Distributional Effects of Residential Zoning Policies: Insights from the Greater Boston Area*

Dongyang He†

October 20, 2024
Link to the latest version

Abstract

This paper examines the impact of zoning policies on housing affordability and welfare inequality across income groups in the Greater Boston area. I focus on two specific regulations: Floor Area Ratio (FAR) restrictions and density regulations, both of which limit the supply of smaller, affordable housing units. Using a housing supply model, I show that these policies significantly reduce housing affordability, with the most affordable housing options at the tract level being, on average, five times more expensive under zoning constraints. To evaluate the welfare effects of these policies, I incorporate the housing supply model into a quantitative spatial equilibrium framework that captures both housing demand and supply across census tracts in the city. My results indicate that in the absence of zoning regulations, welfare for the lowest 10% income group could have been 41.6% higher, while welfare for higher-income groups would be mildly lower. Additionally, removing zoning regulations today could still increase welfare for the lowest 10% income renter group by 34.7%, but would negatively affect around 80% of current residents due to neighborhood demographic shifts. Property owners would also face declining property values, with an average decrease of 2.7% and significant variation across tracts. These findings suggest that while radical zoning reforms could enhance welfare for lower-income households and reduce inequality, they would also impose welfare losses on the majority of current residents. This research highlights the critical role of zoning policies in exacerbating housing affordability and the importance of considering distributional effects in zoning policy reforms.

*I am grateful to Stephen Yeaple, Fernando Parro, Jingting Fan, James Tybout, and Brent Ambrose for their tremendous guidance and support throughout this project. I also thank Weili Chen, Yinong Tan and Hyungjin Kim for their insightful comments. All errors are my own.

†Department of Economics, Pennsylvania State University. E-mail: hedongyang00@gmail.com

1 Introduction

The housing affordability crisis in the United States is a multifaceted issue, shaped by more than just rising home prices and stagnant incomes among lower-income groups. One contributing factor to this crisis is the distortion caused by various zoning regulations, such as restrictions on dwelling density and building Floor Area Ratio (FAR). Although these policies vary in their specifics, they generally constrain the development of middle or high-density, affordable dwelling unit, disproportionately affecting low-income households, who are more likely to rely on such dwelling options. [1]

In recent years, housing affordability has gained significant attention in national politics. Presidential candidate Kamala Harris, during her 2024 campaign, has emphasized the urgency of addressing this crisis, promising to add 3 million dwellings to the market as part of her plan to alleviate the shortage. Although Harris's proposals focused on other solutions, several states have already begun tackling the issue through zoning reform. In 2021, California passed SB 9, which permits duplexes in single-family zones, and Utah introduced reforms in the same year to relax zoning restrictions and encourage multi-family development. Oregon and Washington also implemented significant changes in 2019 and 2021, respectively, aimed at increasing dwelling density and reducing parking mandates. These recent efforts reflect a growing movement in certain states to address housing affordability by loosening restrictive zoning laws.

Despite recent policy changes, the extent to which zoning laws have worsened housing affordability and exacerbated welfare inequality is still not fully understood. This paper aims to bridge that gap by examining the distributional impact of zoning regulations in the Greater Boston Area, where such regulations are among the most restrictive in the United States. Specifically, I focuses on two key regulations that significantly shape the residential landscape in cities worldwide: Floor Area Ratio (FAR) and density restrictions. This paper seeks to answer four important questions: (i) to what extent have these restrictions influenced the cost of the most affordable dwelling options, referred to here as the *"minimum dwelling cost"*; (ii) to what extent these restrictions have distorted housing and location choices for lower-income households, exacerbating welfare inequality; (iii) which groups might benefit from having the zoning policies; (iv) what are the long-term effects if these zoning policies are removed? Central to this analysis is the widely recognized observation that zoning regulations limit the supply of middle- to high-density affordable housing, impacting lower-

[1] In this paper, I use the term "dwelling units" or "dwelling" to avoid the ambiguity associated with "housing." While "housing" can refer to an entire building or an individual unit, "dwelling units" specifically refer to individual units of accommodation within a building, such as an apartment, condominium, or single-family home. Further discussion on this definition is provided in Section 2.2.

income households that depend on these options.

To investigate these questions, I first outline zoning patterns in Greater Boston. The distortion caused by density restrictions is stark: around 20% of dwelling units built post-1950 cluster around the density cap, a pattern not observed for dwellings constructed before 1918, when zoning laws were nonexistent. Density restriction is stringent in the area: roughly 70% of the land in Greater Boston is zoned for densities below four dwelling units per acre, which effectively limits development to single-family homes. Another 10% is zoned for somewhat higher densities, but still insufficient to support middle- to high-density, affordable housing. FAR restrictions, in contrast, seem to have negligible impact on supply, indicating that density rules play a much more decisive role in shaping Greater Boston's housing landscape.

I develop a parsimonious dwelling production model that captures how zoning regulations constrain the supply of affordable housing units. In this model, dwelling units are differentiated by both quality and location, with smaller built FAR, lower built density, and higher costs associated with higher-quality units. While quality admits a wide range of relationships, these naturally emerge from the non-homotheticity I introduced into the dwelling production process. Specifically, the quality of a dwelling unit is derived from a Stone-Geary production function with floor space and low-density amenities as inputs, and a minimum floor space requirement. As the quality of a unit increases, more resources are diverted from producing floor space to enhancing low-density amenities, such as larger yards, establishing the proper correlations between quality, cost, built density, and built FAR. Even in the absence of zoning laws, the model generates a positive "minimum dwelling cost" due to the minimum floor space requirement. Using this model, both FAR and density restrictions can be represented as constraints on dwelling unit quality, enabling me to also calculate the "minimum dwelling cost" under these regulatory constraints.

I estimate the model parameters, treating each of the 530 census tracts in Greater Boston as distinct locations, comparing the "minimum dwelling cost" with and without zoning regulations. The results show zoning policies lead to a five fold increase in the "minimum dwelling cost" on average. While this figure may seem surprising, it aligns with the stringency of density restrictions in the study area. These zoning restrictions sharply limit the construction of affordable, middle- to high-density units like apartments, forcing a reliance on more expensive, low-density options such as single-family homes. In this context, the five-fold increase isn't surprising, as it reflects the price difference between the restricted middle- to high-density units and the more common, costlier low-density dwellings.

This increase in the "minimum dwelling cost" could be burdensome for lower-income households. To fully evaluate the welfare effects, I embedded the dwelling production model into a spatial general equilibrium model, matching the observed housing distribution across census tracts. This model also includes features specific to Boston's housing market—factors like zoning variances, grandfathered houses, special permits, and homeownership patterns. The key ingredients are households with heterogeneous earning potentials and property ownerships making endogenous choices of housing, residential and working locations based on income, cost-of-living, commute cost and residential amenities. Households are required to pay for one dwelling unit to reside in a location, but can choose a dwelling quality that maximize their utilities. Higher "minimum dwelling cost" mainly effects lower income households by either pricing them out of certain locations or forcing them to choose higher-quality units than they otherwise would. The magnitude of this impact depends on how sensitive lower-income households are to changes in the "minimum dwelling cost."

I estimate the parameters, and recover the unobservables to rationalize the observed data as an equilibrium outcome of the model. Then I use the calibrated model to derive the equilibrium outcome in a scenario where zoning policies were never introduced, and compare it with the realized outcome to understand the effects of having zoning policies. My model suggests that, in the absence of zoning constraints, the welfare of lower-income households would be substantially greater, with the welfare of the lowest 10% income group being 41.6% higher compared to a scenario with zoning regulations. In contrast, the welfare of the highest 80% income groups would be slightly lower, ranging from 1.2% to 2.7% less. This substantial welfare difference for the lowest 10% is driven by the extreme sensitivity of low-income households to the availability of more affordable housing, as they have up to 15 times less potential income than the upper 20%, according to my model estimates. For the higher 80%, the differences in welfare is primarily due to the reduced residential amenities that result from a greater presence of lower-income households in their neighborhoods.

After analyzing the effects of existing zoning policies, I explore the long-term impacts of radical zoning reforms that completely remove existing FAR and density regulations. Unlike the previous analysis, this assessment must account for the path dependence of the current equilibrium. To address this, I incorporate redevelopment costs and other necessary adjustments. I use the calibrated model to derive the long-term outcome after zoning policies are removed from the current equilibrium to evaluate the effects of zoning reforms. My model indicates that removing zoning constraints would greatly benefit lower-income renter households, with the welfare of the lowest 10% increasing by 34.7%. In contrast, the welfare of the highest 80% income groups would see a slight decline, with reductions ranging from 1.4% to 2.6%. Moreover, property owners across all income levels would experience negative impacts due to decreased property values.

Opponents of zoning reform often raise two key concerns. The first is the fear that eliminating zoning regulations would reduce property values for current homeowners. This concern is based on the expectation that reforming zoning laws would increase the overall housing supply, making housing less scarce and thus driving down property prices. My model supports this concern, estimating that, on average, property values in the Greater Boston Area would decrease by 2.7% following the removal of zoning policies. However, the change in property values varies significantly across tracts. In 60% of the tracts, property values would increase by 2% to 5%, while in 30% of the tracts, property values would decrease, with 20% of those experiencing declines of more than 5%, and 10% seeing drops greater than 10%.

The second major concern relates to potential demographic shifts. Opponents argue that relaxing zoning regulations could lead to an influx of lower-income households into previously higher-income, homogeneous neighborhoods, raising fears of increased crime or social disruption. This concern is also reflected in my model. In addition to lowering the "minimum dwelling cost," which would allow lower-income households to afford housing in previously exclusive neighborhoods, the removal of zoning policies would also make the Greater Boston Area more attractive to lower-income households from other parts of the U.S., resulting in an influx that lead to a 157.3% increase in the population of the lowest 10% income group.

These findings indicate that while radical zoning reforms could improve welfare for lower-income households and reduce inequality, they would also result in losses for the majority of current residents. This hightlights the need to carefully consider the distributional impacts and take a more balanced approach when implementing zoning policy reforms.

Literature Review. The paper contributes to a large literature on the effects of land-use regulations on housing markets and residential construction (e.g., Glaeser et al. (2005); Glaeser and Ward (2009); Herkenhoff et al. (2018)). Hindered by limited understanding of new housing production, particularly how housing differentiation arises, earlier work faces several limitations. First, most examine the effects of a single regulation—typically either Floor Area Ratio (FAR) or density restrictions—without considering how these regulations interact. Second, they often fail to account for housing heterogeneity, focusing on the impact of regulations on the supply of "average housing" and overlooking differentiated effects across various housing types. Third, many studies neglect spatial equilibrium effects, missing how regulatory changes influence broader urban and welfare outcomes.

Recent research has aimed to address these gaps. For instance, Kulka et al. (2023) utilize spatial regression discontinuity designs to explore interactions between FAR and density regulations in the Greater Boston Area. They find that stricter regulations increase the size of the smallest housing units, raising the "minimum dwelling cost" and reducing affordability, findings that align with my model's results. However, their analysis does not consider the spatial equilibrium effects of zoning reforms or their implications for different demographic groups. Other recent studies, such as Ospital (2023) and Tokman (2023), employ quantitative urban models to capture both housing heterogeneity and equilibrium effects. Yet, their housing heterogeneity arises from exogenous differences in housing locations and parcel sizes, limiting their ability to assess the impact of FAR regulations or examine radical zoning reforms, such as the removal of density restrictions. My contribution to this literature is the development of a parsimonious housing supply model for heterogeneous housing that fully endogenizes key features such as built FAR and built density. This model can account for the interactive effects of FAR and density regulations and can be integrated into a quantitative urban framework to capture equilibrium effects.

This paper also contributes to the quantitative spatial economics literature, which leverages general equilibrium spatial models to evaluate the effects of economic shocks and policies (e.g., Ahlfeldt et al. (2015); Redding and Rossi-Hansberg (2017); Tsivanidis (2022); Heblich et al. (2020); Rossi-Hansberg et al. (2021); Pang (2021)). I estimate my model using the model inversion methodology developed in these studies. Additionally, I enrich the quantitative spatial framework by incorporating both housing and worker heterogeneity. This enhancement enables a more detailed analysis of how zoning regulations influence not only housing supply but also the distribution of welfare across different economic groups.

Finally, this paper contributes to the literature examining how housing supply affects neighborhood-level income sorting and segregation (e.g., Bayer et al. (2007); Guerrieri et al. (2013); Kuminoff et al. (2013); Couture et al. (2024)). I evaluate how zoning laws, by increasing the "minimum dwelling cost," create residential barriers that shape neighborhood demographics and influence the distribution of welfare across income groups.

The remainder of the paper is organized as follows: Section 2 discusses the data and zoning patterns. Section 3 introduces the dwelling production model and analyzes its implications. Section 4 presents the quantitative spatial model. Section 5 calibrates the model. Section 6 conducts the welfare analysis, and Section 7 concludes.

2 Zoning and Housing Supply in Greater Boston Area

2.1 Data

This subsection outlines the datasets used in the analysis of housing supply.

The primary source of zoning regulation information is parcel-level data from digitized maps provided by the Metropolitan Area Planning Council (MAPC) through their Zoning Atlas project. This dataset covers 101 towns in Greater Boston, offering a detailed view of zoning regulations as they existed in 2010. It includes comprehensive zoning codes and bylaws for the region.

Data on housing units and their characteristics are derived from town tax assessment records compiled by CoreLogic, spanning the period from 2006 to 2021. This dataset includes nearly all residential and mixed-use buildings within Greater Boston. Key variables include building type (e.g., single-family or multi-family), number of units per building, parcel size, building area, year of construction, tax-assessed value, sale price and date, geographic coordinates, and various building features, such as the number of rooms and bathrooms.

Land price data comes from AEI-adjusted land price and land share indicators, based on the methodology developed by Davis et al. (2021). This approach estimates land values using the "cost approach," calculating land value as the difference between the appraised property value and the estimated depreciated replacement cost of the building structure. The dataset, consisting of millions of appraisal records, provides a balanced annual panel of land prices at the census tract level for the years 2012 to 2019.

2.2 Dwelling Unit Definition and Characteristics

In this paper, I use the term "dwelling units" or "dwelling" to avoid the ambiguity associated with the term "housing." While "housing" can refer to either an entire building or an individual unit, "dwelling" or "dwelling units" specifically refer to individual units of accommodation within a building, such as an apartment, condominium, or single-family home.

All of the housing analysis conducted in this paper is performed at the dwelling unit level. Although different dwelling units have varying specifications, I focus primarily on two well-defined characteristics that are present across all dwelling units regardless of their specifics: floor space and land input.

For dwellings in single-family structures, floor space refers to the built area, and land input refers to the lot size of the structure. For dwellings in multi-family structures, these are defined as the built area and the lot size of the structure divided by the number of dwelling units it contains. These two characteristics allow for the calculation of the built FAR and built density for each dwelling unit.

2.3 Zoning Regulation Framework

This section focuses on two major zoning regulations that influence the development of residential housing units: Floor Area Ratio (FAR) restrictions and Maximum Dwelling Units Per Acre (DUPAC), also known as density restrictions.

Maximum Dwelling Units Per Acre (DUPAC): DUPAC regulations limit residential density by defining the maximum number of dwelling units that can be constructed per acre of land. DUPAC is calculated based on the number of lots allowed on one acre according to minimum lot size requirements and is further adjusted by the maximum allowable dwelling units per lot. This measure accounts for both the zoning restrictions imposed by minimum lot size and the maximum dwelling units permitted per lot, enabling comparisons across municipalities that regulate residential density in varying ways. Figure 1 illustrates how DUPAC restrictions vary spatially. Approximately 70% of the land in Greater Boston is zoned for densities below four dwelling units per acre, effectively restricting development to detached single-family homes. An additional 10% is zoned for densities of around ten units per acre, which remains insufficient to support even middle-density building, such as three-story apartment buildings.

FAR Restrictions: FAR restrictions regulate the maximum allowable floor area that can be constructed on a parcel relative to the size of the parcel itself. Figure 2 shows the variation in FAR restrictions across Greater Boston. FAR limits are often set to ensure that buildings remain within a certain scale, with lower FARs constraining the size of building developments. Compared to DUPAC, FAR restrictions tend to be more relaxed. Only around 10% of the land has FAR restrictions below 0.5, a level typical for single-family homes, while over half of the land has FAR restrictions above one, which allows for the construction of middle- and high-density developments, including multi-story apartment buildings that exceed three floors.

2.3.1 Interaction of FAR and Density Restrictions

While Floor Area Ratio (FAR) and dwelling units per acre (DUPAC) restrictions are relatively straightforward when considered individually, their combined impact on development is more complex. Depending on the situation, either regulation can act as the binding constraint, limiting both the scale and density of housing projects. Relaxing one constraint without adjusting the other may not lead to a significant increase in housing supply, as the remaining, unchanged regulation can become the new limiting factor. The interaction between these two regulations can have a substantial influence on development outcomes, as illustrated in the following example.

Consider a municipality where the DUPAC limit is set at 5 units per acre and the FAR restriction is 0.5. Under these regulations, developers can construct up to 5 housing units per acre, with the total floor area of all buildings on a parcel limited to 50% of the parcel's land area. In practice, for low-density housing such as detached single-family homes, the built FAR typically remains below 0.5. In this scenario, DUPAC becomes the binding constraint,

Figure 1: Maximum density (DUPAC) restrictions

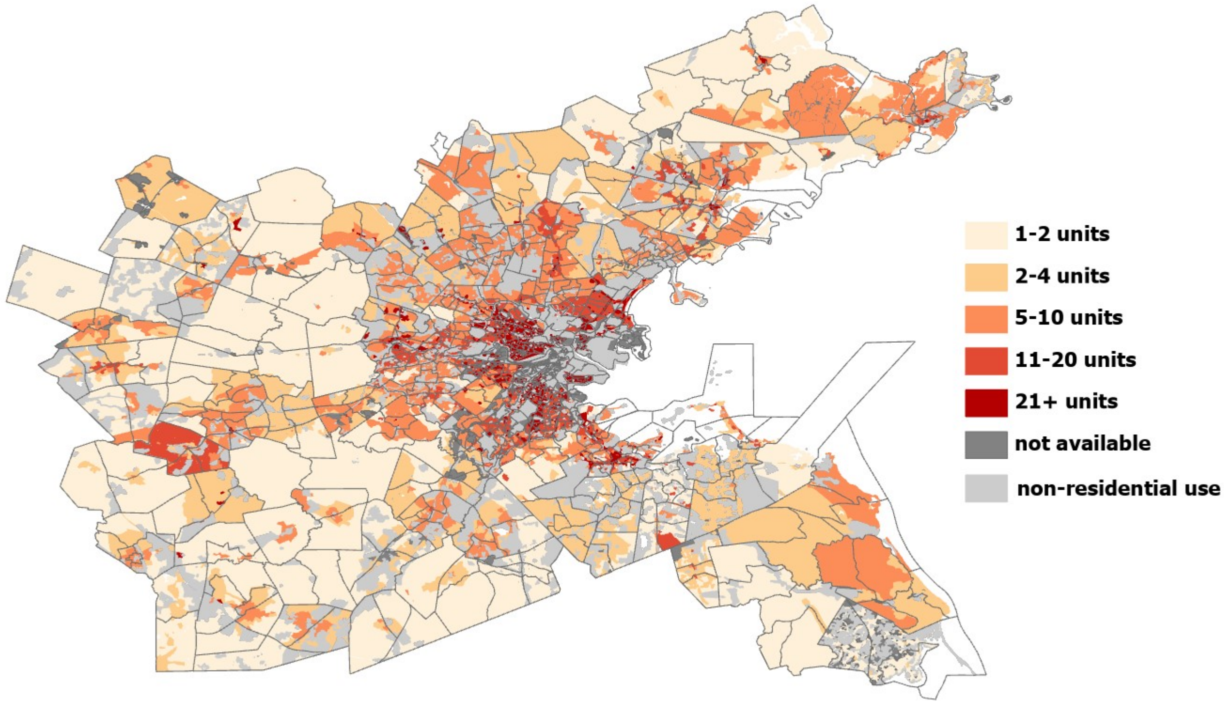
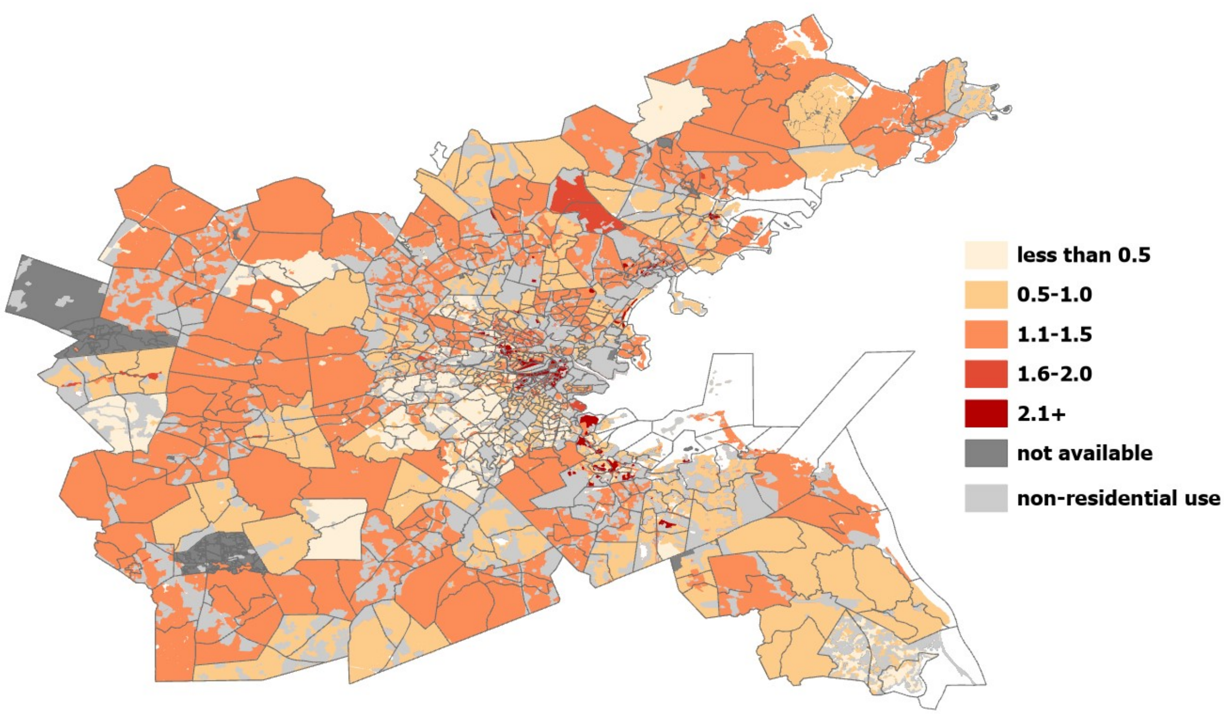


Figure 2: Floor Area Ratio restrictions



as the 5 units per acre cap effectively restricts housing types to detached single-family homes with yard space.

Now, suppose the municipality increases the DUPAC limit to 40 units per acre while leaving the FAR restriction unchanged. Developers might seek to take advantage of the higher density allowance by planning to build more units. However, the unchanged FAR of 0.5 still limits the total floor area that can be developed. As a result, developers may only be able to construct, for example, 15 units per acre within the permitted floor area, far short of the new DUPAC limit. In this case, the FAR restriction becomes the binding constraint, preventing full utilization of the increased density allowance.

This example highlights that when both FAR and DUPAC restrictions are in place, adjusting only one may not effectively increase housing supply. The unaltered regulation can become the new limiting factor, constraining development potential. To promote meaningful growth in housing units, policymakers may need to revise both FAR and density restrictions simultaneously to remove development bottlenecks.

2.3.2 History of Zoning Development in Boston

Boston's zoning regulations have evolved significantly since their initial adoption. The city first established broad zoning categories in 1918, dividing land into residential, industrial, and commercial zones while also introducing maximum height restrictions. This move followed New York's pioneering zoning regulations, implemented just two years earlier in 1916. Neighboring municipalities, such as Cambridge in 1920, and towns like Brockton, Brookline, and Newton in the early 1920s, soon adopted similar zoning frameworks (Bobrowski (2002); MacArthur (2019)).

By the 1950s, however, these broad categories and height limits were proving inadequate for managing the housing potential of parcels. In response, Boston and other municipalities within Greater Boston revised their zoning regulations in 1956, introducing density controls such as DUPAC (dwelling units per acre) regulations. These new measures limited the amount of habitable floor area relative to parcel size, marking a crucial shift in Boston's approach to urban planning and zoning (Bobrowski (2002); MacArthur (2019)).

2.4 Zoning and Distortion in Dwelling Distribution

In Section 2.3, we analyzed the spatial distribution of DUPAC and FAR restrictions, noting that DUPAC is particularly stringent. Only about 20% of the land is zoned for densities above 10 units per acre, a threshold typically supporting middle- and high-density developments.

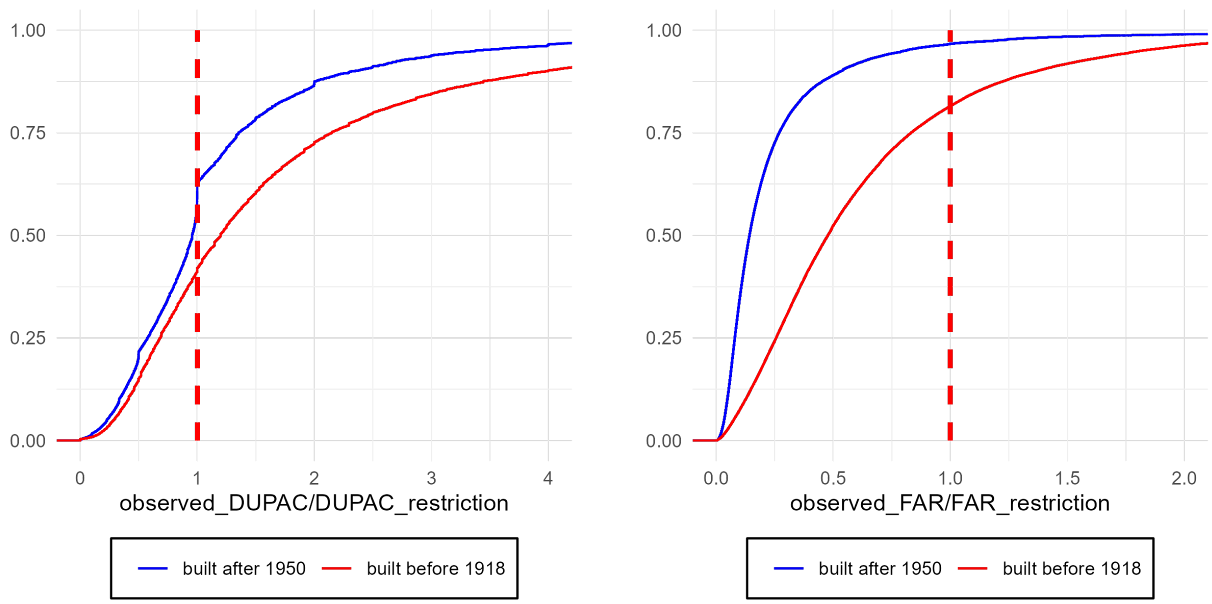
To evaluate how these restrictions influence housing supply, I examined distortions in the distribution of dwelling units. Figure 3a illustrates the ratio of observed DUPAC to regulatory DUPAC for each unit, comparing two categories of buildings: those constructed after 1950, when DUPAC regulations were implemented, and those built before 1918, prior to any zoning regulations. This analysis reveals two key findings.

First, approximately 20% of post-1950 units cluster below a ratio of 1, a pattern absent in pre-1918 buildings, which are categorized as "grandfathered nonconforming uses." This suggests that DUPAC has constrained the density of dwellings constructed after 1950, resulting in lower densities than otherwise would occur in the equilibrium.

Second, around 30% of post-1950 units have a ratio exceeding 1, indicating non-compliance with base zoning regulations. While this may seem surprising, there are two plausible explanations. First, some parcels may have been affected by updated DUPAC regulations introduced after the 1950s, where allowable densities were reduced or minimum lot sizes were increased, but the pre-existing properties remain as "grandfathered nonconforming uses." Second, some developments may have been permitted under alternative regulatory frameworks, such as overlay districts, zoning variances (which are common in the City of Boston), or through Comprehensive Permits.

These findings highlight the importance of grandfathered nonconforming uses and non-compliance with base zoning regulations in the Boston housing market. In Section 4, when I develop a spatial general equilibrium model to evaluate welfare effects, the model will incorporate these features to better reflect the market realities.

Next, I performed a similar analysis by examining the ratio of observed FAR to regulatory FAR for each unit. The results, shown in Figure 3b, reveal no significant distortion around a ratio of 1. Moreover, for units built after 1950, over 95% exhibit observed FAR values below the regulatory level. This aligns with the discussion in Section 2.3, where FAR appears more lenient in comparison to DUPAC. In this case, DUPAC appears to dominate as the binding constraint on density, with FAR rarely serving as a limiting factor.



(a) DUPAC

(b) FAR

Figure 3: Cumulative Distribution for DUPAC and FAR Ratio

This graph plots the Cumulative Density Distribution for the ratio between observed DUPAC and FAR with the respective restrictions. Each data point represents a dwelling unit.

3 Dwelling Production Model

In this section, I present a parsimonious model for the production of heterogeneous dwelling units, linking dwelling characteristics such as built floor-area ratio (FAR) and built density to the quality of the dwelling unit.

The production of a dwelling unit with quality level h requires a combination of floor space f and low-density amenities a. The production function is given by:

$$h = a^{\beta}(f - \underline{f})^{1-\beta}, \tag{1}$$

where $\underline{f}$ represents the minimum floor space required for any dwelling unit.

Low-density amenities a are produced directly from land input in a one-to-one ratio. Floor space f, on the other hand, is generated by combining land input g with construction goods input c, according to the following production technology:

$$f = Bc^{\gamma}g^{1-\gamma}, \tag{2}$$

where B denotes housing productivity.

The total land input k employed in the production process is the sum of low-density amenities and land used for floor space production, i.e., $k = a + g$. Here, low-density amenities a can be interpreted as elements such as a yard or preferences for lower density living.

As the quality of a dwelling unit increases, a larger fraction of the total resources is allocated to the production of low-density amenities, shifting resources away from floor space. This shift, induced by the Stone-Geary production function in (1), leads to lower built FAR and built density, which are characterized by the ratios f/k and $1/k$, respectively.

3.1 Unconstrained Dwelling Production

In this section, we examine the construction costs of dwelling units in the absence of constraints on dwelling attributes.

From equations (1) and (2), we can express low-density amenities $a(h, f) = \left(\frac{h}{(f-\underline{f})^{1-\beta}}\right)^{\frac{1}{\beta}}$ and construction goods $c(h, f, k) = \left(\frac{f}{B(k-a(h,f))^{1-\gamma}}\right)^{\frac{1}{\gamma}}$. For each quality level h, builders choose attributes f (floor space) and k (total land input) to minimize the total construction cost $p_h^{H,con}$. This minimization problem is given by:

$$p_h^{H,con} = \min_{(f,k)} \left(p^c c(h, f, k) + p^L k\right), \tag{3}$$

where p^c denotes the price of construction goods, and p^L denotes the price of land. The term $p^c c(h, f, k)$ represents the structure's construction cost, while $p^L k$ represents the cost of land.

In this model, construction goods serve as the numeraire of the economy, with their price normalized to $p^c = 1$.

By solving the cost-minimization problem, we can derive the optimal choices of floor space f, total land input k, and the resulting production cost. The detailed solution process is provided in Appendix A.2.1. The optimal choices are given by:

$$f = \frac{(1-\beta)h}{z_\beta} \left(Bz_\gamma (p^L)^\gamma\right)^\beta + \underline{f}, \tag{4}$$

$$k = \frac{1}{Bz_\gamma (p^L)^\gamma} \left(\left(Bz_\gamma (p^L)^\gamma\right)^\beta \frac{h}{z_\beta} (1 - \gamma(1-\beta)) + \underline{f}(1-\gamma) \right), \tag{5}$$

and the corresponding optimal production cost:

$$p_h^{H,con} = \frac{(p^L)^{1-\gamma}}{Bz_\gamma} \left(\left(Bz_\gamma (p^L)^\gamma\right)^\beta \frac{h}{z_\beta} + \underline{f} \right), \tag{6}$$

where $z_\beta = \beta^\beta (1-\beta)^{1-\beta}$ and $z_\gamma = \gamma^\gamma (1-\gamma)^{1-\gamma}$.

3.1.1 Properties of the Dwelling Production Model

This model of dwelling production aligns with three well-established observations in the housing market. Firstly, it reflects that land inputs k increase with the size f of the dwelling unit. Secondly, it reflects that the Floor Area Ratio (f/k) declines as the size f of a dwelling unit increases. Thirdly, it captures that the land cost share $(kp^L)/p_h^{H,con}$ increases with the size f of the dwelling unit.

We now use equations to characterize these model implications. We begin by expressing h as a function of k, through the inversion of equation (5). Then, we incorporate this expression into equations (4) and (6) to arrive at the following equations:

$$f = \frac{\underline{f}\beta}{1-\gamma(1-\beta)} + \frac{Bz_\gamma (p^L)^\gamma (1-\beta) k}{1-\gamma(1-\beta)} \tag{7}$$

$$p^{H,con} = \frac{1}{1-\gamma(1-\beta)} k p^L + \frac{\gamma \underline{f} \beta (Bz_\gamma (p^L)^\gamma)^{-1}}{1-\gamma(1-\beta)} p^L \tag{8}$$

To obtain built FAR, divide both sides of (7) by k:

$$\frac{f}{k} = \frac{\underline{f}\beta}{(1-\gamma(1-\beta))k} + \frac{Bz_\gamma (p^L)^\gamma (1-\beta)}{1-\gamma(1-\beta)} \tag{9}$$

Here, the second term represents the baseline built FAR, and the first term indicates the portion of built FAR that varies with the land input k. More specifically, $\frac{1}{k}$ is the number of dwelling units a unit of land can accommodate if each use k land inputs (namely, dwelling density). So $\frac{\underline{f}\beta}{1-\gamma(1-\beta)}$ can be interpreted as the amount of additional floor space will be present on this unit of land if we increase the dwelling unit by one.

For the land cost share, divide both sides of (8) by kp^L and invert the equation:

$$\frac{kp^L}{p^{H,con}} = \left(\frac{1}{1-\gamma(1-\beta)} + \frac{\gamma \underline{f}\beta (Bz_\gamma (p^L)^\gamma)^{-1}}{(1-\gamma(1-\beta))k} \right)^{-1} \tag{10}$$

The minimum level of land input k can be derived by setting $h = 0$ in (5). In that case, the land cost share is $1-\gamma$. As k increases indefinitely, the land cost share asymptotically approaches $1-\gamma(1-\beta)$. Thus, the land cost share spans from $1-\gamma$ to $1-\gamma(1-\beta)$.

3.2 Zoning Regulations

3.2.1 Constrained Dwelling Production and Dwelling Attribute Distortions

Zoning regulations impose restrictions on the attributes of dwelling units. Specifically, density regulations can be represented by minimum land input requirements (i.e., $\underline{k} < k$), and FAR regulations can be represented by $\bar{r} > \frac{f}{k}$.

As discussed in Section 3.1.1, dwelling units of lower quality require less land and have higher built FARs, making them more likely to be constrained by zoning regulations. It is possible that dwelling units with a quality level near the regulatory cutoff might distort their attributes to comply with zoning policies. In such cases, the units may either have larger land inputs or they may have smaller built FARs than otherwise necessary.

If this is the case, the builders' problem can be represented as

$$p_h^{H,con} = \min_{(f,k)\in Z} c(h,f,k) + p^L k,$$

where Z represents all the possible (f,k) combinations that satisfy the zoning policy.

However, upon analyzing dwelling units close to the regulatory cutoff, I do not find significant distortions in dwelling attributes, indicating that zoning regulations do not significantly distort the attributes of dwelling units. Detailed analysis is included in the appendix A.3.

3.2.2 Zoning Policies As Constraints on Dwelling Quality

Figure 3 suggests that dwelling supply is distorted by zoning policies. If dwelling attributes are not distorted, then dwelling quality should be distorted. In other words, to comply with zoning regulations, dwelling units are built with qualities higher than would otherwise be preferred. Consequently, zoning regulations effectively restrict the range of dwelling qualities available on the market.

I now formulate zoning policies as constraints on the choices of dwelling qualities:

According to Equation (9), a FAR regulation imposing the constraint $\bar{r} \geq \frac{f}{k}$ can be expressed in the form of a minimum land input requirement:

$$k \geq \frac{\underline{f}\beta}{(1-\gamma(1-\beta))(\bar{r} - \frac{Bz_\gamma(p^L)^\gamma(1-\beta)}{1-\gamma(1-\beta)})}$$

From Equation (5), a density regulation that specifies a minimum land input $\underline{k} \leq k$ can be reformulated as a constraint on dwelling quality:

$$h \geq \frac{z_\beta(\underline{k}Bz_\gamma(p^L)^\gamma - \underline{f}(1-\gamma))}{(Bz_\gamma(p^L)^\gamma)^\beta\,(1-\gamma(1-\beta))} \tag{11}$$

To accommodate noncompliance due to zoning variance and special permit, I posit that deviations from zoning rules incur a nonconforming cost v. The final cost of dwelling is thus

$$p_h^H = p_h^{H,con} + v\mathbf{1}_{\{h \notin Z\}},$$

where Z denotes the set of dwelling qualities h that comply with the zoning policies. I assume the housing market is competitive, so this final cost also represents the market price of the dwelling unit.

3.2.3 Zoning Policies and Property Values

Given the assumption that housing market is competitive, it may not be immediately clear how zoning policies influence property values, aside from the addition of the nonconforming cost v for certain units. From equation (3), we know that the construction cost $p_h^{H,con}$ can be decomposed into the structure cost and land cost.

In this model, zoning policies primarily impact property values by altering land demand and land prices at equilibrium. Property owners, therefore, either gain or lose depending on whether land prices increase or decrease as a result of these zoning regulations.

3.3 Model Estimation and Model Fit

I utilize property-level housing data from the Greater Boston Area, sourced from CoreLogics, to estimate dwelling production parameters B_I, γ, $\underline{f}$, and β. It is assumed that the land price, denoted as p_I^L, and land productivity, denoted as B_I, vary only across different census tracts, indexed by I.

Following (9), the built FAR for each dwelling unit ω, located in census tract $I(\omega)$, is characterized by the equation:

$$\frac{f(\omega)}{k(\omega)} = \frac{B_{I(\omega)} z_\gamma (p^L_{I(\omega)})^\gamma (1-\beta)}{1-\gamma(1-\beta)} + \frac{\underline{f}\beta}{1-\gamma(1-\beta)} \frac{1}{k(\omega)} + \epsilon_\omega^{FAR} \tag{12}$$

By regressing built FAR $f(\omega)/k(\omega)$ against dwelling density $1/k(\omega)$ and incorporating tract-level fixed effects, we estimate $\frac{\underline{f}\beta}{1-\gamma(1-\beta)}$ and retrieve B_I up to a scalar. Table 1 presents the regression results. From column 1, R^2 of a simple regression is already 0.89, indicating built FAR is roughly linear in dwelling density. My preferred specification, as presented in column 2, has a R^2 over 0.93.

From (7) and (10), we have

$$\frac{kp^L}{p^{H,con}} = \left(\frac{1}{1-\gamma(1-\beta)} + \frac{\underline{f}\beta\gamma(1-\beta)}{(1-\gamma(1-\beta))^2 (f - \frac{\underline{f}\beta}{1-\gamma(1-\beta)})} \right)^{-1} \tag{13}$$

Setting $\frac{\underline{f}\beta}{1-\gamma(1-\beta)}$ at the estimated level x_1, we characterize the land cost share with:

$$\frac{kp^L_{I(\omega)}}{p^{H,con}(\omega)} = \left(\frac{1}{1-\gamma(1-\beta)} + \frac{x_1\gamma(1-\beta)}{1-\gamma(1-\beta)} \frac{1}{(f(\omega)-x_1)} \right)^{-1} + \epsilon_\omega^{LCS} \tag{14}$$

From the data, we observe only the final price, $p^H(\omega)$, not the construction cost, $p^{H,con}(\omega)$. Therefore, we focus on a subset of observations that are either not constrained by zoning or have a presumably low nonconforming cost, allowing us to treat the final price as being roughly equivalent to the construction cost under the assumption of perfect competition. Specifically, I assume that census tracts where more than 80% of post-1950 dwelling fails to comply with zoning regulations likely have minimal nonconforming costs.

I employ a nonlinear regression model to estimate the parameter $\gamma(1-\beta)$, denoting the estimated value as x_2. [2] Figure 4 illustrates the estimated land cost share function derived

[2]The type of dwelling unit—such as single-family homes, duplexes, or larger multifamily buildings—is a significant factor affecting cost share. To ensure my analysis of the subset aligns with the overall dataset patterns, I use a stratified weighting approach. Specifically, I categorize dwelling in the subset and the whole sample into type-floorspace groups. The "type" includes single-family homes and multi-family homes, while

f/k	(1)	(2)
$1/k$	931.567***	827.204***
	(0.559)	(0.787)
Constant	2800.640***	
	(20.479)	
Observations	327688	327094
Full Model R^2	0.8946	0.9378
Fixed Effects		tract

Table 1: Built FAR and dwelling density

The sample limited to post-1950 dwelling units. (Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)

from equation (14), alongside a nonparametric regression curve. This comparison highlights the model's effectiveness in capturing the data pattern for smaller dwelling units. For larger units, the nonparametric curve indicates that land cost share decreases as floor space increases. This trend can be attributed primarily to differences in the composition of dwelling types. For instance, larger single-family homes often include more luxury features compared to their smaller counterparts, resulting in a lower proportion of land cost. Conversely, large multifamily buildings tend to contain more densely packed dwelling units and more stories compared to smaller ones, also leading to a reduced land cost share. Consequently, as the housing mix shifts from large multifamily buildings to smaller multifamily and single-family houses, and then to larger single-family homes, the land cost share forms an inverse U-shaped pattern. My model specification only allows the land cost share to increase monotonically, thus fail to capture the patterns for larger units. However, for the purpose of welfare analysis, matching larger unit is less important. I will discuss this in the full model section.

Using estimated values for x_1 and x_2 alone does not provide enough information to separately identify the parameters β, γ, and $\underline{f}$. To address this, I use the data patterns to determine a range of appropriate parameter sets while keeping x_1 and x_2 fixed. The model land cost share curve depicted in Figure 4 also visually represents the potential sets of γ and $\underline{f}$ given x_1 and x_2. By selecting a point on this curve as the minimum land cost share $1 - \gamma$, the corresponding floor space value becomes the minimum floor space requirement $\underline{f}$. This choice of minimum land cost share reflects the permissible building types in the model. Data indicates that single-family houses typically have a land cost share of around 0.6, small multifamily buildings around 0.5, and large multifamily buildings approximately 0.2. These

"floorspace" is grouped into bins of 100 square feet intervals. I weight the observations in each group by the ratio of the whole sample count to the subset count in that group. This stratified weighting ensures that the analysis reflects the proportional representation of each group in the overall dataset, minimizing bias from over- or under-representation of specific dwelling types or floorspace ranges.

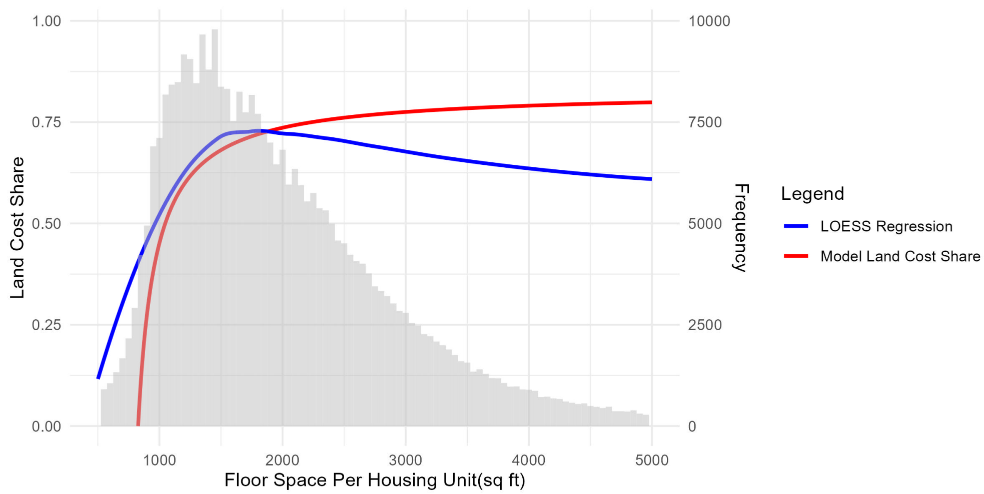


Figure 4: Land cost share: model v.s. nonparametric regression

This graph is derived from CoreLogic building-level data within the study area, limited to post-1950 structures. The model land share is derived when imposing $\gamma(1-\beta) = 0.1736612$ and $\frac{\underline{f}\beta}{1-\gamma(1-\beta)} = 827.2036$. The histogram beneath charts the distribution of dwelling unit floor space within the sample.

figures highlight differences in land use efficiency among various dwelling types. Setting the minimum land cost share to 0.6 would imply that only single-family houses could be built. To allow for the construction of large multifamily buildings in the counterfactual scenario, I choose a minimum land cost share of 0.2 for the main analysis. The implications of this selection will be further explored in the discussion of the full model. Given the choice of minimum land cost share, we have $\gamma = 0.8$, $\underline{f} = 873.07$, and $\beta = 0.78$.

3.3.1 Model Validation

In this section, I aim to validate the model by predicting the cost of any dwelling unit ω in the dataset using only three observable factors: land input $k(\omega)$, floor space $f(\omega)$, and census tract $I(\omega)$. By comparing these predicted costs to actual data, we can assess the model's reliability.

To estimate the cost $p^{H,con}(\omega)$ based on land input $k(\omega)$ and floor space $f(\omega)$, I make specific assumptions about how these inputs are determined. I assume that floor space is optimally chosen given the land input, while allowing for the possibility that land input may not always be chosen optimally.

Starting from Equation (3) and conditioning on k (land input) while assuming that f (floor space) is optimally selected, the construction cost can be expressed as:

$$p_h^{H,con} = \min_f \left(c(h, f, k) + p^L k \right)$$

Further details of this derivation are provided in Appendix A.2.2. Solving this minimization problem yields the following expression for the construction cost:

$$p_h^{H,con}(\omega) = \left(\frac{f(\omega)}{B_{I(\omega)} \left(k(\omega) - a\left(f(\omega), k(\omega) \right) \right)^{1-\gamma}} \right)^{\frac{1}{\gamma}} + p_{I(\omega)}^{L} k(\omega)$$

Using this formula, I calculated the construction costs for all dwelling units built after 1950 and compared these with their most recent sale prices. The regression results, shown in Table C.1, indicate that the predicted construction costs explain 62% of the variance in sale prices. This result is significant, given that the predictions are based solely on the floor space, land input, and census tract for each unit.

Additionally, this procedure allows us to recover the dwelling quality for each dwelling unit in the dataset. I then calculated the minimal construction cost necessary to maintain the same quality for each unit and compared this with the predicted construction cost. This comparison provides a metric for evaluating the accuracy of the model's predicted dwelling attributes relative to the observed attributes in the data.

A substantial discrepancy between the predicted and observed attributes would suggest that the model's predictions are inconsistent with the actual data. Figure B.9 illustrates the relative price differences at both the individual dwelling unit level and the aggregated census tract level. The results show that the majority of observations exhibit less than a 10% difference, indicating a close alignment between the model's predictions and the observed attributes.

3.4 Quantitative Exercise

3.4.1 Comparison between FAR and DUPAC regulations

Section 3.2.2 outlines the process for converting FAR regulations into per-unit land input requirements. This conversion enables a direct comparison of the stringency of each regulation type within the study area. For a comparative analysis, both FAR and DUPAC regulations are converted into per-unit land input requirements. Figure 5 depicts the ratio of FAR to DUPAC at the tract level, once both are expressed as per-unit land input requirements. The results indicate that the median ratio is 0.13, suggesting that FAR regulations are typically around seven times less stringent than DUPAC regulations.

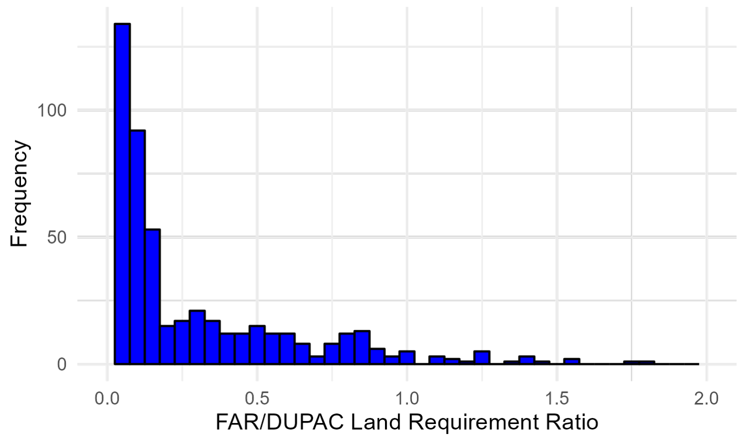


Figure 5: Tract-level Per-unit Land Requirement: Regulatory FAR v.s. DUPAC

3.4.2 Comparison of Minimum Dwelling Costs

To better understand the impact of zoning regulations on housing expenses for low-income households, I have calculated the minimum dwelling costs for each tract, both without constraints and in compliance with these regulations. Specifically, the minimum dwelling cost without constraints is derived by setting $h = 0$ in equation (8), while the cost with constraints is determined by setting h to the minimum compliance value from equation (11). It is important to note that these calculations do not account for the common use of special permits and zoning variances in the study area, focusing instead on the base zoning conditions as they stand.

Figure 6 displays a histogram illustrating the ratio of unconstrained to compliant costs across tracts. The average ratio is 0.19, while the median ratio is 0.16, indicating that, on average, zoning constraints increase the ”minimum dwelling cost” by fivefold.

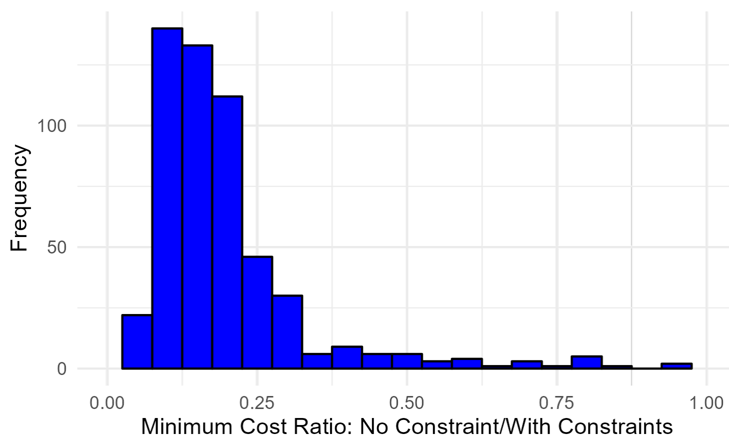


Figure 6: Tract-level Minimum Dwelling Cost: Unconstrained vs. Compliant Scenarios

4 Full Model

Geography. The model consists of $\mathcal{I}+1$ locations, with $\mathcal{I}$ representing within-city locations and one location outside the city.

To account for properties that deviate from the MAPC base zoning data due to grandfathering, I assume each location I comprises two sub-locations denoted as $i \in \{New, Old\}_I$. [3] These sub-locations share identical labor productivity $A_{I(i)}$ and communte costs to every other location $\tau_{I(i)J}$, yet they differ in their housing productivity B_i, residential amenity u_i, land endowment D_i, and a set of dwelling supply constraints Z_i along with a nonconforming cost v_i. [4] For sub-location New, the constraints include FAR and Density Regulations Z_i^{new}, which can be alleviated by incurring a nonconforming cost v_i. For sub-location Old, the constraint is inherited density Z_i^{old}, which cannot be mitigated, effectively making v_i infinite. [5]

4.1 Worker

There exists a fixed measure $\bar{N}$ of workers with heterogeneous effective labor e and property ownership status $o \in \{Owner, Renter\}$ jointly distributed according to G. Taking into considerations of income, dwelling price, residential amenity, and commute cost, workers initially decide whether to reside in the city. If yes, the worker will then make sequential decisions regarding their residential location, sub-location, work location, and consumption. [6] I solve the workers problem in a backward manner, first considering workers' consumption after selecting their residential location $I(i)$, sub-location i, and work location J. The worker ω's labor income is determined by $y^L(\omega) = e(\omega)w_J\tau_{I(i)J}\epsilon_J(\omega)$, where $\epsilon_J(\omega)$ represents the match-specific productivity for worker ω at location J. Additionally, worker with a $Owner$ status also receive transfers $t(y^L)$. These transfers are financed by the profit from the land

[3] A substantial proportion of properties in the study area exists as "grandfathered nonconforming uses",considering that approximately 58% of dwelling units constructed prior to 1956, which was the year a comprehensive zoning law with density limit was introduced,and approximately 25% of dwelling units constructed prior to 1918, which predates the implementation of any zoning policies.

[4] τ_{IJ} is the iceberg commuting cost with $0 < \tau_{IJ} \leq 1$ for all I, J, and $\tau_{II} = 1$ for all I.

[5] In the absence of an inherited density, the model would allow for smaller dwelling units than those observed in the grandfathered units. To ensure the model's dwelling stock aligns with empirical data, the inherited density is set to the 30th percentile of densities observed in the grandfathered units within each tract.

[6] In the study region, 70% of households own their residence. Given the housing transactions are costly, these households are less likely to change their residence when they faced a productivity shock. Additionally, 62.7% of households have resided in their current residence for more than five years. This indicates that a majority of households likely selected their residence prior to choosing their jobs, considering the national median tenure of 4.2 years with their current employer. Thus, the modeling assumption of selecting residences prior to observing the productivity draw appears reasonable. Source: 2015-2019 ACS data and 2018 Current Population Survey.

market. The total income of worker ω is $y(\omega) = y^L(\omega) + t_o(y^L(\omega))$, where $t_o(y^L(\omega)) = \mathbf{1}_{\{o(\omega)=Owner\}} t(y^L(\omega))$.

Workers choose their dwelling quality h based on their income $y(\omega)$ in order to maximize the following utility function:

$$U_{i,e}(y) = \max_h u_{i,e} \left(y - p^H_{i,h}\right)^\alpha (h)^{1-\alpha}, \tag{15}$$

where $p^H_{i,h}$ represents the price of a dwelling unit of quality h in sub-location i in location $I(i)$.

Employment Decisions

Conditional on having chosen the residential location $I(i)$ and sub-location i, individuals draw a vector of match-productivities with firms across the locations within the city i.i.d. from a Frechet distribution $F(\epsilon_J) = \exp\left(-\epsilon_J^{-\theta}\right)$.

With these draws in hand, linearity of (15) implies that workers choose to work in the location that offers the highest labor income $\max_J\{e(\omega) w_J \tau_{I(i)J} \epsilon_J(\omega)\}$. Properties of the Frechet distribution imply that the probability a worker residing in I and commuting to work in J is given by:

$$\pi_{J|I} = \frac{(w_J \tau_{IJ})^\theta}{\sum_{J' \in \mathcal{I}} (w_{J'} \tau_{IJ'})^\theta} \tag{16}$$

Conditioning on residential location I, the effective labor e does not impact the choice of work location J based on (16). This convenient feature allows us to estimate the average effective labor $\bar{e}_I$ in location I based on observed residents' average income, commuting choices, and workplace productivity. With $\bar{e}_I$ and current modeling assumptions, we can also derive the empirical distribution of effective labor $\hat{G}$ directly from the data. [7]

The distribution of labor income of workers with effective labor e in location I also follow a Frechet distribution, with cumulative distribution function (CDF) given by:

$$F_{e,I}(y^L) = F_I(y^L/e) = \exp\left[-\left(\frac{y^L/e}{(\sum_{J\in\mathcal{I}}(w_J\tau_{IJ})^\theta)^{\frac{1}{\theta}}}\right)^{-\theta}\right] \tag{17}$$

Therefore, the expected utility for choosing residential location $I(i)$ and sub-location i is:

[7] If we were to utilize alternative timing specifications, such as simultaneous determination of residential and workplace choices, the conditional commuting choice would also depend on effective labor e. Consequently, to derive $\bar{e}_I$, we would need to solve the complete model assuming certain worker type distribution G. This could complicate the computation and would not relax the modeling assumptions.

$$\bar{U}_{o,e,i} = \int_0^\infty U_{i,e}(y_o(y^L))dF_{I(i)}(y^L/e)$$

where y_o is the total income for workers with status o and labor income y^L.

Residential Location

Workers first draw idiosyncratic values $\xi_I(\omega)$ for every location and choose a location I. Subsequently, they draw idiosyncratic values $\xi_i(\omega)$ for each sub-location within that location and choose sub-location $i \in I$. I assume that the idiosyncratic values $\xi_I(\omega)$ and $\xi_i(\omega)$ are i.i.d. Frechet distributed with a shape parameter $\eta > 1$.

Given the choice of residential location I, the probability of choosing sub-location i can be expressed as:

$$\pi_{i|I,o,e} = \frac{\left(\bar{U}_{e,o,i}\right)^\eta}{\sum_{i' \in I} \left(\bar{U}_{e,o,i'}\right)^\eta} \tag{18}$$

The expected utility for living in location I is given by:

$$\bar{U}_{e,o,I} = \delta_\eta \left[\sum_{i \in I} (\bar{U}_{e,o,i})^\eta\right]^{1/\eta}$$

where $\delta_\eta = \Gamma((\eta - 1)/\eta)$ and $\Gamma(\cdot)$ denotes the Gamma function.

The probability for choosing residential location I is:

$$\pi_{I|e,o} = \frac{\left(\bar{U}_{e,o,I}\right)^\eta}{\sum_{I' \in \mathcal{I}} \left(\bar{U}_{e,o,I'}\right)^\eta} \tag{19}$$

Similarly, the expected utility for living in the city is given by:

$$\bar{U}_{e,o} = \delta_\eta \left[\sum_{I \in \mathcal{I}} (\bar{U}_{e,o,I})^\eta\right]^{1/\eta}$$

Migration Decision

Workers choosing either the city or an outside location based on the average expected utility within the city and a Frechet taste draw with a shape parameter η.[8] The migration decision is formulated as follows:

[8] I assume that the migration taste shock shares the same distribution as the residential taste shock within the city.

$$V(\omega) = \max_{\mathcal{I},o} \left\{ \bar{U}^{out}_{e(\omega),o(\omega)} v^{out}(\omega), \bar{U}_{e(\omega),o(\omega)} v(\omega) \right\} \tag{20}$$

where $\bar{U}^{out}_{e,o}$ is the utility of outside location for effective labor group e. I select $\bar{U}^{out}_{e,o}$ to match the city effective labor distribution with the population effective labor distribution G.

4.2 Consumption Goods Production

The consumption good is homogeneous, costlessly tradable, and produced under perfect competition. The technology uses effective labor E_J as input, and is $Y^L_J = A_J E_J$.

$$E_J = \bar{\epsilon}_J \sum_I \pi_{J|I} \bar{e}_I \tau_{IJ} N_I$$

$$\bar{\epsilon}_J = \sum_I \tau_{IJ} \pi_{J|I}^{-1/\theta} \frac{\pi_{J|I} N_I}{\sum_{I'} \pi_{J|i'} N_{I'}}.$$

Given the market structure and production technology, the wage rate to one unit of effective labor is $w_J = A_J$.

4.3 Dwelling Production and Constraints

Workers seeking residence in location I and sub-location $i \in I$ choose a dwelling quality h and place orders on the housing market. For each order of a dwelling unit with quality h, there are infinitely many builders competing for the contract, with the lowest bidder securing the order. Builders maximize their profit by minimizing the cost of producing a dwelling unit of quality h.

The detailed specification of the dwelling construction cost $p^{H,con}_{i,h}$ follows from Section 3, with one difference: the construction productivity B now varies by sub-location i.

Zoning regulations Z_i impose constraints on dwelling quality, as detailed in Section 3.2.2. Choosing a dwelling quality that does not conform to zoning regulations incurs a nonconforming cost v_i. In a competitive housing market, the final price of dwelling is given by:

$$p^H_{i,h} = p^{H,con}_{i,h} + v_i \mathbf{1}_{\{h \notin Z_i\}}, \tag{21}$$

where Z_i also represents the set of dwelling qualities that comply with the zoning policies Z_i.

4.4 Agglomeration

One of the major concerns raised by opponents of zoning reform is the potential change in neighborhood demographic composition. To quantify the effects of these changes, I decompose the city's unobserved amenity into an exogenous component and a component that depends on the endogenous demographic composition of each sub-location, following the approach of Tsivanidis (2022) and Pang (2021). Specifically, I assume that the residential amenity of a sub-location i depends endogenously on the proportion of residents with high effective labor. The utility function is given by:

$$u_{i,e} = u_{i,e,0} \left(\frac{N_{i,H}}{N_i} \right)^{\mu} \tag{22}$$

where $N_{i,H}$ represents the population of residents in location i with high effective labor (to be defined later), and N_i represents the total population in that location.

4.5 Land Market Clearing Condition

I denote the dwelling price as a function of land price by $p^H_{i,h}(p^L_i)$, as derived from equations (6) and (21). The optimal dwelling quality, based on income y and the dwelling price schedule p^H, is represented by $h^*(y, p^H)$, which is determined by solving equation (15). The function $k_i(h, p^L_i)$ denotes the cost-minimized land input required for a given dwelling quality and land price, according to equation (5). Let $\widetilde{k}_i(y, p^L)$ be defined as $k_i(h^*(y, p^H(p^L_i)), p^L_i)$.

The land market clearing condition is given by the following equation,

$$D_i = \int N_{iy}(p^L)\widetilde{k}_i(y, p^L), dy, \quad \forall i \in I, \quad \forall I \in \mathcal{I}, \tag{23}$$

where D_i is the total land endowment in sub-location i and N_{iy} is the measure of worker with income y living in sub-location i. The function $N_{iy}(p^L)$ is defined as:

$$N_{iy}(p^L) = \sum_o \int f_I \left(\frac{g_o^{-1}(y)}{e} \right) N_{e,o} \pi_{I|e,o}(p^L) \pi_{i|e,o,I}(p^L) de, \tag{24}$$

where $N_{e,o}$ is the measure of city residents with effective labor e and property ownership status o, $f_I(\cdot)$ is the probability density function of labor income from (17), and $g_o(y) = y + t_o(y)$. [9]

[9] Migration decision, characterized by equation (20), implies $N_{e,o} = \bar{N}(\bar{U}_{e,o})^{\eta}/((\bar{U}_{e,o})^{\eta} + (\bar{U}^{out}_{e,o})^{\eta})$.

4.6 Labor Market Clearing Condition

Follow from equation (16), the labor market clearing condition is given by the following equation,

$$N_J^W = \sum_I \frac{(w_J \tau_{IJ})^\theta}{\sum_{J' \in \mathcal{I}} (w_{J'} \tau_{IJ'})^\theta} N_I, \quad \forall J \in \mathcal{I}, \tag{25}$$

where N_J^W is the measure of workers working at location J, and N_I is the measure of residents at location I.

4.7 Budget Balance of Transfers

The profit from the land market is redistributed to the property owners. The following budget balance condition must be satisfied:

$$\sum_i D_i p_i^L = \sum_{I,e} N_{I,e,owner} \int_0^\infty t(y) dF_I \left(\frac{g_{owner}^{-1}(y)}{e} \right) \tag{26}$$

4.8 Equilibrium

Definition 1 *Given vectors of exogenous location characteristics* $\{A_I, B_i, D_i, u_{i,0}, \tau_{I,J}\}$*, land-specific dwelling constraints* $\{Z_i, v_i\}$*, populations characteristics* $\{\bar{N}, G\}$ *and model parameters* $\{\alpha, \beta, \theta, \eta, \gamma, \underline{f}, \mu\}$*, an equilibrium is defined as a vector of endogenous objects* $\left\{w_J, \pi_{J|I}, \pi_{i|I,e,o}, \pi_{I|e,o}, p_{i,h}^H, p_i^L, N_{i,e,o}, N_{i,y}, \widetilde{k}_i\right\}$ *such that:*

1. *Labor Market Clearing: The market for labor clears, as given by equation (25).*

2. *Housing Market Clearing: The cost of dwelling production equates the price of dwelling, as given by equation (21).*

3. *Land Market Clearing: The market for land clears, as given by equation (23).*

4. *Balance of Transfer: The transfer budget balance condition is satisfied, according to equation (26).*

5 Quantitative Analysis

5.1 Additional Data

In addition to the MAPC zoning data, CoreLogic housing data, and AEI land price data outlined in Section 2.1 for estimating the dwelling supply function, several additional datasets are necessary to estimate the spatial equilibrium model.

Demographic data is obtained from the Integrated Public Use Micro-data Series American Community Survey (IPUMS ACS). Specifically, I use the 2014-2019 five-year dataset to generate aggregate statistics, such as tract-level residential population and income. This dataset is also used to calculate key variables like the average housing consumption share, housing ownership probability across income levels, and the income distribution within the Greater Boston Area.

Commuting data, essential for understanding mobility patterns, is obtained from three main sources. First, the 2018 LEHD Origin-Destination Employment Statistics (LODES) provides aggregated commute flows at the census tract level. To estimate commute costs between tracts, I employ the Bing Distance API and the 2010-11 Massachusetts Travel Survey (MTS). The Bing Distance API calculates average travel times between geographic coordinates, accounting for existing traffic infrastructure and various transportation modes, offering insight into the time-costs for different transportation modes. The MTS supplies detailed information on residents' commute times and their choice of transportation modes, enabling the estimation of disutility linked to commuting times. I use these combined information to derive the overall commute costs.

5.2 Quantification

In this section, I estimate the model parameters and recover the model unobservables. The modelled city is the Greater Boston area, as defined by the 101 cities and towns administered by the Metropolitan Area Planning Council (MAPC). Additionally, I incorporate a outside hypothetical location to represent the rest of U.S.

5.2.1 Parameters Determined Externally

A subset of parameters are determined externally without solving the model.

Housing Consumption Parameter$\{\alpha\}$. I estimate α to align with the housing expenditure share at the 90th percentile of the income distribution in the Greater Boston area. Given that dwelling quality choice of high-income groups are less likely to be constrainted, their housing expenditure share is approximately $1-\alpha$. This feature allows me to obtain α from the expenditure share of the high-income group. In Figure B.10, I plot the relationship between the share of income spent on housing rent and labor income for renter. I set $\alpha = 0.79$.

Dwelling Production Parameters$\{\gamma, \beta, \underline{f}\}$. Following the procedures detailed in section 3.3, I derive an estimate of $\frac{\underline{f}\beta}{1-\gamma(1-\beta)}$ and $\gamma(1-\beta)$. To separate these three parameters, I need to choose a sensible γ. As discussed in section 3.3, γ determines the maximum land use efficiency for dwelling units within the model. I choose $\gamma = 0.8$ to accommodate the existence of dwelling units similar to those in large multifamily buildings, which typically have a land cost share of only 20%. With this selection of γ, we have $\underline{f} = 873.07$ and $\beta = 0.78$.

Migration Elasticity$\{\eta\}$. I set $\eta = 3.4$, which is consistent with estimates in Pang (2021).

Commute Cost$\{\tau_{IJ}\}$. Following Ahlfeldt et al. (2015), I assume the commute cost between two locations takes an exponential function form $\tau_{IJ} = exp(-\kappa t_{IJ})$, where κ denotes the disutility of longer commute, and t_{IJ} is the average commute time under certain traffic condition. Appendix A.4 lays out a multinomial logit model of transit mode choice to estimate κ and average commute time t_{IJ}, where a car owner can choose to commute via car, public transit and walk, otherwise only the latter two modes are available. The logit model is estimated using 2010-11 Massachusetts Travel Survey. κ is identified from how mode choice probability responds to mode-specific travel time, and t_{IJ} is the average travel time across modes using the mode choice probabilities predicted by the logit model. Table A.2 reports the results. Value of $\kappa = 0.035$ is higher compared to 0.01 in Ahlfeldt et al. (2015) and 0.015 in Pang (2021).

Productivity and Commute Elasticity$\{A_j, \theta\}$. In equilibrium, $w_J = A_j$. Following equation (16) and (25), the labor market clearing condition is represented by:

$$N_J^W = \sum_I \frac{(A_j \tau_{IJ})^\theta}{\sum_{J' \in \mathcal{I}} (A_{J'} \tau_{IJ'})^\theta} N_I, \quad \forall J \in \mathcal{I},$$

Given $\{N_J^W, N_I, \tau_{IJ}\}$, I select values for A_j to clear the labor market. Then, I select commute elasticity θ to maximize the likelihood of generating the commuting flows in LODES data. From this procedure, I recover $\theta = 2.5$. This number is lower than estimates in Tsivanidis (2022) and Pang (2021).

Worker Type Distribution$\{G\}$

Workers are heterogeneous in terms of their effective labor and property ownership status. The distribution G represents the joint distribution of these two worker characteristics.

To derive the distribution of effective labor G_e, I match the model's income distribution with the income distribution recovered from IPUMS ACS data. Details regarding this procedure are provided in Appendix A.5. This approach allows us to obtain both the city-level effective labor distribution G_e and a more granular, within-tract distribution of effective labor. For simplicity, I discretize the effective labor space into six distinct levels that can best approximate the empirical distribution, with each level representing approximately 10 to 20 percent of the population. In my model, the level of effective labor determines ex-ante incomes, which allows us to interpret these six groups as corresponding approximately to the 5th, 15th, 30th, 50th, 70th, and 90th income percentiles. It is noteworthy that the highest effective labor level has about 16 times more labor productivity than the lowest level, reflecting significant income disparities in the Boston area.

After estimating G_e, I obtain the conditional distribution of property ownership, $G_{o|e}$, by matching it with the property ownership shares across income percentiles in the data. Figure B.13 illustrates the correlation between homeownership rates and household income. I selected $G_{o|e}$ to match the data, with ownership rates for these income levels set at 0.25, 0.37, 0.54, 0.65, 0.75, and 0.85, respectively.

Agglomeration (μ) The strength of this amenity externality is set at $\mu = 0.3$, consistent with values used in existing studies Pang (2021); Tsivanidis (2022).

To define the "high effective labor group" in the context of amenity externality, I assume that workers not in the very low-income bracket are less likely to contribute to social disruptions. Therefore, I classify workers with an effective labor level of 3 or higher into this group, roughly corresponding to the upper 80% of the income distribution.

Transfer Rule$\{t(\cdot)\}$.

I assume the transfer amounts have this functional form: $t(y^L) = (y^L)^\lambda t_0$, where t_0 is a scaling factor to ensure the balance of transfer and λ is the transfer elasticity. This formulation ensures that households receive land rent in proportion to the share of real estate assets they would be expected to own at their income levels, reflecting the greater asset holdings of higher-income households.

To determine the transfer elasticity, I use data from IPUMS ACS, which include details on household income, homeownership, and house values. I restrict the dataset to Massachusetts households that fall within the 10th to 90th income percentiles and own their homes. The analysis results are shown in Table C.2.

5.2.2 Parameters Determined Internally

Residential amenities $\mu_{i,e}$ are selected to match the population distribution of each effective labor level across tracts as recovered from the data, following Appendix A.5. Although the model includes six effective labor levels, for stability reasons, these levels are grouped into three categories: the lowest two, the middle two, and the highest two. The residential amenities for effective labor levels within the same category are restricted to be identical and are chosen to exactly match the population distribution of each category across tracts.

Land prices p^L are set to clear the land market, given an estimated $B_i(p_i^L)^\gamma$, derived from regressing equation (12). Once land prices are established, housing productivity B_i can be determined. Finally, the nonconforming cost v_i is set to the smallest positive value that ensures the nonconforming rate at each location does not exceed the rates observed in the data.

5.3 Model Fit

Land Prices. Land prices in the model are determined by equilibrium conditions in the land market. Figure B.12 shows the relationship between land prices predicted by the model and those observed in the actual data. The model performs reasonably well, with a Pearson correlation coefficient of 0.73, indicating a strong positive correlation.

Commute Patterns. Figure B.14 illustrates the correlation between modeled and observed commute flows when aggregated at the PUMA level, where each PUMA encompasses roughly 10 nearby tracts. Tracts in close proximity tend to have similar commuting costs, which leads to comparable probabilities of commuting to other tracts. Aggregating data at the PUMA level provides a more robust evaluation of commute patterns, with the model achieving a high Pearson correlation coefficient of 0.97.

Housing Expenditure Share. Figure B.15 presents the model's estimated rent share for renters. The overall shape of the model's rent share curve closely resembles that of the data shown in Figure B.10, although the model's share has less variation. In the model, the housing expenditure share ranges from 50% to 24%, compared to a range of 60% to 18% observed in the data.

6 Counterfactual Analysis

6.1 Distributional Effects of Residential Zoning Policies

In this analysis, I use the calibrated model to simulate counterfactual scenarios in which zoning regulations had not been implemented since the 1950s. I then compare these results to the realized equilibrium in 2018 to understand the impacts of residential zoning policies.[10] It is important to clarify that the purpose of this analysis is not to predict the long-term equilibrium effects if zoning policies were removed today. This analysis does not consider the current equilibrium or the potential adjustment costs that could arise from deviating from it. Instead, the objective is to explore how welfare distribution might have differed had zoning regulations never been imposed.

Three distinct counterfactual scenarios are considered: (1) the absence of the density restriction (DUPAC), (2) the absence of FAR regulations, and (3) the absence of both DUPAC and FAR regulations. To generate these counterfactuals, I remove zoning constraints in the sub-location labeled New across various areas and calculate the resulting equilibrium outcomes. Table 2 presents a comparison of equilibrium outcomes for each worker type under these scenarios.

There are three key findings from this analysis:

First, FAR regulations are largely non-binding when DUPAC is present. The removal of FAR restrictions alone has a negligible effect on welfare across all worker types. However, when DUPAC is absent, FAR regulations become binding. This is evident when comparing the "No DUPAC" scenario to the "Neither" scenario, where the welfare of lower-income renter groups is higher in the latter.

Second, zoning policies substantially harm the welfare of lower-income groups. In the scenario without both regulations, the welfare of the renter group with the lowest effective labor levels increases by 41.6% compared to the current equilibrium. This outcome underscores the sensitivity of lower-income households to the average fivefold increase in the "minimum dwelling cost" attributed to zoning policies, as discussed in Section 3.4.2.

Third, the welfare of higher-income groups is mildly lower in the absence of zoning policies, with changes ranging from -1.2% to -2.5% relative to the current equilibrium. The housing choices of higher-income groups are typically less distorted by zoning policies, meaning that the observed welfare differences arise primarily from variations in equilibrium land prices and residential amenities. These variations, in turn, are driven by the different housing

[10] Although FAR regulations existed prior to the 1950s, no significant housing supply distortions attributable to these regulations were detected, as discussed in Section 2.4. Thus, I assume only housing built after the 1950s is affected by zoning regulations.

choices and mobility patterns of lower-income groups. Table 3 compares aggregate statistics between the current equilibrium and the counterfactual scenario where neither zoning policy was introduced. In the counterfactual scenario, average land prices are 9.5% lower, indicating a reduced rent burden for renters. Consequently, the reduced welfare for higher-income groups in the counterfactual is primarily due to a decline in residential amenities, which is linked to the larger share of lower-income households in neighborhoods. [11]

Effective Labor Level	No DUPAC		No FAR		Neither	
	Renter	Owner	Renter	Owner	Renter	Owner
1	+36.7%	+4.1%	0.0%	0.0%	+41.6%	+3.2%
2	+6.4%	-0.5%	0.0%	0.0%	+4.7%	-0.9%
3	-1.1%	-1.5%	0.0%	0.0%	-1.2%	-1.6%
4	-1.5%	-1.6%	0.0%	0.0%	-1.5%	-1.7%
5	-2.7%	-2.6%	0.0%	0.0%	-2.7%	-2.5%
6	-2.8%	-2.7%	0.0%	+0.1%	-2.5%	-2.4%

Table 2: Welfare Comparison, Zoning Impacts

The percentage changes are calculated using the 2018 equilibrium as the base. Column groups represent the scenarios in which DUPAC (Dwelling Units Per Acre), FAR (Floor Area Ratio) regulations or both are absent.

Population	Total	+9.5%
	By Level:	
	1	+157.3%
	2	+7.6%
	3	-4.7%
	4	-5.3%
	5	-8.5%
	6	-7.9%
Land Price		-9.5%

Table 3: % Comparison of Aggregates, Zoning Impacts

The percentage changes are calculated using the 2018 equilibrium as base and the counterfactual scenario that both DUPAC and FAR regulations are never introduced serving as the comparison. "By Level" refers to the breakdown of population changes by workers' Effective Labor Levels.

[11] Although average land prices are lower in the counterfactual scenarios, indicating a reduced total amount of rent distributed to property owners, property owners in this setup do not actually suffer from this reduction. Similar to renters, property owners are free to relocate to other cities. In the counterfactuals, there are fewer property owners in the city to claim the available land rent, which offsets the decline in the aggregate rent amount.

6.2 Effects of Radical Zoning Reforms

In this section, I examine the long-term equilibrium outcomes after lifting zoning policies completely from the 2018 baseline equilibrium. A key challenge in this analysis is accounting for the state dependency of the current equilibrium. To address this, I made two important adjustments to derive welfare outcomes accurately for different types of workers.

First, I restrict the mobility of property owners. Like renters, property owners make decisions about their residential location, workplace, housing, and consumption. However, unlike renters, they are no longer permitted to relocate outside the city. Without this restriction, the population of property owners could adjust in response to changes in land values, affecting how rents are distributed. Such adjustments would be necessary if the goal were to assess outcomes in a scenario where zoning policies had never been introduced, as the population of property owners would likely differ from the existing equilibrium. However, when assessing the impact of zoning reform, allowing such adjustments could diminish or even reverse the gains or losses that property owners experience from changes in land values, as rents are distributed solely to those residing within the city. By fixing the property owner population, rents are shared among a stable group, ensuring that property owners directly experience the effects of changes in land values.

Second, I account for redevelopment costs, but instead of tying these costs to housing, I associate them with land use. For each census tract, I track the amount of land allocated to dwelling units that comply with existing zoning policies and those that do not—referred to as compliant and non-compliant land, respectively. In the counterfactual scenario where zoning policies are lifted, if the amount of non-compliant land increases (where non-compliance is corresponding to the old zoning policies, not the updated ones), a redevelopment fee is imposed on the price of the non-compliant land. This fee reflects the cost of converting compliant land into non-compliant land, covering expenses such as purchasing compliant buildings, demolishing them, and preparing for the construction of non-compliant structures.

Specifically, the land price for non-compliant land in sub-location i in the counterfactual scenario is given by:

$$p_i^{L,N'} = \frac{p_i^D \min(k_i^{N'} - k_i^N, 0)}{k_i^{N'}} + p_i^{L,C'}$$

where $p_i^{L,C'}$ represents the land price for compliant land in the counterfactual scenario, $k_i^{N'}$ is the amount of non-compliant land in the counterfactual scenario, k_i^N is the amount of non-compliant land in the baseline equilibrium, and p_i^D is the redevelopment cost per unit of land. Since p_i^D captures the cost of purchasing and demolishing compliant buildings, I take a conservative approach by selecting it to match the highest per-unit-of-land structure cost among compliant buildings for each location.

With this framework, I evaluate three distinct counterfactual scenarios: (1) the removal of the density restriction (DUPAC), (2) the removal of FAR regulations, and (3) the removal of both DUPAC and FAR regulations. Table 4 presents the welfare changes for each worker type under these scenarios, while table 5 summarizes the changes in aggregate statistics.

Lifting the zoning policies significantly increases the welfare of renters with the lowest effective labor level by 34.7%, and the population of this group grows by 120.5%. In contrast, the welfare of renters with the second lowest effective labor level sees only a modest increase of 2.6%, while other worker types experience welfare losses ranging from -0.3% to -2.9%. For renters, these welfare losses are primarily driven by a decline in residential amenities, which is linked to the larger share of lower-income households in neighborhoods. This shift reflects a common opposition to zoning reform, where critics argue that loosening regulations could lead to an influx of lower-income households in previously affluent areas, raising concerns about potential changes in crime rates or social disruption. The welfare reductions observed here align with these concerns. For property owners, additional negative impacts arise from a decline in land prices.

One main concerns of zoning reform among existing homeowners is the potential for declining property values. Property values can be divided into structure value and land value. In my model, structures are always priced competitively, meaning that any gains or losses due to demand shifts are fully reflected in changes to land prices. To understand the impact of the 4.5% average decline in land prices, consider that the median dwelling unit in the Greater Boston Area has a land cost share of 60%. Therefore, a 4.5% decline in land value results in a 2.7% reduction in overall property values.

Moreover, there is considerable variation in land price changes across tracts. Figure 7 illustrates that a significant number of tracts experience declines in land values of more than 10%, with some suburban tracts seeing reductions exceeding 30%. For clarity, I assume the median land cost share of 60% and calculate the cumulative distribution function of tract-level average property value changes, as shown in Figure 8. In 60% of the tracts, property values increase by 2% to 5%, while in 30% of the tracts, property values decrease. Of those, 20% experience declines greater than 5%, and 10% see reductions exceeding 10%. These results align with concerns about declining property values due to zoning reform.

Effective Labor Level	No DUPAC		No FAR		Neither	
	Renter	Owner	Renter	Owner	Renter	Owner
1	+28.5%	-0.3%	0.0%	+0.1%	+34.7%	-0.3%
2	+2.3%	-3.1%	0.0%	+0.1%	+2.6%	-3.6%
3	-1.2%	-2.0%	0.0%	0.0%	-1.4%	-2.4%
4	-1.3%	-1.8%	0.0%	0.0%	-1.5%	-2.2%
5	-2.4%	-2.7%	0.0%	0.0%	-2.6%	-2.9%
6	-2.4%	-2.6%	0.0%	+0.1%	-2.3%	-2.6%

Table 4: Welfare Change, Radical Zoning Reform

The percentage changes are calculated using the 2018 equilibrium as the base. Column groups represent the scenarios in which DUPAC (Dwelling Units Per Acre) or FAR (Floor Area Ratio) regulations are absent, or neither regulation is in effect.

Population	Total	+10.0%
	By Level:	
	1	+120.5%
	2	+4.9%
	3	-2.2%
	4	-1.7%
	5	-2.1%
	6	-1.2%
Land Price		-4.5%

Table 5: % Change of Aggregates, Radical Zoning Reform

The percentage changes are calculated using the 2018 equilibrium as base and the counterfactual scenario that both DUPAC and FAR regulations are removed serving as the comparison. "By Level" refers to the breakdown of population changes by workers' Effective Labor Levels.

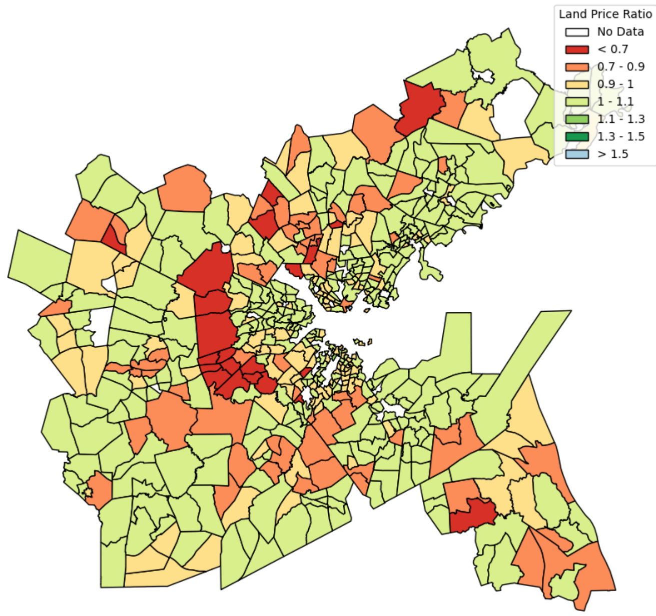


Figure 7: Land Price Ratio, Radical Zoning Reform

The ratiio are calculated using the tract-level land prices of the 2018 equilibrium as denominators and the land prices of the counterfactual scenario that both DUPAC and FAR regulations are removed as numerators.

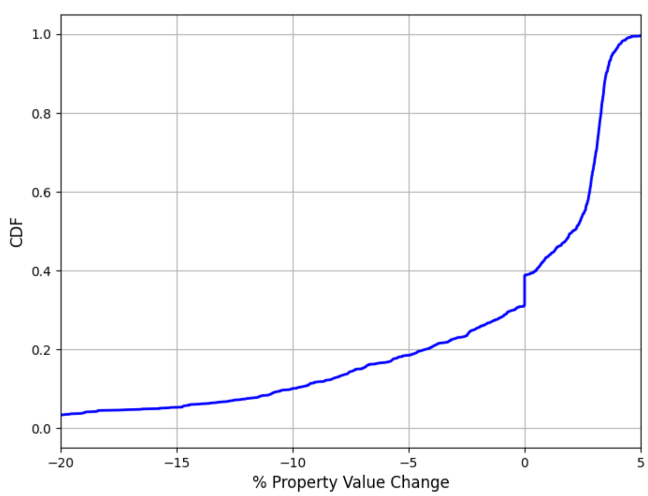


Figure 8: Tract-level Average Property Value Change, Radical Zoning Reform

Each observation represents a census tract in the Greater Boston Area. The percentage changes are calculated using the 2018 equilibrium as base and the counterfactual scenario that both DUPAC and FAR regulations are removed serving as the comparison. The impacts on property values are derived from the changes in land prices, assuming a land cost share of 60% for each property, which represents the median land cost share for dwelling units in the study region.

7 Concluding Remarks

This paper investigates the distributional impact of zoning regulations in the Greater Boston Area, focusing on density and Floor Area Ratio (FAR) restrictions. The findings reveal that zoning regulations have significantly increased the "minimum dwelling cost," disproportionately affecting lower-income households by limiting their access to affordable, high-density housing. The counterfactual analysis shows that, without zoning constraints, welfare for the lowest 10% income group could have been 41.6% higher, while higher-income groups would experience a mild welfare decline.

The analysis of radical zoning reforms reveals both benefits and costs. While removing zoning restrictions would significantly improve welfare for the lowest 10% income renter households—up to a 34.7% increase—it would also result in welfare losses for the majority of residents due to reduced residential amenities. Property owners would face further impacts from declining property values, with an average decrease of 2.7% and significant variation across tracts, some experiencing reductions of over 10%. These findings suggest that although radical zoning reforms could enhance welfare for lower-income households and reduce inequality, they would also impose losses on the majority of current residents, highlighting the need to carefully consider the distributional impacts and take a more balanced approach when implementing zoning policy reforms.

References

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

Athey, Susan, Julie Tibshirani, and Stefan Wager, "Generalized random forests," *The Annals of Statistics*, April 2019, *47* (2).

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *journal of political economy*, 2007.

Bobrowski, Mark, *Handbook of Massachusetts Land Use and Planning Law: Zoning, Subdivision Control, and Nonzoning Alternatives*, Wolters Kluwer, 2002.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, "Income Growth and the Distributional Effects of Urban Spatial Sorting," *Review of Economic Studies*, March 2024, *91* (2), 858–898.

Davis, Morris A., William D. Larson, Stephen D. Oliner, and Jessica Shui, "The price of residential land for counties, ZIP codes, and census tracts in the United States," *Journal of Monetary Economics*, March 2021, *118*, 413–431.

Glaeser, Edward L. and Bryce A. Ward, "The causes and consequences of land use regulation: Evidence from Greater Boston," *Journal of Urban Economics*, May 2009, *65* (3), 265–278.

Glaeser, Edward L, Joseph Gyourko, and Raven E Saks, "Why Have Housing Prices Gone Up?," 2005, *95* (2).

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst, "Endogenous gentrification and housing price dynamics," *Journal of Public Economics*, April 2013, *100*, 45–60.

Heblich, Stephan, Stephen J Redding, and Daniel M Sturm, "The Making of the Modern Metropolis: Evidence from London*," *The Quarterly Journal of Economics*, November 2020, *135* (4), 2059–2133.

Herkenhoff, Kyle F., Lee E. Ohanian, and Edward C. Prescott, "Tarnishing the golden and empire states: Land-use restrictions and the U.S. economic slowdown," *Journal of Monetary Economics*, January 2018, *93*, 89–109.

Kulka, Amrita, Aradhya Sood, and Nicholas Chiumenti, "How to Increase Housing Affordability? Understanding Local Deterrents to Building Multi-family Housing," 2023.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins, "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets," *Journal of Economic Literature*, December 2013, *51* (4), 1007–1062.

MacArthur, Will, *The Kind of City Which is Desirable and Obtainable*, Cambridge, 2019.

Ospital, Augusto, "Urban Policy and Spatial Exposure to Environmental Risk," April 2023.

Pang, Xinle, "The Welfare Effect of Spatial Mismatch: Evidence From the New York Metropolitan Area," 2021.

Redding, Stephen J and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," 2017.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman, "Cognitive Hubs and Spatial Redistribution," February 2021.

Tokman, Anthony E, "Density Restrictions and Housing Inequality," November 2023.

Tsivanidis, Nick, "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio," 2022.

A Appendix

A.1 Deriving DUPAC and FAR restrictions

Dwelling Units per Acre (DUPAC)

Although the Dwelling Units per Acre (DUPAC) is explicitly specified in the Bylaw or Ordinance for only about 7% of the dwelling units, we have additional information on the Maximum Dwelling Units per lot and Minimum Lot Size restriction for approximately 93% of the dwelling units within the study region. These values allow us to derive an inferred DUPAC using the formula:

$$\text{inferred DUPAC} = \frac{\text{Maximum Dwelling Units per lot}}{\text{Minimum Lot Size restriction}}$$

In Figure A.1a, I compare the distribution of the inferred DUPAC with the specified DUPAC. As illustrated, the inferred DUPAC closely aligns with the specified DUPAC, demonstrating that the inferred DUPAC can serve as a valid proxy when the specified DUPAC is unavailable.

Floor Area Ratio (FAR) Restriction

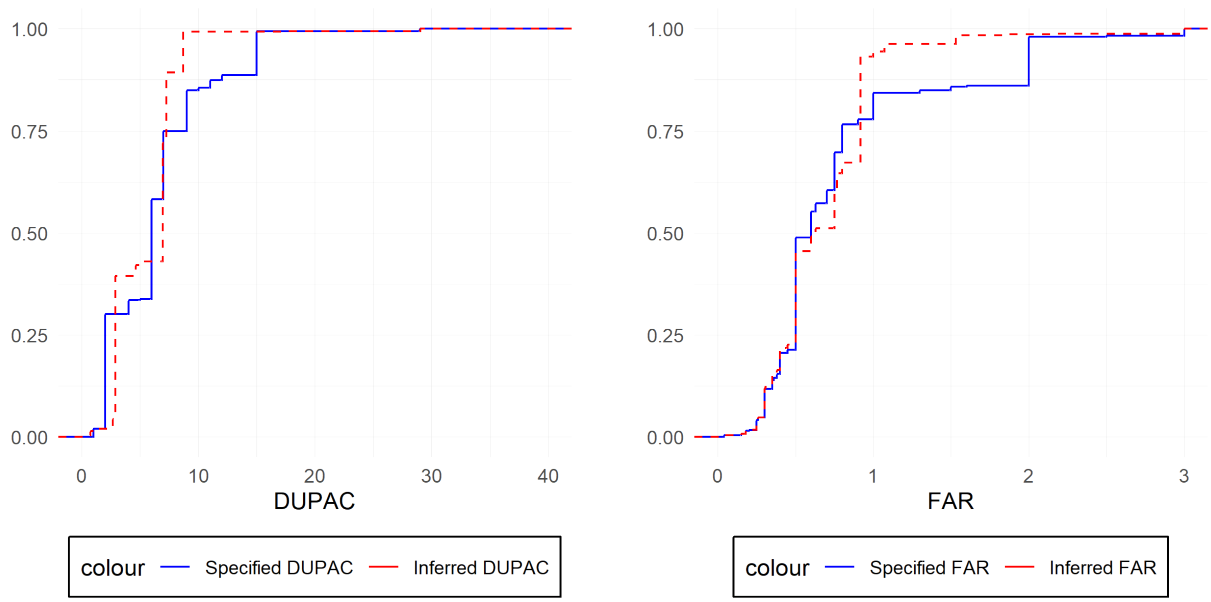
The Floor Area Ratio (FAR) restriction is explicitly provided for approximately 26% of the dwelling units in the Bylaw or Ordinance. However, we do have the Maximum Floor value specified for about 98% of the dwelling units. In cases where the FAR restriction is not directly available, I derive an inferred FAR using the Maximum Floor value.

To calculate this inferred FAR, I rely on dwelling units where both the FAR restriction and Maximum Floor restriction are specified. From these, I derive a conservative estimate of their ratio, which is 0.307.

The formula for calculating the inferred FAR is given by:

$$\text{Inferred FAR Restriction} = \text{Maximum Floor} \times 0.307$$

Figure A.1b compares the specified FAR values with the inferred FAR values. The results show that for 80% of the dwelling units, the inferred FAR aligns well with the specified FAR. For the remaining 20%, the inferred FAR tends to be stricter, especially for units with specified FAR values of 0.9 or higher. These higher FAR values are typically associated with lands designated for high-density dwelling units. Consequently, the inferred FAR regulation is often more stringent than what the Maximum Floor restriction alone would imply for these high-density units.



(a) DUPAC

(b) FAR

Figure A.1: Specified Restriction versus inferred Restrictions

This graph plots the cumulative density distribution for zoning restrictions. Each observation is a dwelling unit. Sample is restricted to dwelling units built after 1960 and have both specified and inferred information available.

A.2 Technical Details

A.2.1 Solving Unconstrained Housing Production

We first look at the construction cost of dwelling units when there is no constraints on the housing attributes.

From (1) and (2), we obtain $a(h,f) = (\frac{h}{(f-\underline{f})^{1-\beta}})^{\frac{1}{\beta}}$, and $c(h,f,k) = (\frac{f}{B(k-a(h,f))^{1-\gamma}})^{\frac{1}{\gamma}}$. For each quality level h, builders would choose attributes f and k to minimize the construction cost $p_h^{H,con}$:

$$p_h^{H,con} = \min_{(f,k)} p^c c(h,f,k) + p^L k, \tag{27}$$

where p^c represents the price of construction goods and p^L represents the price of land. The first term $p^c c(h,f,k)$ represents the cost of building the structure, and the second term $p^L k$ represents the cost of land. In this paper, construction goods are treated as the numeraire in the economy, and their price is normalized to $p^c = 1$.

The optimal land input for floor space production satisfies the equation $\frac{c}{g} = \frac{\gamma p^L}{(1-\gamma)}$. Consequently, the cost of producing one unit of floor space can be expressed as

$$p^f = \frac{(p^L)^{1-\gamma}}{B z_\gamma},$$

where $z_\gamma = \gamma^\gamma (1-\gamma)^{1-\gamma}$.

The builder's problem in (27) can be formulated as minimizing the costs for floor space and density amenities at each quality h:

$$p_h^{H,con} = \min_a f(h,a) p^f + a p^L,$$

where $f(h,a) = (\frac{h}{a^\beta})^{\frac{1}{1-\beta}} + \underline{f}$.

Solving this problem yields the optimal density amenities and production cost

$$a = \frac{\beta h}{z_\beta} (B z_\gamma (p^L)^\gamma)^{\beta-1} \tag{28}$$

$$p_h^{H,con} = \frac{(p^L)^{1-\gamma}}{B z_\gamma} \left((B z_\gamma (p^L)^\gamma)^\beta \frac{h}{z_\beta} + \underline{f} \right) \tag{29}$$

where $z_\beta = \beta^\beta (1-\beta)^{1-\beta}$.

By substituting the optimal amenities choice a into the function $f(h,a)$, we obtain

$$f=\frac{(1-\beta)h}{z_\beta}(Bz_\gamma(p^L)^\gamma)^\beta+\underline{f} \tag{30}$$

Given (2) and the optimality condition of floor space land input, we derive

$$g=\frac{f(1-\gamma)}{Bz_\gamma(p^L)^\gamma} \tag{31}$$

Adding (28) and (31), we determine the total land input used in housing production at given quality h:

$$k=\frac{1}{Bz_\gamma(p^L)^\gamma}\left((Bz_\gamma(p^L)^\gamma)^\beta\frac{h}{z_\beta}\left(1-\gamma(1-\beta)\right)+\underline{f}(1-\gamma)\right) \tag{32}$$

A.2.2 Solving Housing Price When k is not Optimally Chosen

Beginning with Equation (3), and conditioning on k while assuming f is optimally chosen, the construction cost is given by:

$$p_h^{H,con}=\min_f c(h,f,k)+p^L k$$

The first order condition is given by:

$$\left(\beta(\gamma f-\underline{f})+(1-\gamma)f\right)\left(h(f-\underline{f})^{\beta-1}\right)^{\frac{1}{\beta}}=\beta k(f-\underline{f})$$

After substituting from Equation (1) and rearranging, we obtain:

$$a(f,k)=\frac{\beta k(f-\underline{f})}{\beta(\gamma f-\underline{f})+(1-\gamma)f} \tag{33}$$

Using Equations (1) and (33), we map $k(\omega)$ and $f(\omega)$ to housing quality $h(\omega)$:

$$h(\omega)=a\left(f(\omega),k(\omega)\right)^\beta\left(f(\omega)-\underline{f}\right)^{1-\beta}$$

Revisiting the initial equation (3), we have:

$$p_h^{H,con}(\omega)=\left(\frac{f(\omega)}{B_{I(\omega)}\left(k(\omega)-a\left(f(\omega),k(\omega)\right)\right)^{1-\gamma}}\right)^{\frac{1}{\gamma}}+p^L_{I(\omega)}k(\omega)$$

A.3 Checking Housing Attribute Distortion

A.3.1 Check Attribute Distortions With Causal Forest

The goal is to assess whether housing attributes are influenced by proximity to regulatory thresholds. Specifically, I examines whether the proximity to these thresholds affects the built Floor Area Ratio (FAR), conditioning on built density. The assessment is not straightforward as regions with different regulatory thresholds also have different regional fundamentals —such as housing productivity, demographics, and residential amenities— that could affect housing attributes. Currently, there is no theoretical framework that links regional fundamentals with housing attributes. Moreover, zoning enforcement varies across regions, as depicted in Figure B.11, potentially leading to treatment heterogeneity where stricter enforcement might cause more noticeable distortions.

To tackle these complexities, I propose the following nonparametric model to estimate heterogeneous treatment effects:

$$y = T^{reg}\theta^{reg}(X) + g(Z, V, WV) + \epsilon$$

Here, y is the built FAR of housing, T^{reg} is dummy variables that equal one when the housing's attribute is within 10% of respective regulatory threshold, where $reg \in \{Density, FAR\}$. The functions θ^{reg} and g are nonparametric. X is a vector of tract-level regulation compliance rates. Z includes information on built density and additional housing-level information such as construction year, housing type, and distance from the city center. V includes tract-level data such as land price, population, total land size, total housing stock, and per capita income. WV describes the average characteristics of neighboring tracts, acknowledging that the broader neighborhood context can also influence housing traits.

I estimated the above model using causal forest, which is discussed in Athey et al. (2019). This approach automatically capture complex interactions and nonlinear relationships in the data without requiring explicit specification of these relationships. Table A.1 presents the estimated average treatment effects (ATEs) across various subsamples categorized by tract-level compliance rates, derived from conditional average treatment effects (CATEs) estimated using causal forest. The estimates for $\theta^{FAR}(x)$ exhibit very large standard errors, primarily because only a limited number of housing units are located near the FAR threshold, as shown in 3. Specifically, only 0.4% of housing units fall within 10% of the FAR threshold, compared to over 17% within 10% of the Density threshold. Focusing on the estimates of $\theta^{Density}(x)$, we see ATE is only significant for the subsample where tract-level compliance rate is between 0 % to 30 %. Even for this subsample, the effect is only -0.04, which isn't a large number.

The population ATE in the column (1) is only -0.006 and insignificant. From this analysis, I did not find evidence that zoning regulation induces any housing attribute distortion relevant to my supply model. Therefore, I proceed with the assumption that zoning regulation only distort the choice of qualities, but not how houses are built.

Focusing on the estimates of $\theta^{Density}(x)$, the ATE is only significant for the subsample where the tract-level compliance rate is between 0% and 30%. Even for this subsample, the effect is minimal, at -0.04. The population-wide ATE presented in column (1) is a mere -0.006 and statistically insignificant. Based on this analysis, I found no evidence that zoning regulations cause distortions in housing attributes that are relevant to my supply model. Consequently, I will proceed with the assumption that zoning regulations primarily distort the choice of qualities, rather than the construction features of houses.

	(1) $0 \leq x \leq 1$	(2) $0 \leq x \leq 0.3$	(3) $0.3 < x \leq 0.6$	(4) $0.6 < x \leq 0.9$	(5) $0.9 < x \leq 1$
$\theta^{Density}(x)$	-0.006	-0.043*	-0.018	0.008	0.100
	(0.082)	(0.022)	(0.012)	(0.011)	(0.254)
$\theta^{FAR}(x)$	-2.313	-1.977	-2.362	-3.175	-0.530
	(9.311)	(10.624)	(6.280)	(9.706)	(11.194)

Table A.1: ATE by Compliance Rate Subsamples for Proximity to Regulatory Thresholds

The sample limited to post-1950 housing units. The number of observation is 210,575. (Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)

A.4 Commute Cost

I estimate the commute costs between two census tracts, denoted as τ_{IJ} in the main text, by following the procedure outlined below. I determine the geographic location of each tract using the geo-coordinates of tract centroids. The travel distance (in miles) by car, as well as travel durations (in minutes) by both car and public transit, are obtained under the current average traffic conditions via the Bing Distance API.

Consistent with Ahlfeldt et al. (2015), I adopt an exponential functional form for the commute cost, expressed as

$$\tau_{IJ} = \exp(-\kappa \cdot t_{IJ}),$$

where t_{IJ} represents the average commute time between tracts I and J across different transit modes, and κ quantifies the disutility associated with longer commute times. To estimate these parameters, I employ a multinomial logit model for transit mode choice. For a commuter traveling between locations I and J, the possible transit modes include walking, public transit, and driving (if the commuter owns a car); otherwise, the options are limited to walking and public transit. The probability of choosing mode m is modeled as:

$$\pi_{IJ}^{m} = \frac{\exp(-\kappa t_{IJ}^{m} + \xi^{m})}{\sum_{m'} \exp(-\kappa t_{IJ}^{m'} + \xi^{m'})},$$

where t_{IJ}^{m} denotes the commute time if mode m is selected, and ξ^{m} represents the mean preference for mode m, encapsulating all other factors that the commuter values besides commute time.

Utilizing data from the 2010-2011 Massachusetts Travel Survey, which includes 8,192 observed commute patterns from the study area, I estimate the mode choice model. This model is fitted using standard maximum likelihood methods, and the results are presented in Table A.2. The estimated value of $\kappa = 0.035$ is higher than the values reported in previous studies, specifically 0.01 in Ahlfeldt et al. (2015) and 0.015 in Pang (2021).

For the year 2017, and for each pair of census tracts, I compute the average commute time as follows:

$$\bar{t}_{IJ} = \sum_{m} \pi_{IJ}^{m} t_{IJ}^{m},$$

where t_{IJ}^{m} is the commute time for mode m as computed previously, and π_{IJ}^{m} is the mode choice probability derived from the logit model. Considering three modes of transportation—car, public transit, and walking—I calculate the average commute time $\bar{t}_{IJ}$. The corresponding commute cost is then determined as

$$\tau_{IJ} = \exp(-\kappa \bar{t}_{IJ}).$$

	Logit Estimate
κ	0.034***
	(0.001)
ξ^{car}	0.244***
	(0.067)
$\xi^{\text{publictransit}}$	0.058
	(0.055)
Observations	8,060

Table A.2: Commute mode choice

Distuility of walk ξ^{walk} is normalized to 0. Data comes from 2010-11 Massachusetts Travel Survey. Observation is a commute between home and workplace of a employed sampled person. Standard errors are reported in parentheses. $*p < 0.1; **p < 0.05; ***p < 0.01$.

A.5 Estimate Effective Labor

A.5.1 Obtain household income distribution within each tract

Estimating effective labor requires tract-level household income distributions. I utilize the American Community Survey (ACS) 5-year data, which includes both an aggregated data file and a public microdata file. The aggregated data file provides tract-level income distributions categorized into predefined income brackets. While these brackets offer sufficient granularity for lower income levels, they are sparse for higher income levels, posing challenges for accurate effective labor estimations. The public microdata file contains household-level income information but only identifies geographic locations at the Public Use Microdata Area (PUMA) level, typically encompassing around 20 census tracts.

To construct the income distribution for each census tract, I weight the household-level data from the PUMA according to the population in the income brackets of the respective tract. This procedure assumes the income distribution is consistent across different tracts within the same income bracket and PUMA area. This assumption is defensible, as the income brackets are fine enough that significant variation within the same bracket across tracts in the same PUMA is unlikely.

A.5.2 Estimating Effective Labor Distribution

The household labor income distribution, given effective labor e and location I, is characterized by equation (17).

Labor incomes y_I^L for residents living in different locations I are not directly comparable for estimating effective labor. Therefore, I consider labor income adjusted for the local wage index, denoted as $y^{\text{adj}} = \frac{y_I^L}{\left(\sum_{J\in\mathcal{I}}(w_J\tau_{IJ})^\theta\right)^{1/\theta}}$. The cumulative distribution function (CDF) for adjusted labor income, given effective labor e, is:

$$F_e(y^{\text{adj}}) = \exp\left[-\left(y^{\text{adj}}/e\right)^{-\theta}\right]$$

Using the Law of Total Probability, the CDF of adjusted labor income is given by:

$$F(y^{\text{adj}}) = \int_e F_e(y^{\text{adj}})p(e)\,de,$$

where $p(e)$ characterizes the distribution of effective labors.

Once we recover $F(y^{\text{adj}})$ from the income distribution data, we can invert the equation to recover $p(e)$.

To solve the inversion problem, I first discretize $F(y^{\text{adj}})$ into a vector $\vec{v_a}$, discretize $F_e(y^{\text{adj}})$ into a matrix $M_{a,e}$, and discretize $p(e)$ into a vector $\vec{v_e}$. This results in the following problem:

$$\vec{v_a} = M_{a,e}\vec{v_e} \tag{34}$$

where I need to solve for $\vec{v_e}$ given $\vec{v_a}$ and $M_{a,e}$.

I focus on finding a positive solution for $\vec{v_e}$. In practice, I found that an exact positive solution does not exist, regardless of the fineness of the discretization, so I minimize the mean square error instead. However, as I increase the fineness, the recovered CDF converges. Moreover, the CDF can be very well captured by a small selection of effective labor grids.

In Figure B.18, I show the results for both the full grid and a subgrid. For the full grid, effective labors are discretized into 1000 equally spaced levels from 100 to 50000, and adjusted labor income is discretized into 200 equally spaced percentile levels. For the subgrid, effective labors are discretized into only 6 levels: {1900, 4200, 7300, 11700, 18700, 30200}. We can see that even with the subgrid, most of the distribution pattern is captured.

In Figure B.19, I compare the CDF of adjusted labor income from the data with that estimated using an effective labor grid of only 6 levels. The CDFs match very well, suggesting that having only 6 effective labor levels is sufficient to characterize the income distribution in the study region.

A.5.3 Model-Based Decomposition of Residents into Effective Labor Levels by Location

For each location, I solve the following problem:

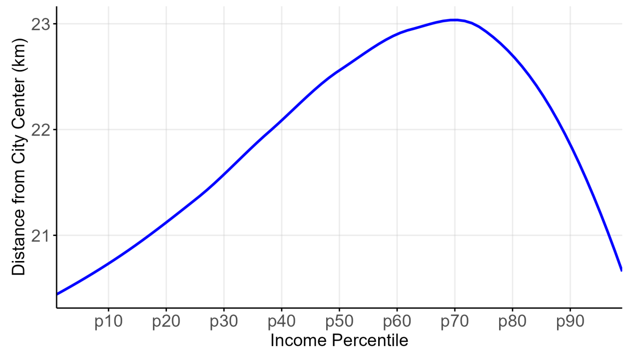
$$\vec{v}_{I,a} = M_{a,e}\vec{v}_{I,e}$$

where $\vec{v}_{I,e}$ is the vector I need to solve for, given $M_{a,e}$ and the vector $\vec{v}_{I,a}$, which characterizes the adjusted income distribution for location I.

Since a subgrid of 6 effective labor levels is sufficient to closely match the income distribution, I assume there are only 6 effective labor levels in each location. Solving the above problem reveals the share of residents in location I in each of the 6 effective labor levels.

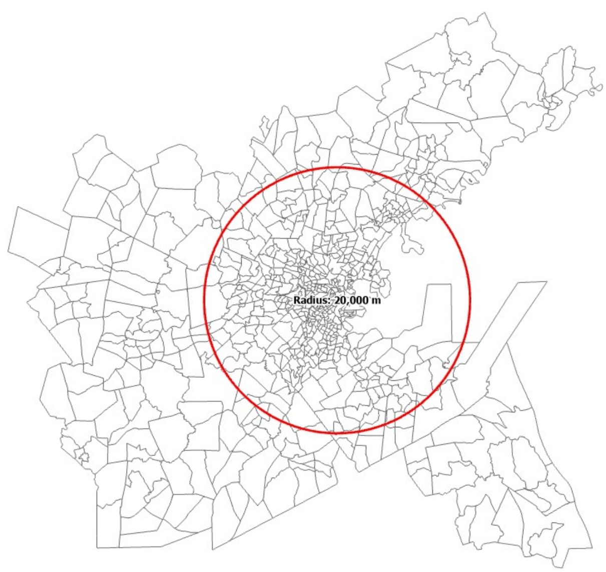
B Additional figures

Figure B.1: Resident income and distance from city center



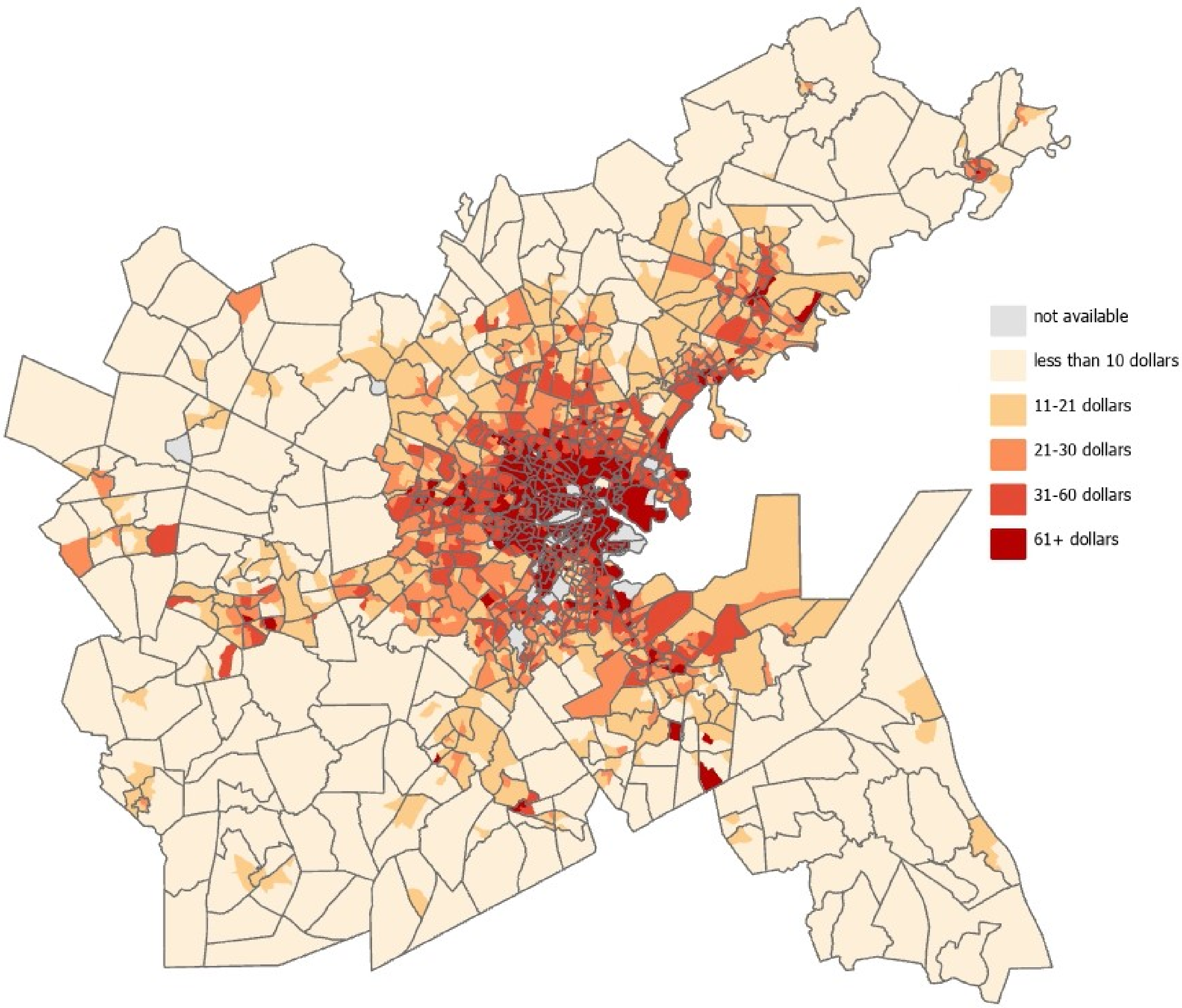
This graph is constructed using LOESS regression on merged CoreLogic-HMDA data from the study region. The sample limits to homeowner living in their primary residence. City center is set at the location of Boston City Hall.

Figure B.2: 20-km radius around city center



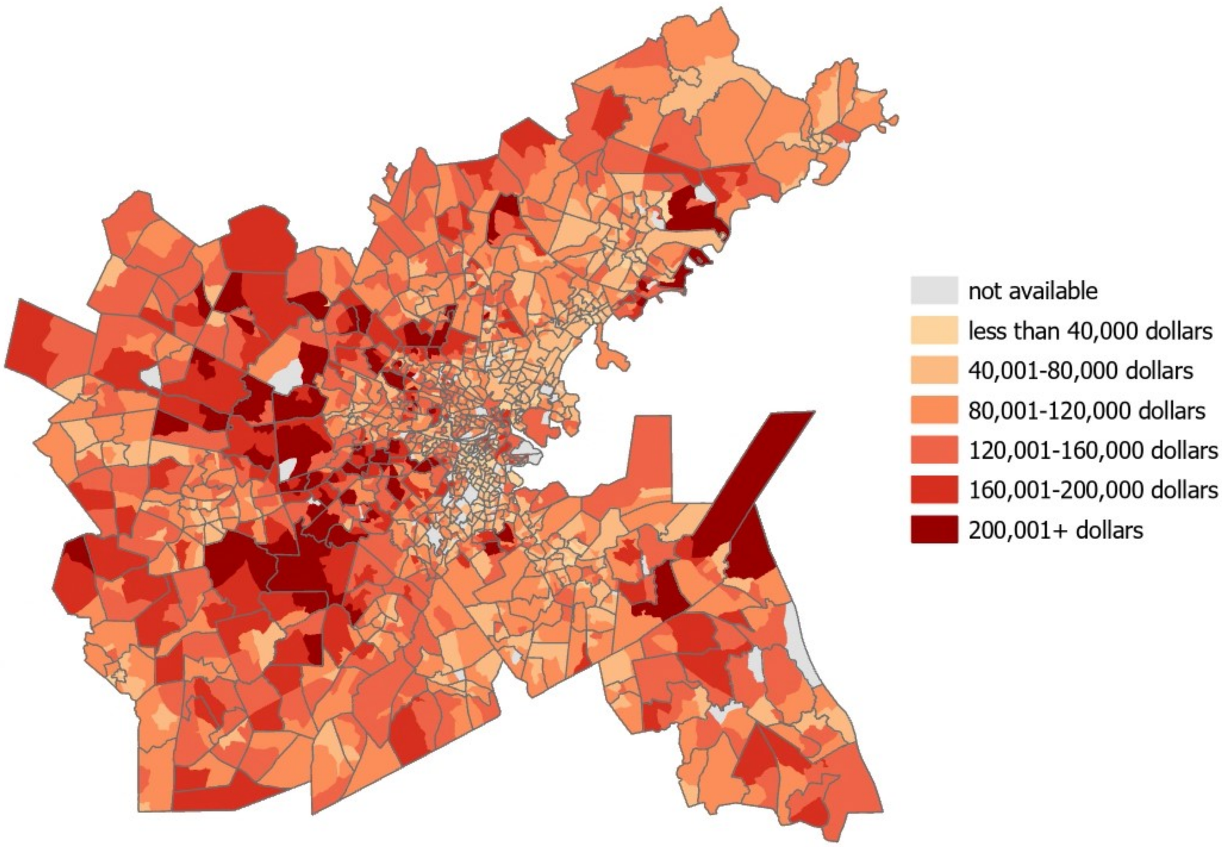
This graph illustrates the areas located within 20 kilometers to the city center. The city center is set at the location of Boston City Hall.

Figure B.3: Aggregate resident income per square foot by census block group



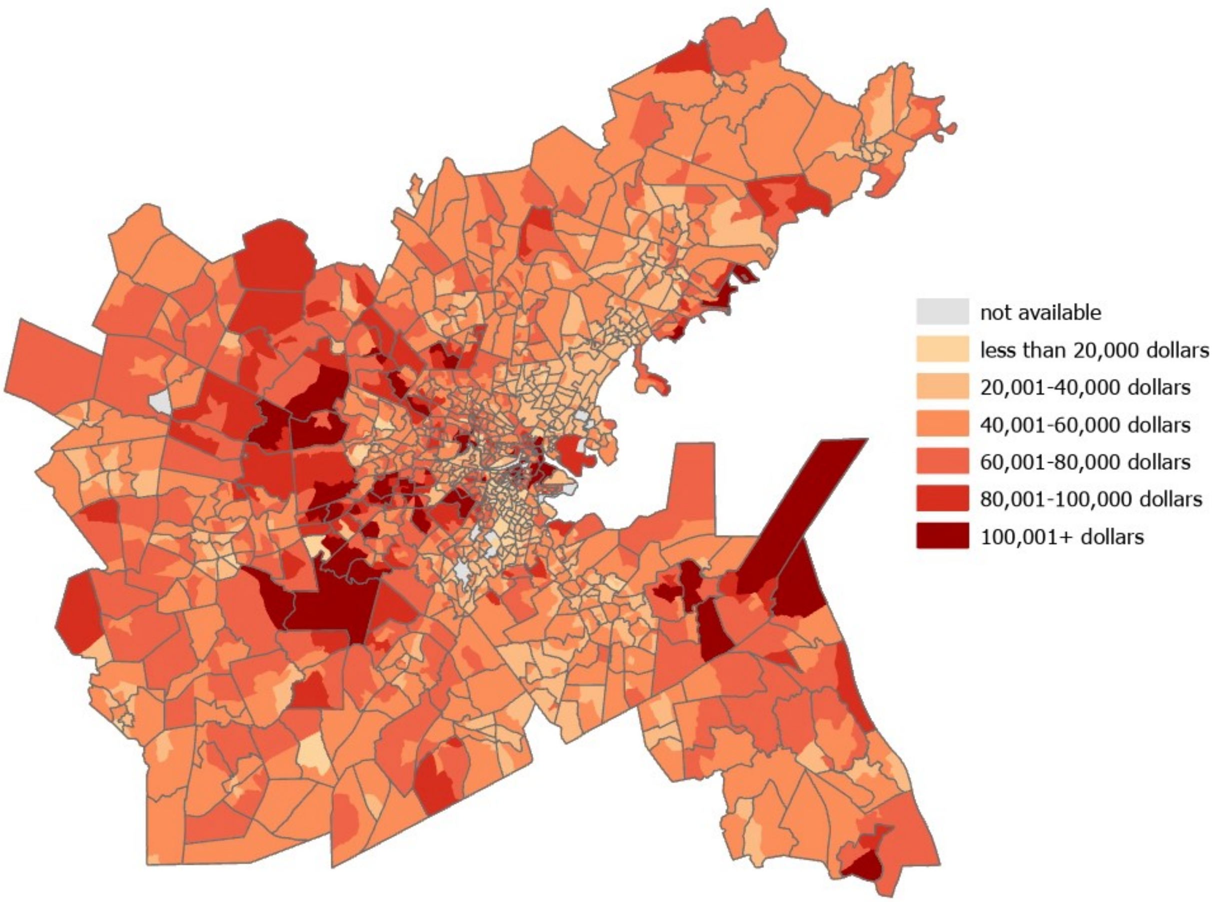
This graph presents aggregate resident income normalized by total square footage of residential land at the block group level. The income data comes from the 2015-2019 American Community Survey 5-Year Data, and the property data comes from CoreLogic.

Figure B.4: Median household income by census block group



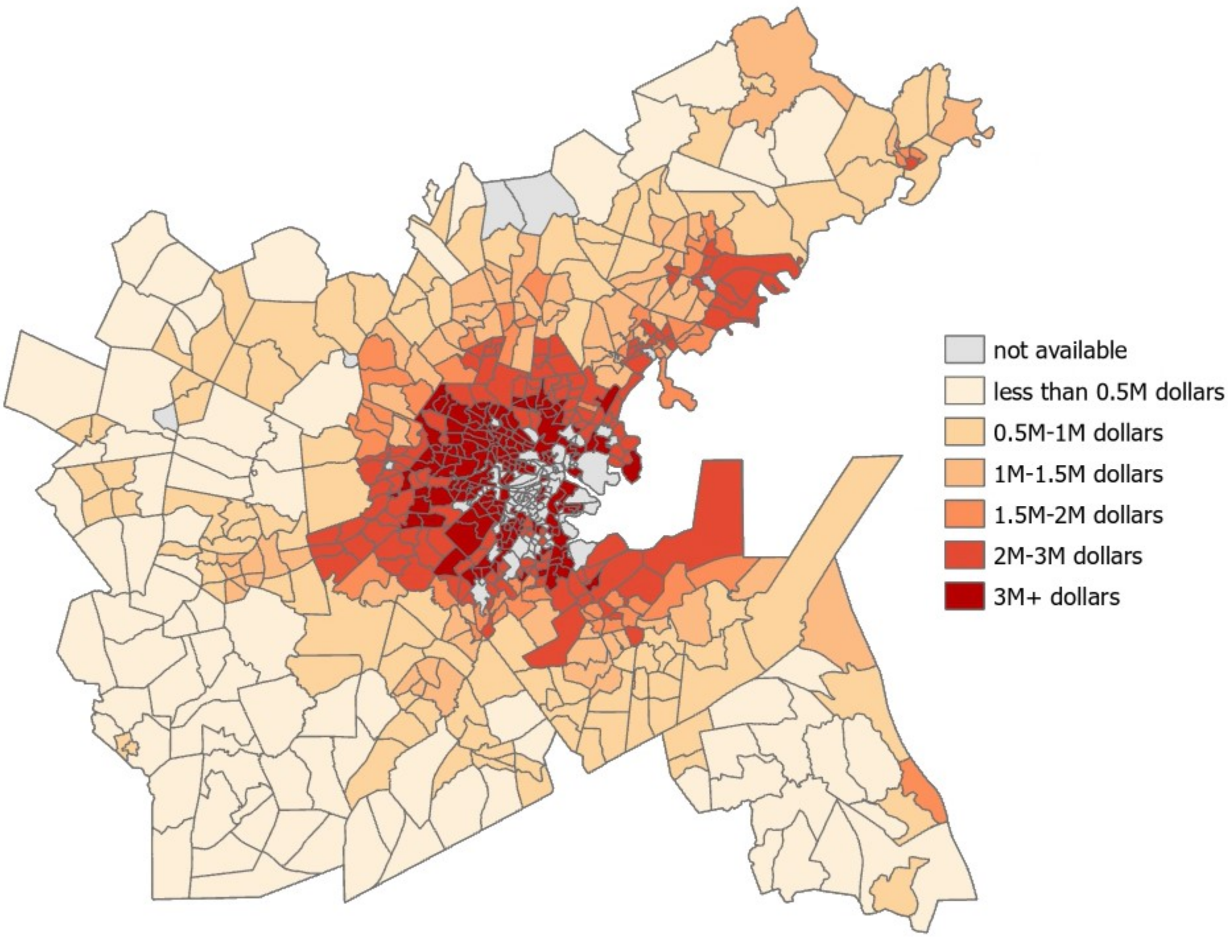
This graph displays median household income at the block group level, sourced from the 2015-2019 American Community Survey 5-Year Data, with delineations of census tract boundaries.

Figure B.5: Per capita income by census block group



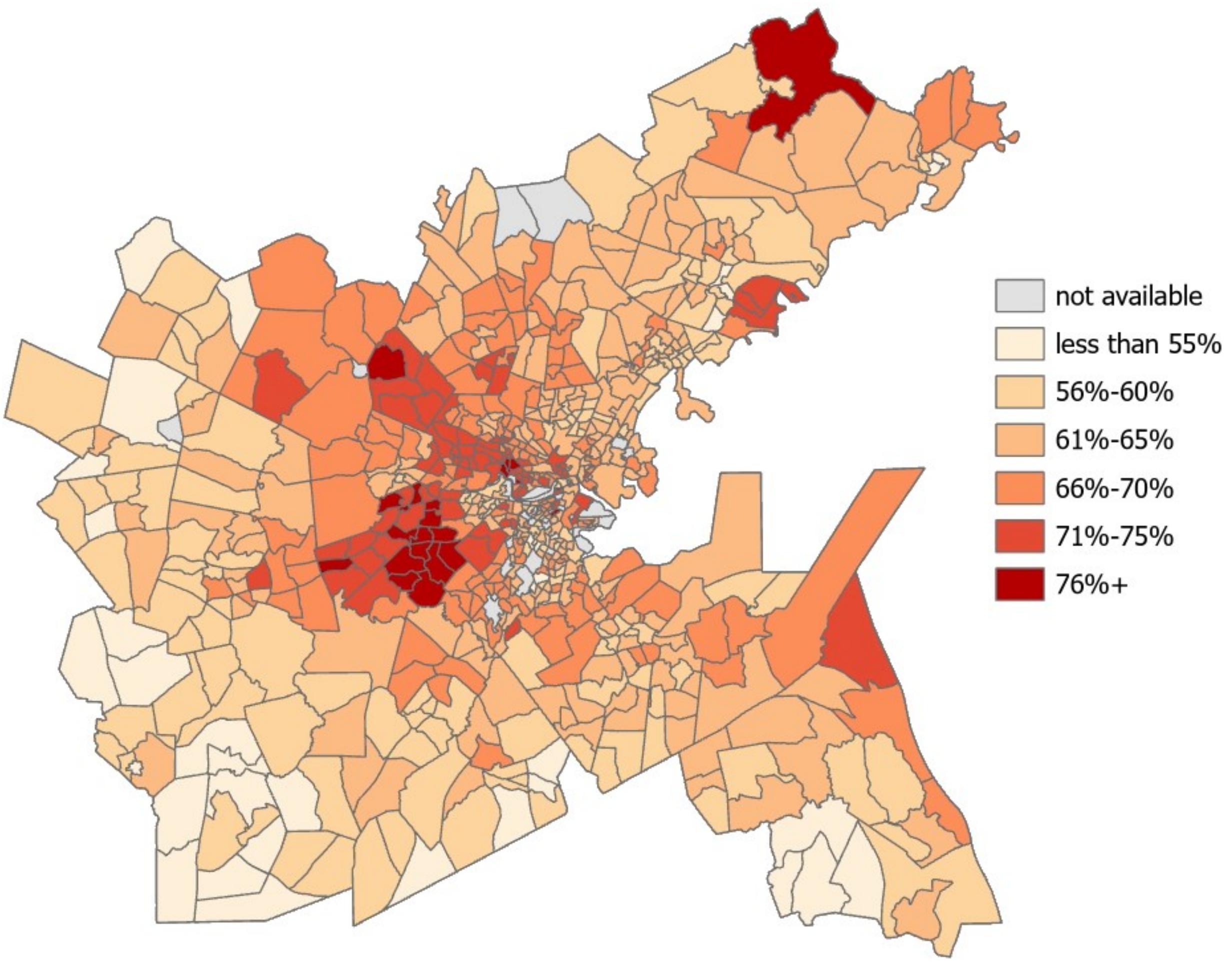
This graph displays per capita income at the block group level, sourced from the 2015-2019 American Community Survey 5-Year Data, with delineations of census tract boundaries.

Figure B.6: Land cost per acre



This graph displays per acre residential land cost at the tract level, sourced from the 2019 AEI Land Price and Land Share Indicators.

Figure B.7: Average land cost share for single-family residences



This graph displays the average land's share of house value for single-family residences at the tract level, sourced from the 2019 AEI Land Price and Land Share Indicators.

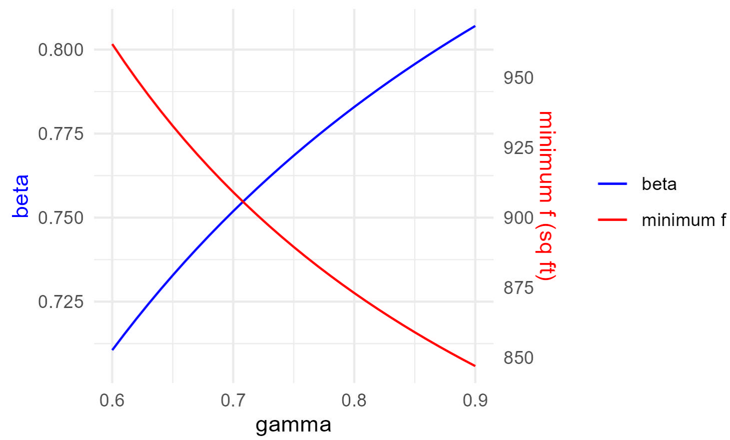
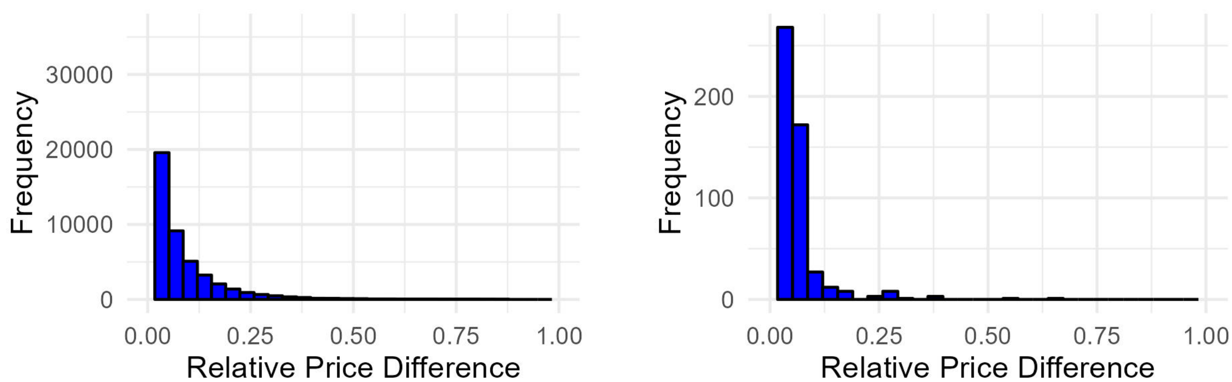


Figure B.8: Parameters selections

This graph depicts the possible set of parameters when $\gamma(1-\beta) = 0.1736612$, $\frac{f\beta}{1-\gamma(1-\beta)} = 827.2036$, and $0.6 \leq \gamma \leq 0.9$.



(a) Housing Unit

(b) Tract

Figure B.9: Difference Between Estimated Cost and Optimal Cost

The sample limited to post-1950 housing units.

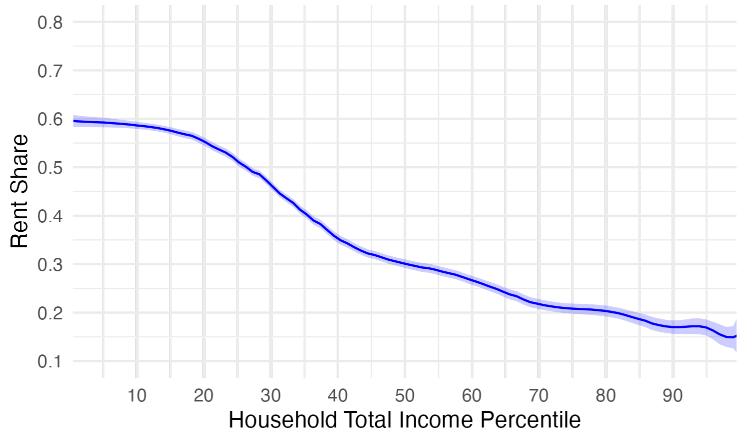


Figure B.10: Housing Rent Share

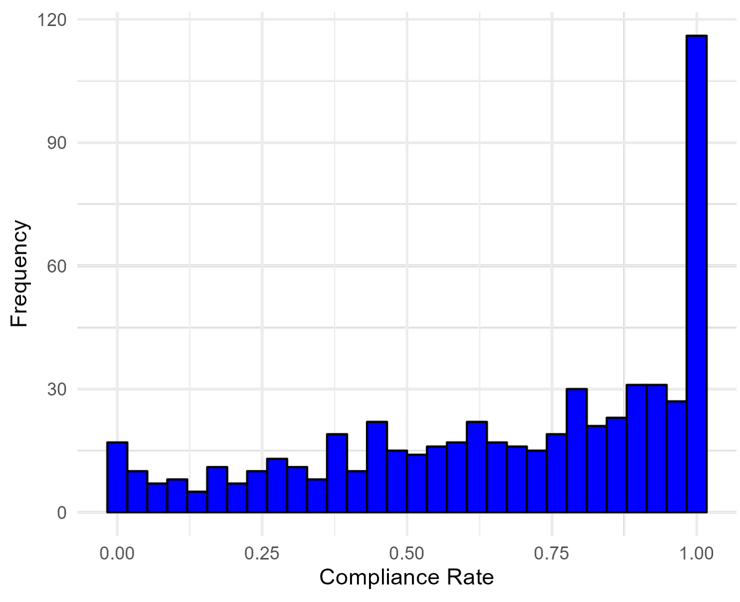


Figure B.11: Tract-level Compliance Rate Distribution

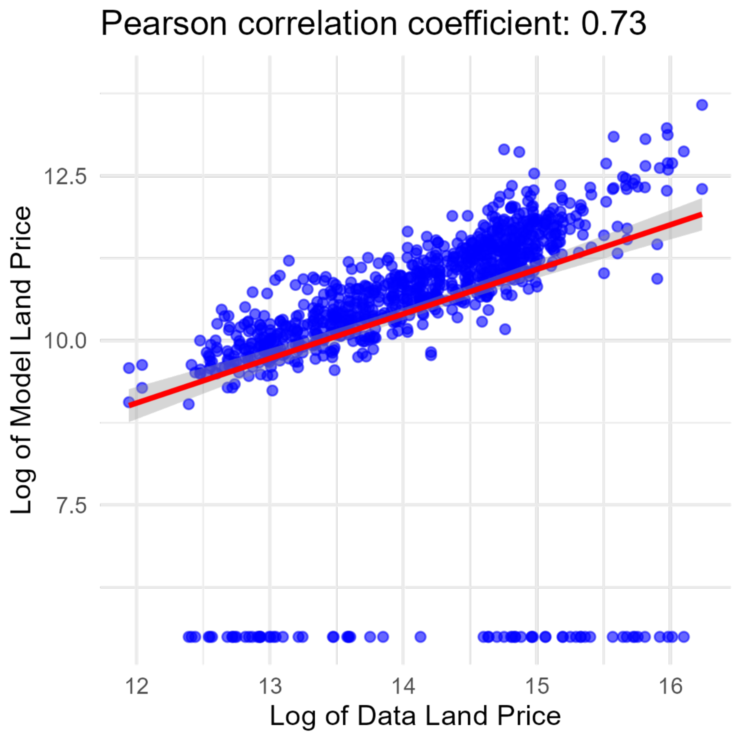


Figure B.12: Tract-level Land Price: Model v.s. Data

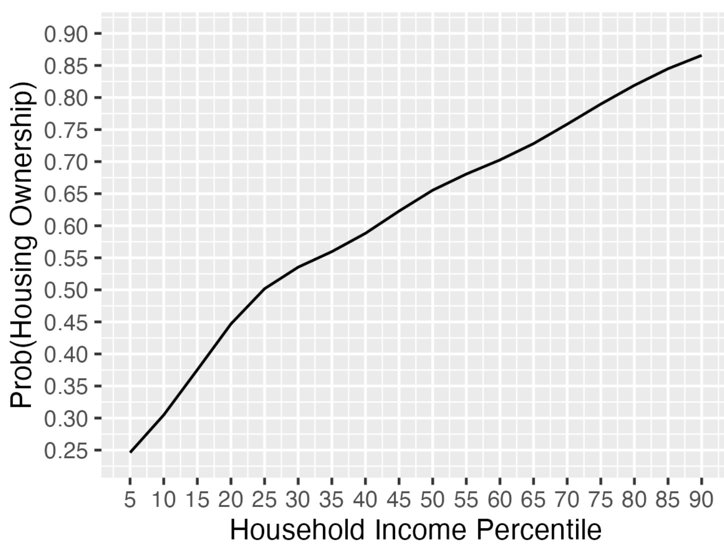


Figure B.13: Probability of Housing Ownership and Household Income Percentile

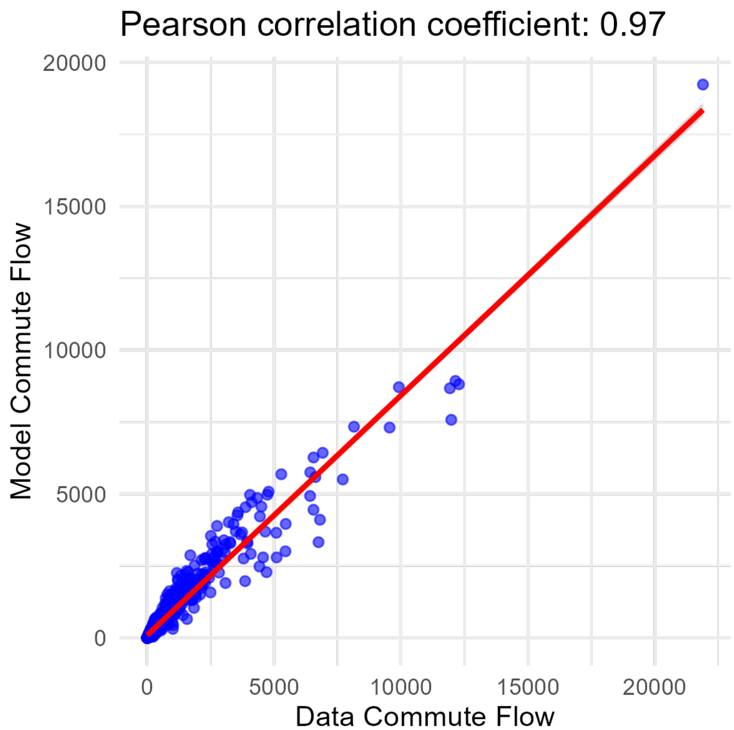


Figure B.14: Puma-level Commute Flow: Model v.s. Data

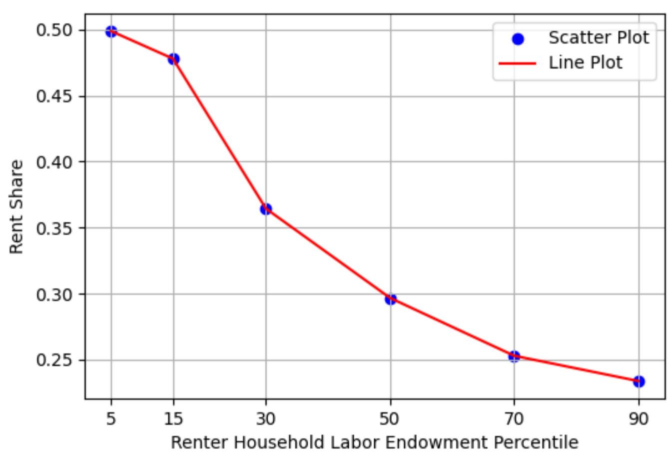


Figure B.15: Model Housing Rent Share

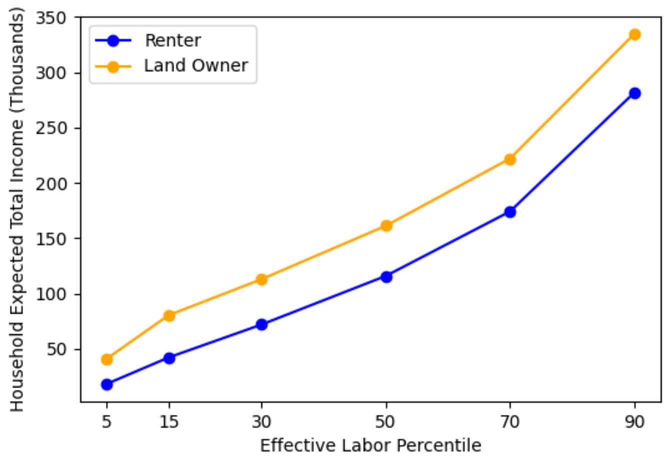


Figure B.16: Model Household Total Income Across Effective Labor Percentile

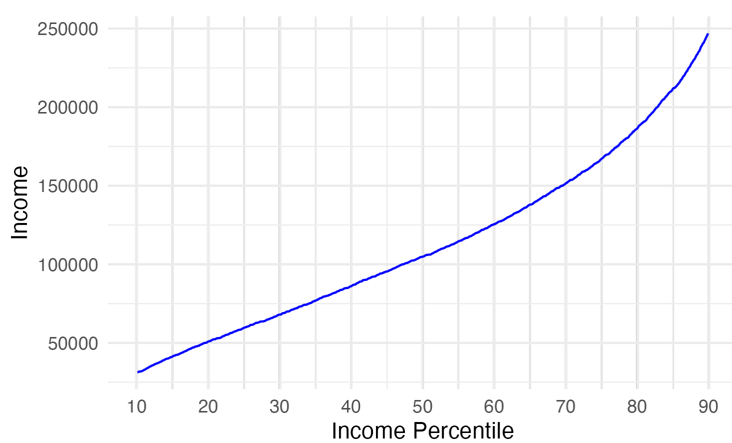


Figure B.17: Data Household Income Percentile

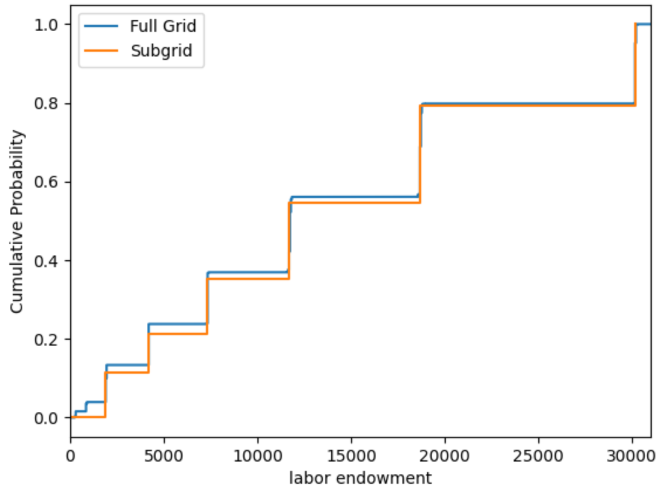


Figure B.18: Estimated Effective Labor Distribution

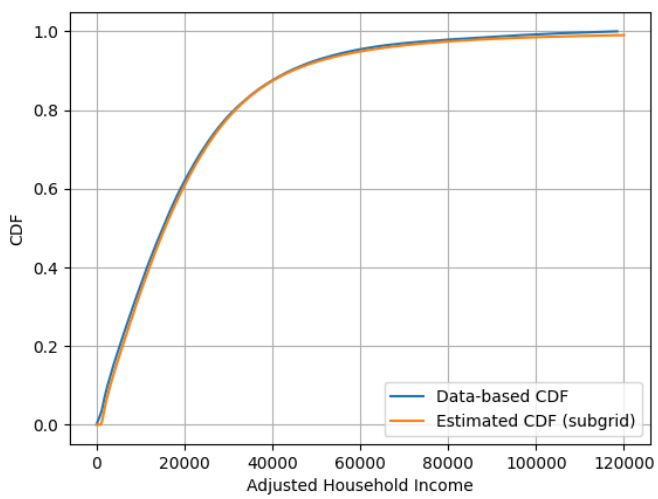


Figure B.19: Adjusted Income Distribution

C Additional tables

$\log(p^H)$	(1)	(2)
$\log(\hat{p^{H,con}})$	0.830***	0.804***
	(0.004)	(0.004)
Constant	4.813***	
	(0.0548)	
Observations	20054	20009
Proj Model R^2	0.619	0.632
Fixed Effects		sale_year

Table C.1: Data Price and Estimated Cost Correlation

The sample limited to post-1950 housing units. (Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)

$\ln(\bar{p^H})$	(1)
$\ln(y^L)$	0.221***
	(0.004)
Constant	10.155***
	(0.046)
Observations	51102
Model R^2	0.055

Table C.2: Labor Income Elasticity of Housing Value

(Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)