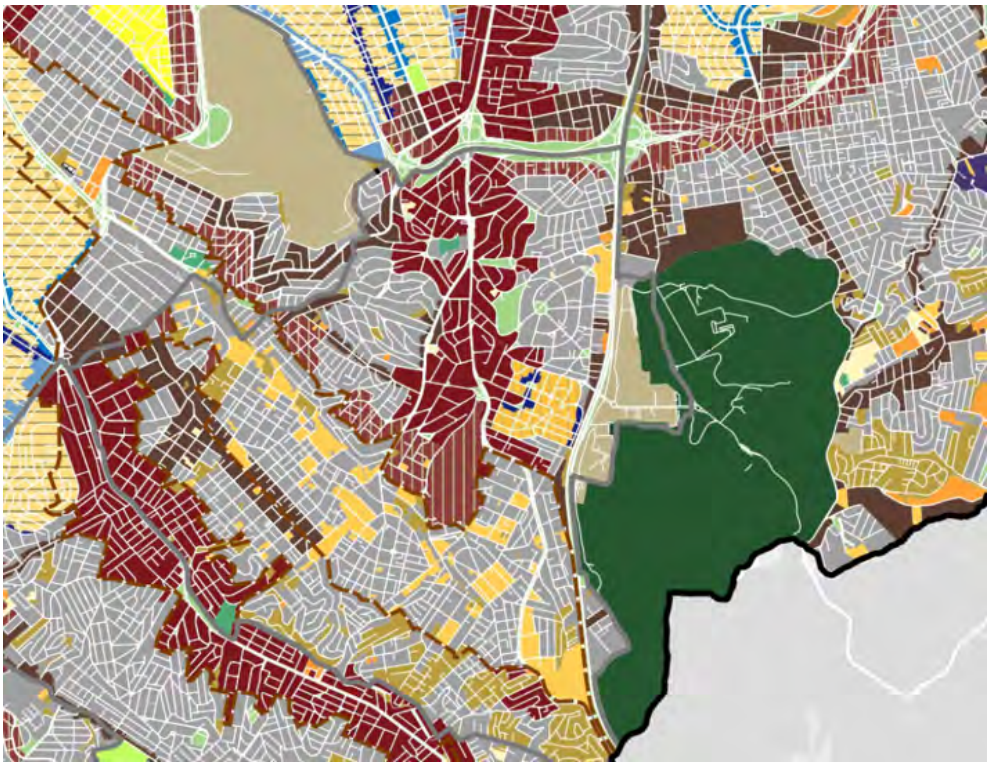


Figure A2: Block-by-block Land Use in Jabaquara Neighborhood

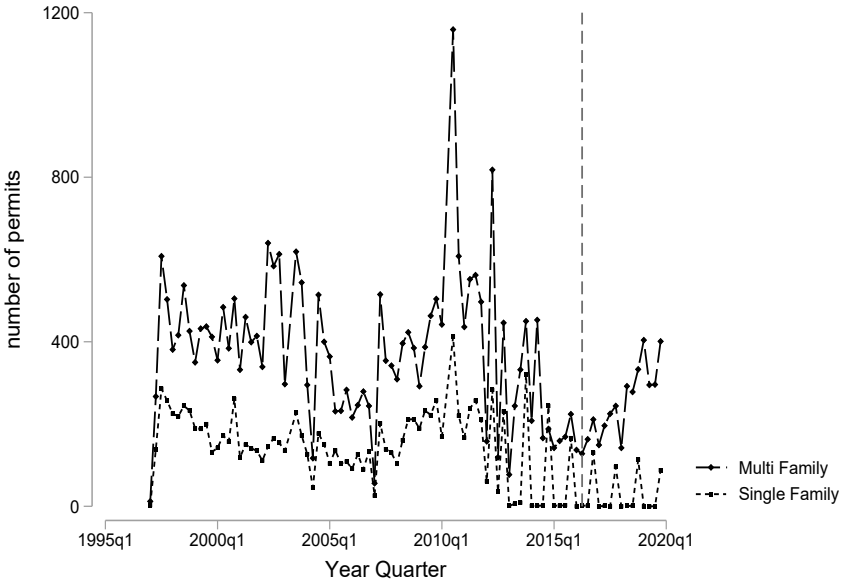


Figure A3: New Single and Multifamily Residential Building Permit Filings by Quarter

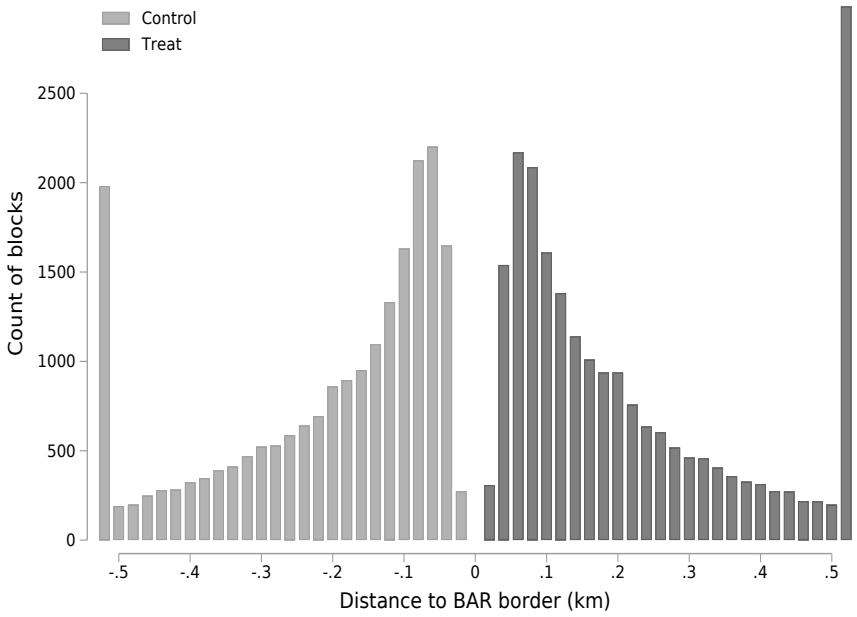


Figure A4: Histogram of blocks by running variable

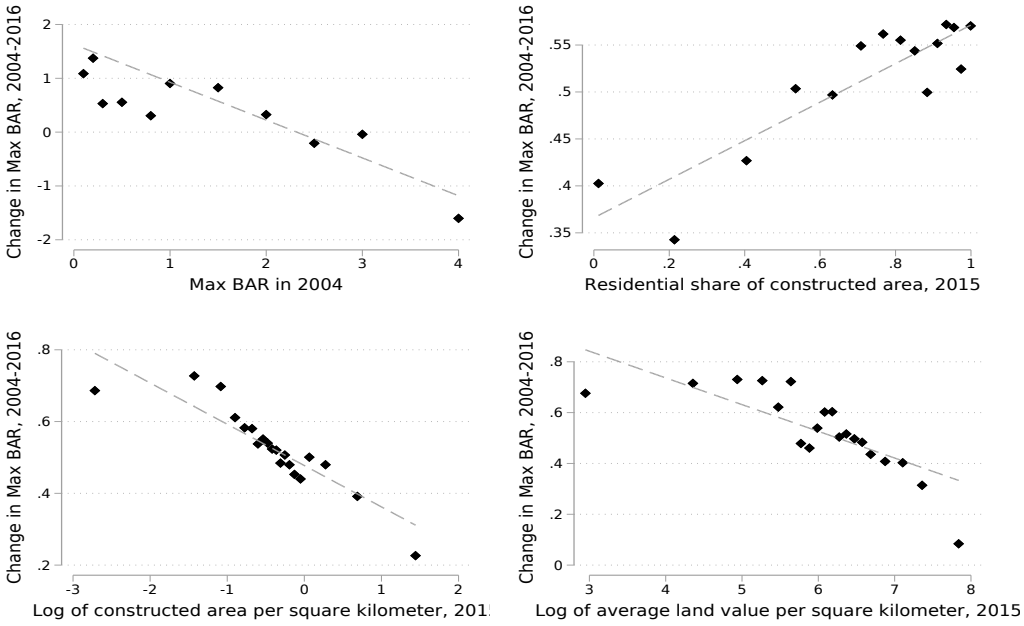


Figure A5: Correlates of Built Area Ratio Changes, Pre-to-Post 2016 Reform

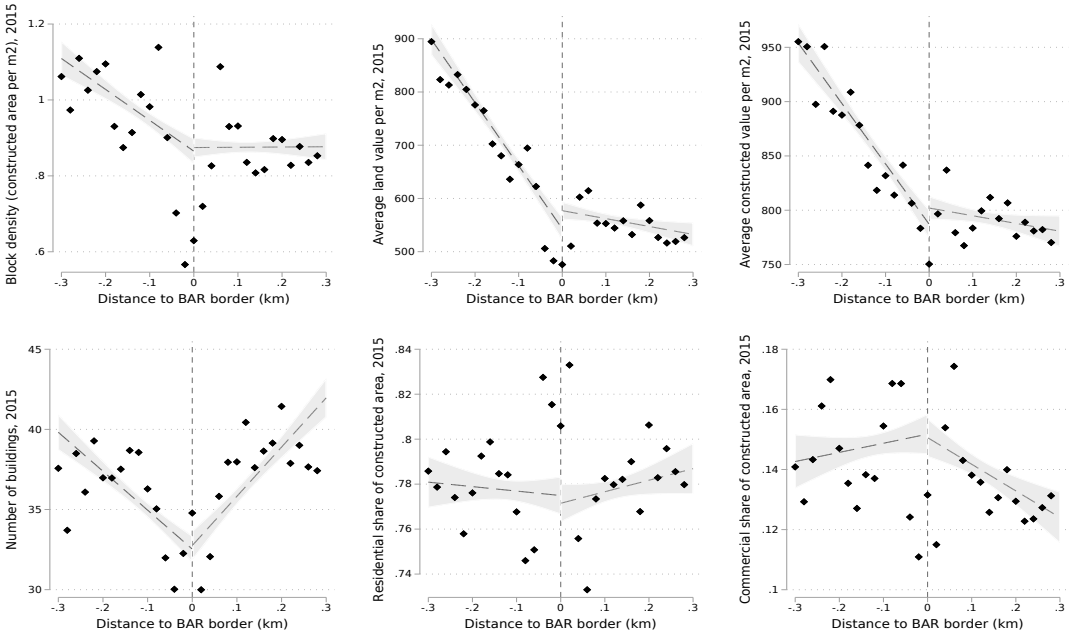


Figure A6: Average Property Characteristics in 2015 (Year Prior to 2016 Zoning Reform)

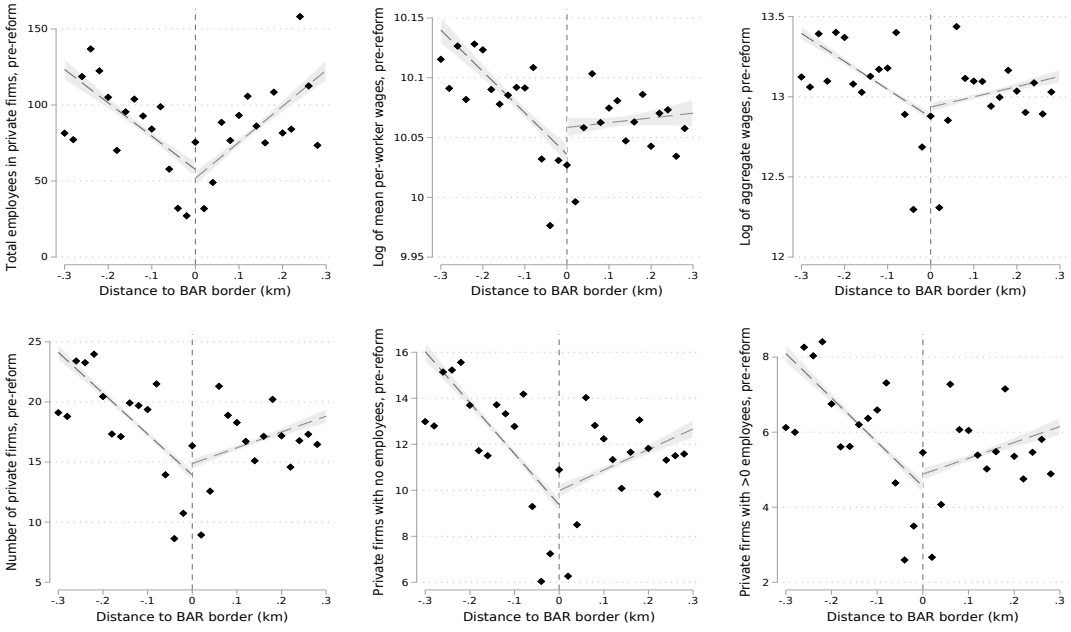


Figure A7: Average Labor Market Outcomes in RAIS Data in 2015 (Year Prior to 2016 Zoning Reform

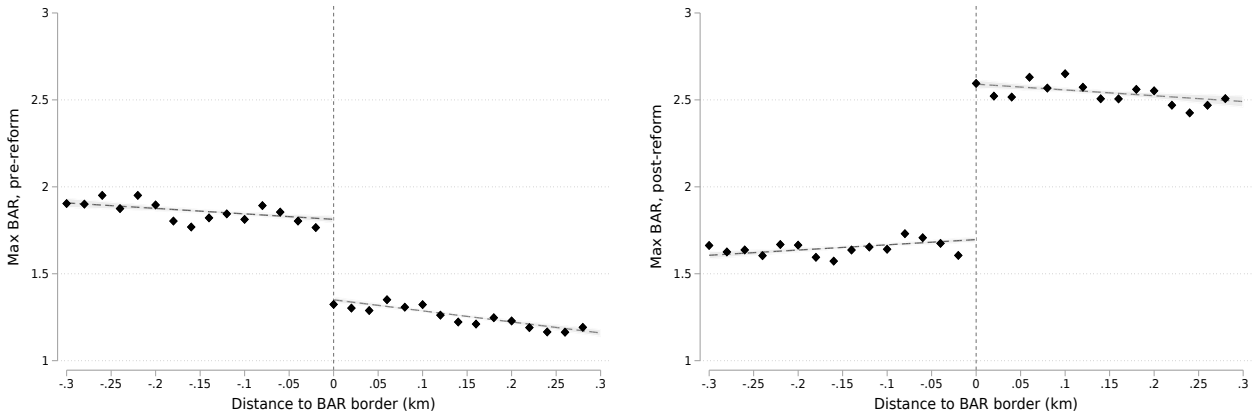


Figure A8: Built Area Ratios, Before and After 2016 Reform

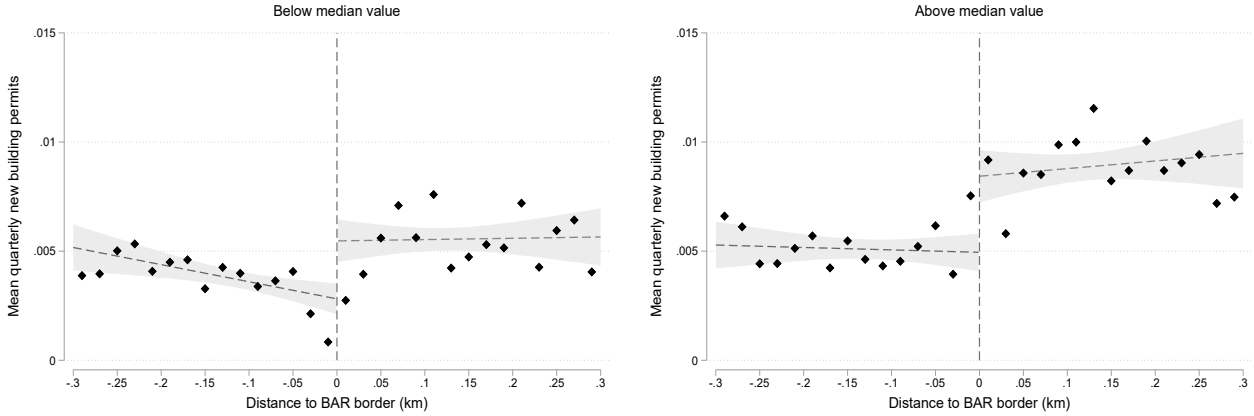


Figure A9: Heterogeneity: land values

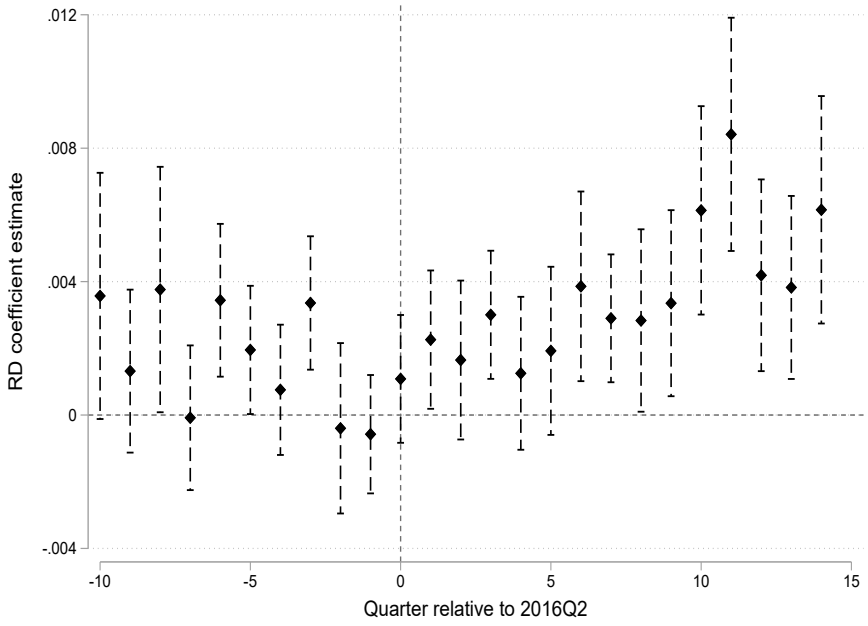


Figure A10: Dynamic RD coefficients

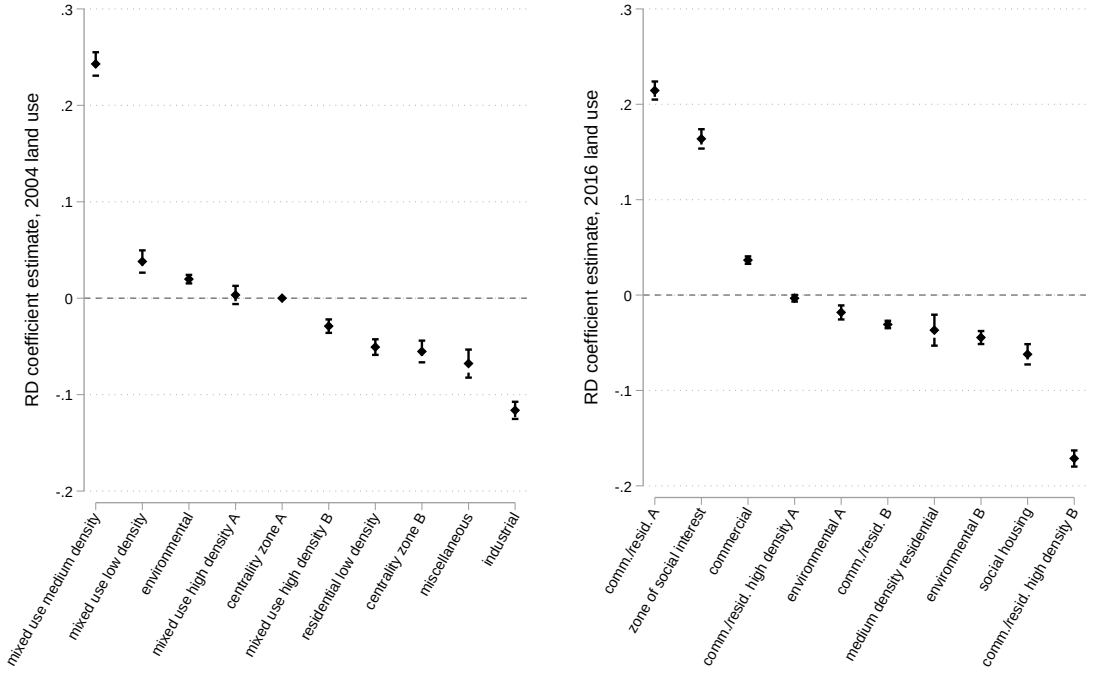


Figure A11: Land use category changes at BAR border

Note on 2004 zone types: 1) the mixed use high density A allows lower basic and max BAR than the mixed use high density B zone. 2) Centrality zones are generally areas away from major transportation corridors, and Centrality Zone B is higher density than Centrality Zone A. Note on 2016 zone types: 1) Main difference between Comm./resid A and B is B also has goal of improving and expanding public transportation infrastructure. 2) Comm./resid. high density B is similar to Comm./resid high density A except located along railways and main rivers

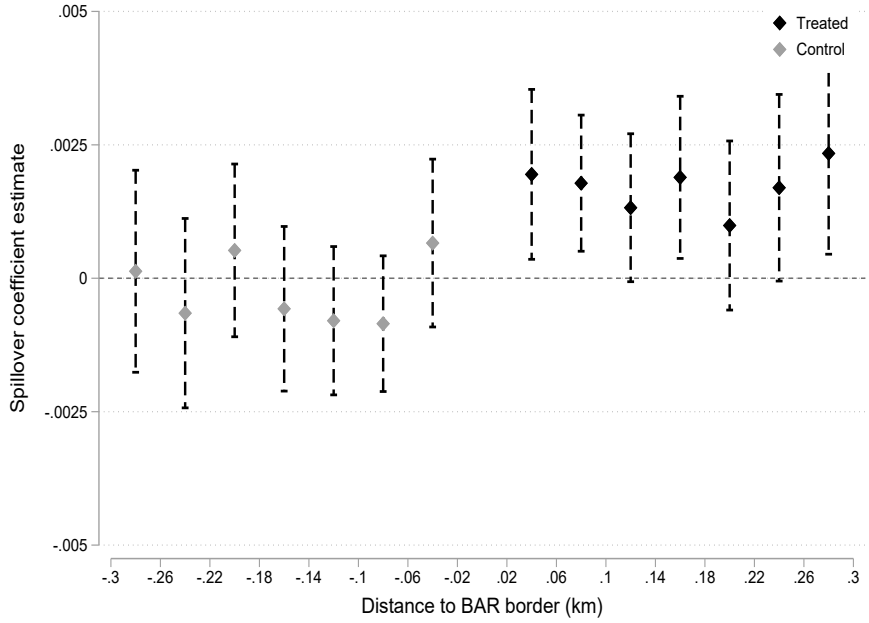


Figure A12: Estimates of Spillover Effects

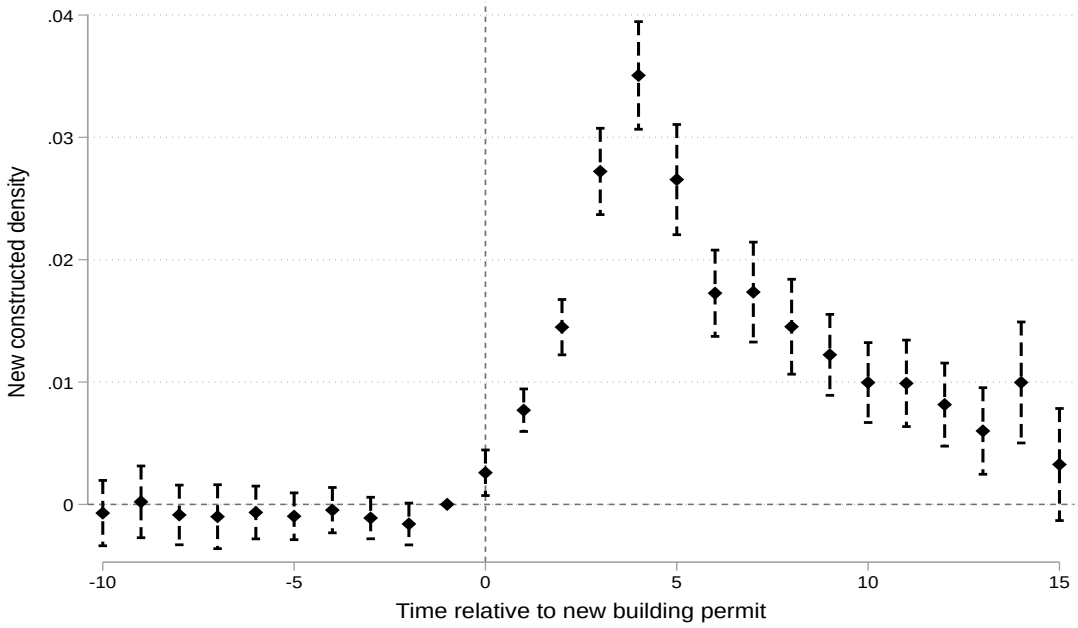


Figure A13: Event-Study Evidence on Relationship Between Block-Level Permit Issuance and Future New Construction Density

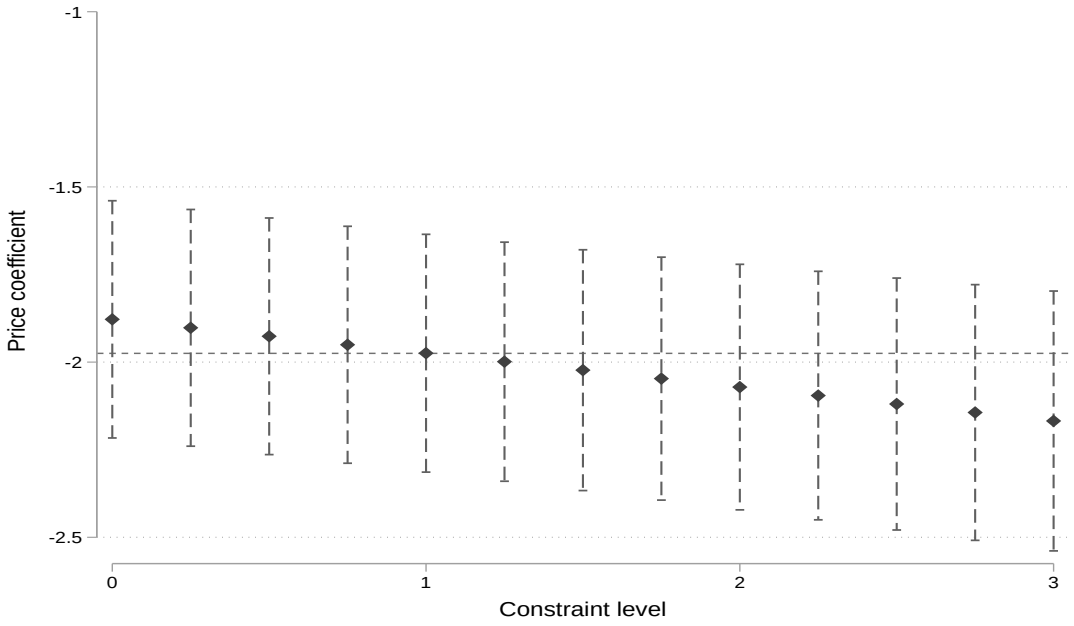


Figure A14: Sensitivity of price estimates to varying constraints of income and college share

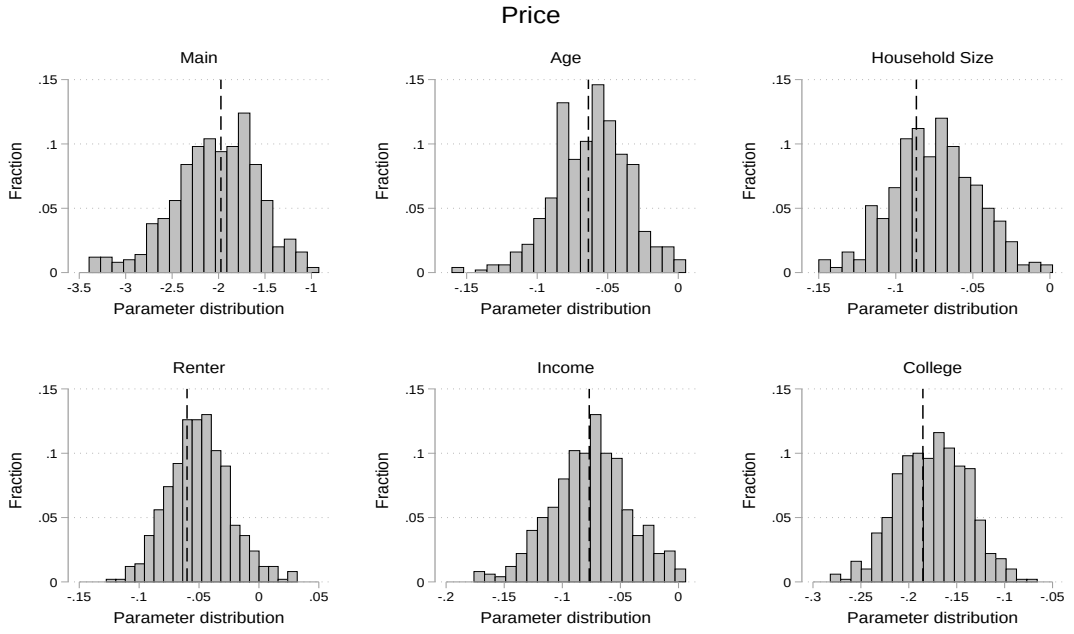


Figure A15: Bootstrap results: price

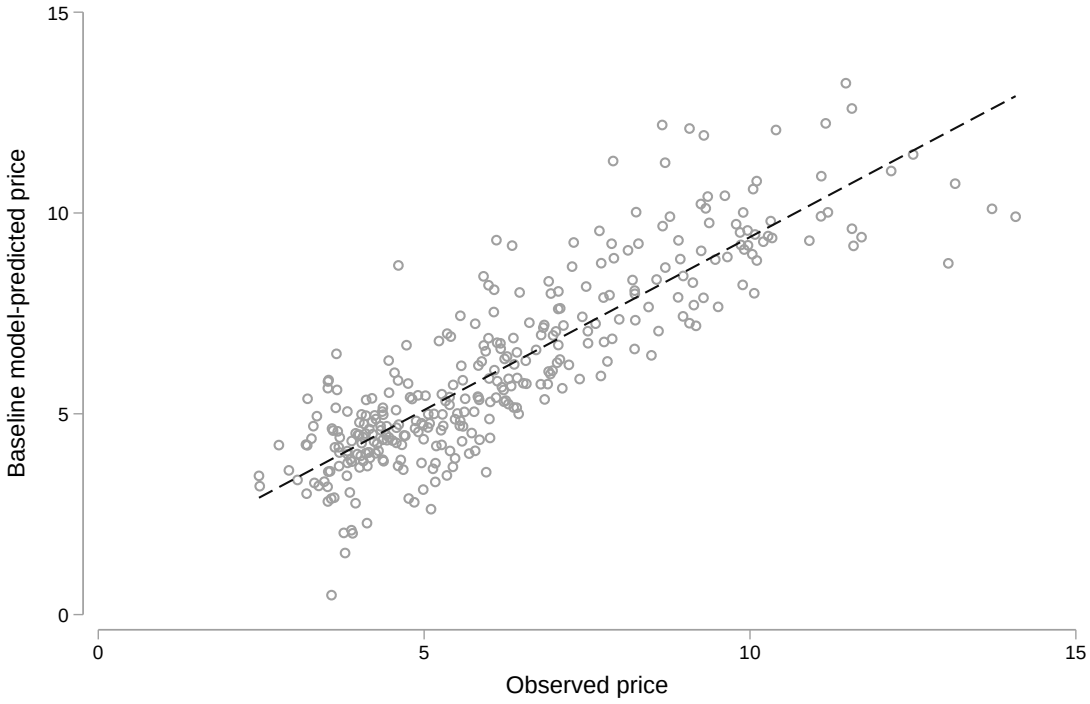


Figure A16: Model validation: prices

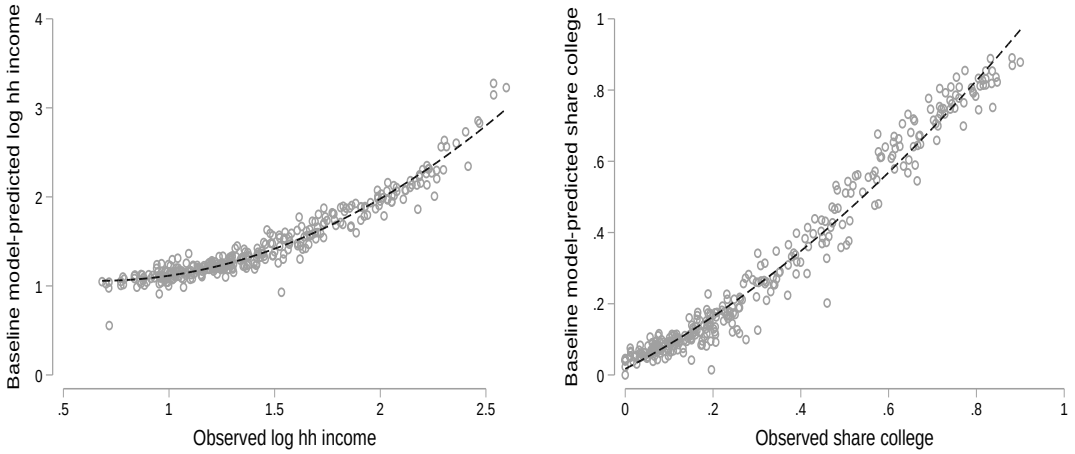


Figure A17: Model validation: demographics

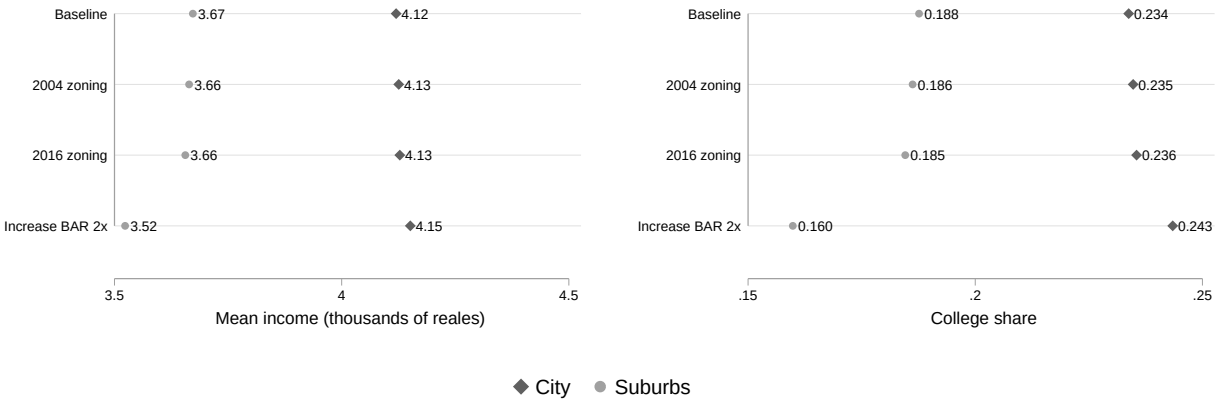


Figure A18: Demographic change in the suburbs and city under policy counterfactuals

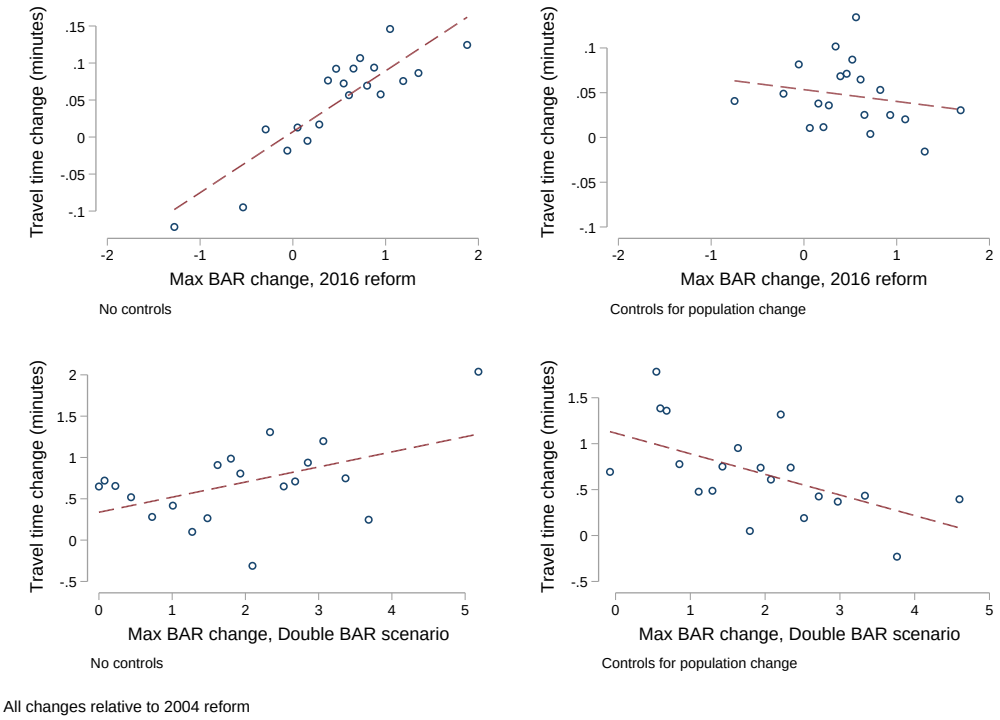


Figure A19: Travel time changes under policy counterfactuals

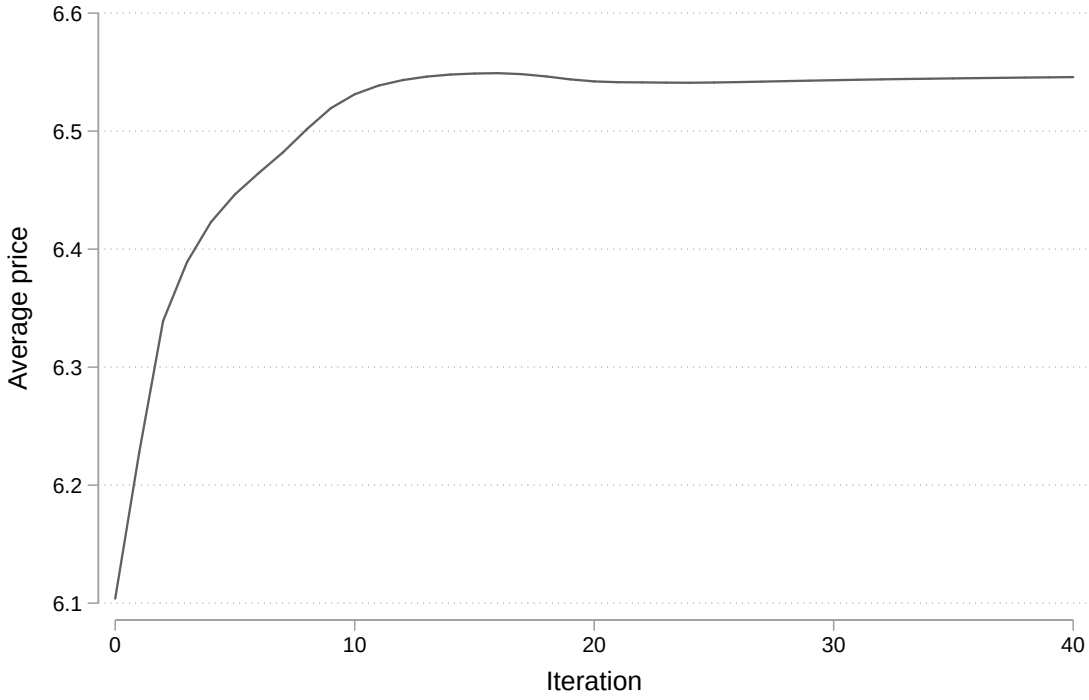
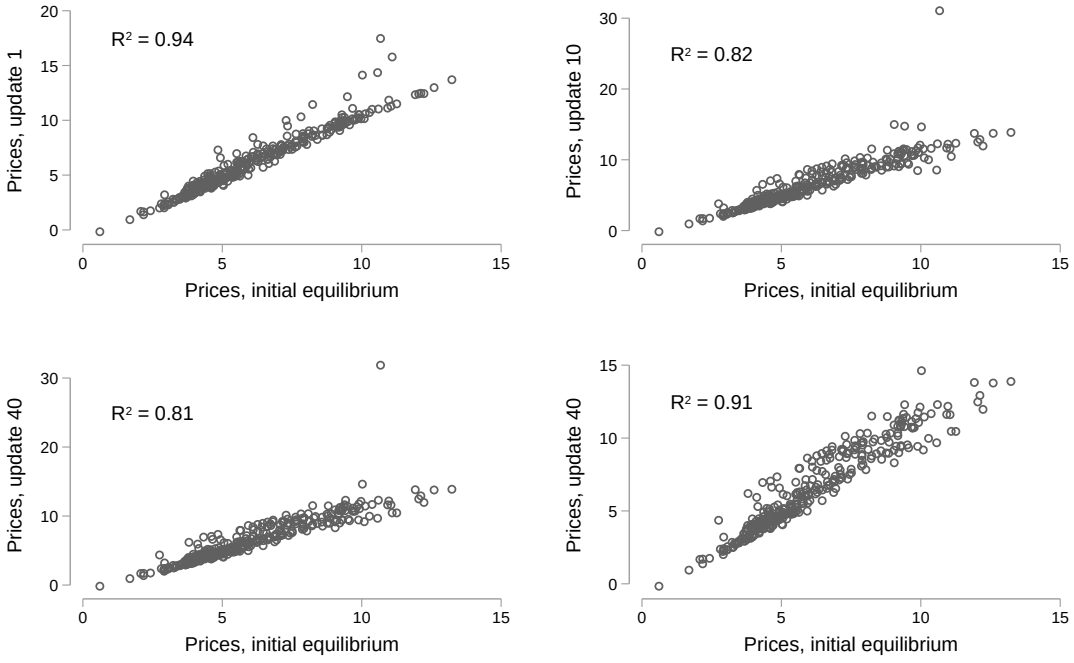


Figure A20: Prices by equilibrium iteration



Bottom right drops outlier

Figure A21: Price correlations by equilibrium iteration

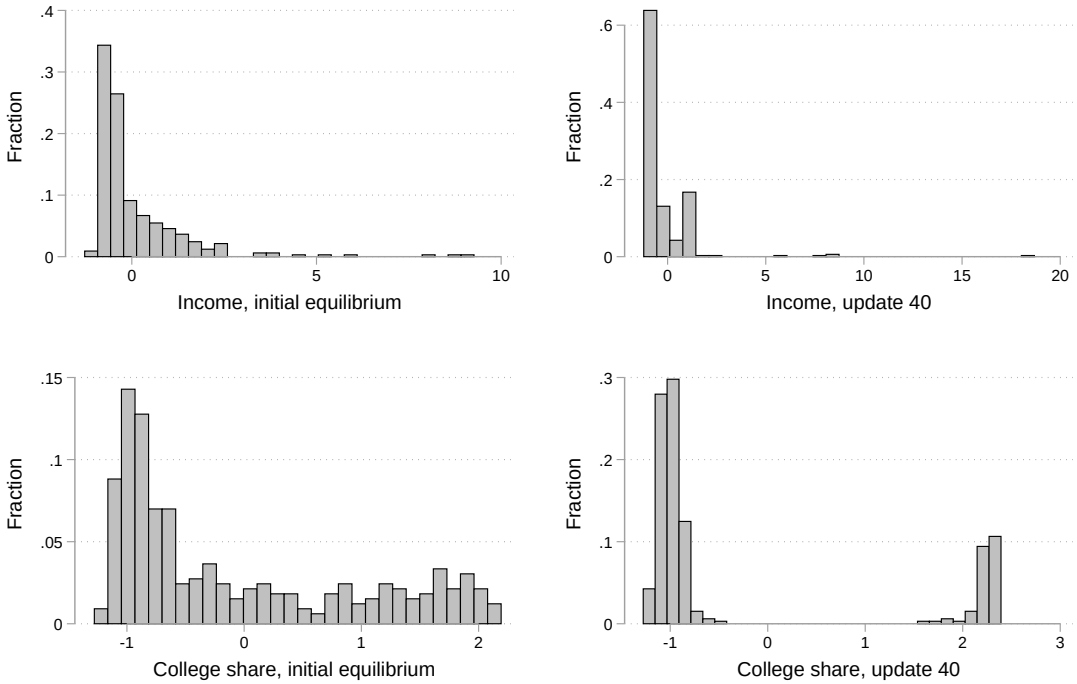


Figure A22: Demographic distributions by iteration