

The Internal Geography of America's Housing Crisis *

November 20, 2025

Álvaro Sánchez Leache[†]

CEMFI

Job Market Paper

[Please click here for the latest version]

Abstract

Why are cities in America struggling to supply housing at affordable prices? The conventional view is that housing regulations restrict the expansion of big, high-productivity cities. I document how shifting demand *within* these cities towards costly margins of urban growth — infill and redevelopment, as opposed to sprawl— has been central to rising unaffordability since at least 1990. I measure this by combining satellite imagery, digitized building footprints and Census data to track development and redevelopment at 30x30 meter resolution nationwide. Big cities increasingly relied on costly infill and redevelopment as commuting speed stagnation and the revival of urban amenities concentrated demand in already-urbanized areas. A quantitative spatial model reveals that relaxing zoning regulations to small-city levels —mostly easing suburban expansion— would only modestly increase big-city populations, since they would still be bound by the costs of redevelopment in dense areas where demand concentrates.

*I am very grateful to Diego Puga for his guidance and support, and to Manuel Arellano and Gilles Duranton for their valuable advice and feedback. I thank Dmitry Arkhangelsky, Daniel Fernández, Cécile Gaubert, Dalia Ghanem, Sevin Kaytan, Antoine Levy, Marianna Magagnoli, Jorge Martín, Tim McQuade, Jacob Moore, Enrico Moretti, Tai Nguyen, Giorgio Pietrabissa, Jorge de la Roca, Cristian Navarro, Nick Tsivanidis and seminar participants at CEMFI, UC Berkeley, at the 2024 Symposium of the Spanish Economic Association for helpful comments and suggestions. I gratefully acknowledge funding from the María de Maeztu Unit of Excellence CEMFI MDM-2016-0684 and CEX2020-001104-M.

[†]Centro de Estudios Monetarios y Financieros (CEMFI). Email: alvaro.sanchez@cemfi.edu.es

1 Introduction

Over the past several decades, housing prices have exploded in the United States's largest and most productive cities, sparking an affordability crisis. For many years, the literature has argued this is the result of city-wide constraints to the supply of urban land, namely a combination of regulatory and geographic barriers to development¹. According to this view, the superstar metro areas that drive much of the country's productivity growth simply have too many regulatory barriers to the supply of urban land, thus converting most of their productivity gains into higher housing prices and constraining the country's economy as a whole. However, this focus on aggregate constraints obscures the internal geography of demand and supply, and its importance for the evolution of prices.

Cities don't just expand into new land, they also infill empty gaps and redevelop existing structures depending on household demand. Housing prices reflect the relative demand for different areas in the city, because infill and redevelopment are inherently costlier than greenfield expansion. Empirically, areas where land is already occupied by urban structures have low supply elasticities regardless of regulations (Baum-Snow and Han, 2024). This reflects the high fixed costs of redevelopment: the process of purchasing old buildings, demolition, and constructing larger structures that will yield higher rents constrains how much this margin can respond (Rollet, 2025). When demand concentrates in parts of the city that are already urbanized, these high fixed costs can make housing very unaffordable even without changes in regulations or a shortage of developable land in the fringe. Moreover, each new wave of infill and redevelopment in these areas makes subsequent local development more costly, compounding the effect.

This paper uses an unprecedented combination of detailed spatial data over the 1990-2020 period to document how those two forces — the concentration of demand in already urbanized areas of cities, and the increasingly constrained supply resulting from further infill and redevelopment — have become an increasingly important driver of housing supply constraints in the United States. Demand concentration partly reflected decades of stagnant commuting speeds (Couture et al., 2018), which raised the cost of further suburban expansion, and the revival of central-city amenities (Couture et al., 2024; Couture and Handbury, 2020). Because this concentration of demand was unique to big cities, it made them increasingly unaffordable compared to smaller cities, significantly amplifying the effect that regulatory differences would have had on their own.

¹Some notable examples include Duranton and Puga (2023), Ganong and Shoag (2017), Glaeser and Gyourko (2025), Glaeser et al. (2005), and Hsieh and Moretti (2019)

The first contribution in this paper is providing a comprehensive account of urban growth across the different margins of development. To do so, I combine gridded land cover data from the Annual NLCD (U.S. Geological Survey, 2025) with digitized footprints for every building in the United States and housing data geocoded to Census Block Groups. The result is a panel dataset that tracks the evolution of urban structures, building footprints and housing units at 30x30 meter resolution from 1990 to 2020. Thanks to this dataset, I can estimate for the first time in the literature how much of the new housing supply that was constructed in each decade was provided by extending the fringes of cities, or by infill and redevelopment in already urbanized neighborhoods².

With the new dataset in hand, I document three key empirical facts about urban development and housing costs since 1990. First, larger cities already relied more heavily on the costlier intensive margins in the 1990s, housing more of their population in dense neighborhoods and building new units disproportionately through infill and redevelopment. Second, this reliance intensified dramatically over three decades: big cities turned increasingly to infill and redevelopment while small cities continued sprawling, creating a sharp divergence in development patterns. Third, housing cost divergence mirrored this spatial reallocation—all of the price divergence across metros occurred in big cities’ downtowns and dense residential areas, while fringe areas showed no divergence. The post-1990 big-city premium was really a rising premium for the densest neighborhoods within those cities.

To understand these results, I develop a conceptual framework that relates the demand for housing within metro areas to sprawl, infill, redevelopment, and the evolution of housing prices. I build a model combining elements from the monocentric Alonso-Muth-Mills model (Alonso, 1964; Mills, 1967; Muth, 1969) with quantitative spatial models (Redding and Rossi-Hansberg, 2017), where the three margins of development happen endogenously in response to demand. I derive conditions for the endogenous boundaries at which each margin of development is optimal, and for how changes in the relative attractiveness of the *Downtown*, *Residential* and *Fringe* areas of the city affect prices and quantities in equilibrium.

The model predicts two reasons why demand would concentrate in already-developed places. Because transportation is costly, cities with larger populations extend further out from their center, but also feature a larger share of housing stock built at the costlier *Residential* and *Downtown* densities—reflecting greater capitalized commuter savings in central locations. When they build

²Baum-Snow and Han (2024) used data for about half of the Census Tracts in the nation in the 2000-2010 period and documented how redevelopment makes up a sizeable share of new housing supply, and how Census Tracts that require redevelopment are more supply-inelastic. However, they did not document how the reliance on this margin varies across cities or how it has changed over time.

new housing, a larger fraction is built through infill and redevelopment, so they are more supply-inelastic. If transportation speeds stagnate, cities redevelop more and more and become supply-inelastic as they grow. This is further amplified if residential amenities become more attractive at the urban core: Infill and redevelopment become a larger share of new construction, at the cost of disproportionate price growth in these areas and an overall decline in the city-wide supply elasticity.

I calibrate the model to the U.S. urban system, keeping into account the role of regulations and geographic constraints. Regulations make the same margin of development—for example, fringe expansion—costlier in some cities than others, and I incorporate them through the housing supply elasticity estimates from Baum-Snow and Han (2024). Geographic constraints are modeled as in Duranton and Puga (2023) and Saiz (2010): mountains and bodies of water force developers to build *around* them, so the city has to extend further out to obtain the same amount of developable land. On the demand side, the evolution of wages, commuting costs and housing costs identifies the residential amenity gradient of each city. In line with findings by Baum-Snow and Hartley (2020), Couture et al. (2024), Couture and Handbury (2020), Diamond (2016), and Moreno-Maldonado and Santamaría (2025), my model indicates that central city amenities increased substantially in big cities, reflecting how urban consumption is an increasingly important driver of residential choice³.

I first use the model to quantify how the *additional* demand concentration since 1990 contributed to rising housing costs. To do so, I decompose the changes in demand determinants into within-city and across-city changes, by adjusting values of wages, commuting speeds, and amenities and simulating counterfactual equilibria. The additional within-city concentration of demand—within-city changes in amenities, and the growth in commuting times—contributed to 23% of the growth in housing expenditures⁴ between 1990 and 2020. Changes in wages across cities also contributed to rising housing costs, by making big, supply-inelastic cities more attractive, contributing to 14% of the growth in housing expenditures. Overall, the combination of all of these changes in demand contributed to 35% of the growth in housing expenditures between 1990 and 2020⁵.

³I measure amenities as an unobserved fundamental that is recovered from the location choices of households, as is standard in the urban economics literature. The literature on U.S. *urban revival* has identified this revival of urban amenities as an equilibrium outcome from young, high-income households increasing their presence downtown and creating demand for consumption amenities—restaurants, shops, entertainment—, which in turn attracts more young, high-income households. Some of the deeper structural drivers of this return are the growth in income inequality, which allows high-income households to out-bid low-income households downtown (Couture et al., 2024), and the delayed child-bearing by high-income households, which makes urban amenities more attractive to them (Moreno-Maldonado and Santamaría, 2025). In this study, I take the amenities that I recover from the data as exogenous fundamentals that drive the *aggregate* location choices of households, and I interpret them as being the outcome of this process.

⁴Measured as the the average monthly rent in the city divided by the city’s average monthly wage

⁵Note that because this decomposition is produced by computing alternative equilibria, the decomposition is not necessarily linear

Finally, I use the model to quantify what would be the effect of zoning relaxation, taking into account the internal geography of demand and supply. I simulate a counterfactual economy where I relax housing regulations in seven large, highly regulated cities⁶ to the levels of the median U.S. city. I also run the same counterfactual in a *naïve* model that treats cities as homogeneous units. The results indicate that, while zoning relaxation *would* allow more population into big cities at lower prices, the effect is very muted compared to previous claims in the literature or my own naive model. Relaxing zoning regulations in these seven cities would have increased their 1990-2020 population growth rate by 12%, allowing them to host 2.7% more population by 2020, much less than the increase predicted by the naive model and by extant literature like Duranton and Puga (2023) and Hsieh and Moretti (2019). The reason is that, even with relaxed regulations that would mostly facilitate fringe expansion, big cities would still face very intense demand for already-urbanized areas where the supply constraints are not regulatory in nature.

To conclude, the results in this paper illustrate how the internal geography of demand and supply within cities is key to understanding how much they can grow at affordable prices. Big cities in the United States increasingly struggle to provide affordable housing, not just because of regulations, but because the parts of these cities that attract households are simply too heavily developed to provide more housing easily. As big cities concentrate a growing fraction of productivity growth in advanced economies (Moretti, 2012), this poses a fundamental challenge. While easing regulations would help somewhat, policymakers may need to consider complementary approaches that reduce the spatial concentration of demand itself. Technologies like remote work show early promise in this direction (Delventhal and Parkhomenko, 2024; Delventhal et al., 2022), as does strengthening alternatives to superstar cities through targeted infrastructure investment or place-based policies.

1.1 Related literature

This paper contributes to several strands of literature. First, it contributes to the literature on housing constraints. Part of this literature has focused on measuring housing constraints at the city or microgeographic level (Baum-Snow and Han, 2024; Gyourko et al., 2008; Rollet, 2025; Saiz, 2010), while another strand has focused on measuring the *consequences* of these constraints for housing prices, migration to big cities, or aggregate productivity (Duranton and Puga, 2023; Ganong and Shoag, 2017; Glaeser et al., 2005; Hsieh and Moretti, 2019). This paper showcases how the internal demand within cities is a first order driver of prices due to the different costs of each margin of urban growth, and how this channel has become increasingly important for

⁶New York, Los Angeles, Washington DC, San Francisco, Boston, Seattle, San Diego

explaining housing unaffordability in the U.S.

This paper also contributes to the study of the tradeoff of city size. Many works like Glaeser et al. (1992) have studied the *benefits* of big cities in terms of productivity, while more recent work like Combes et al. (2019) and Duranton and Puga (2023) has attempted to quantify the *costs* of agglomeration and thus find the optimal tradeoff. This paper provides evidence for a previously underappreciated cost of agglomeration: as cities grow, they become more dependent on costlier margins of development, which makes housing supply more inelastic and prices more sensitive to demand shocks.

A third strand of the literature has studied the internal structure of cities. Starting with the monocentric Alonso-Muth-Mills model (Alonso, 1964; Mills, 1967; Muth, 1969) up to the recent wave of tractable quantitative spatial models (Ahlfeldt et al., 2015; Heblich et al., 2020; Tsivanidis, 2023), this literature has studied how transportation costs, amenities, and productivity shape where households and firms locate within cities. This paper contributes by documenting how cities adapt their internal structure through different development margins—greenfield expansion, infill, and redevelopment—and by quantifying the rising frictions that constrain this adaptation.

More particular to the U.S. context, I contribute to the literature on urban decline and revival in the United States. A large literature has documented the process of central city decline and suburbanization that U.S. cities experienced after World War 2 (Baum-Snow, 2007; Boustan, 2010), followed by a revival of central cities since the 1990s (Baum-Snow, 2020; Couture et al., 2024; Couture and Handbury, 2020). While the literature had recognized the revival of demand for central cities, it assumed that the *overall* process of suburbanization continued apace. Thanks to the novel use of far more detailed data, I show that, in fact, the process of U.S. suburbanization has been slowing down substantially since at least the 1990s.

Finally, I contribute to the use of remotely sensed and geospatial data in economics. Remotely sensed data has proven useful in many settings where administrative data lacks the necessary resolution or does not exist at all (Donaldson and Storeygard, 2016). In urban economics, Burchfield et al. (2006) pioneered the use of satellite imagery to study aspects of urban growth, followed by other prominent examples like Baum-Snow and Han (2024), Harari (2020), and Saiz (2010). With the construction of my dataset, I demonstrate how the careful combination of remotely sensed and administrative data can overcome the limitations of each individual source and produce data that is more than the sum of its parts. This dataset can prove useful for many other applications in the U.S. setting, and the same methodology can be applied to many other settings where similar

data sources exist.

The rest of the paper is organized as follows. In section 2, I describe the construction of the dataset and other choices regarding the measurement of urbanized areas. In section 3, I present and rationalize the main empirical findings. Section 4 presents the quantitative model and its calibration to the U.S. urban system. Section 5 presents results from the calibrated model and the counterfactual experiments. Section 6 concludes.

2 A dataset to track all margins of urban growth

Distinguishing sprawl from infill and redevelopment requires data that tracks both where cities expand and how intensively they develop existing areas. To achieve comprehensive geographic coverage over three decades, I turn to remotely sensed data. I construct a 30×30 meter gridded panel from 1990 to 2020 by combining the recently released Annual NLCD—which for the first time provides land cover data back to 1990—with detailed building footprints and Census housing counts. I combine these sources to isolate building footprints from other artificial structures and to allocate housing units within Census Block Groups to individual grid cells. From this gridded foundation, I then apply a density-based statistical method to delineate the evolving physical boundaries of urban areas, producing dynamic measures of sprawl, infill, and redevelopment that are not constrained by fixed administrative boundaries.

2.1 Constructing the gridded dataset

The main dataset I construct consists of a series of gridded maps, each containing 9 billion 30×30 meter cells that track the number of housing units, the fraction of each cell covered by buildings, the fraction of each cell covered by paved surfaces, and geographic features such as the cell’s elevation or the presence of water bodies. As explained in the introduction, I do this by carefully merging a series of geospatial datasets from the U.S. Geological Survey, Overture Maps, and the Census, each containing complementary pieces of information.

The point of departure for the data construction is the new generation of the National Land Cover Database (NLCD) (U.S. Geological Survey, 2025). The NLCD is a fine-resolution (30×30 meter cells) dataset that describes land cover and land use for the entirety of the United States. One layer of this dataset classifies these cells in major land use and cover categories (the *Land Cover* layer), while another classifies cells by the *percentage* of their surface that is covered by buildings *and / or* paved surfaces (the *percentage imperviousness* layer). Previous versions of the NLCD have been used in prominent papers in economics (Baum-Snow and Han, 2024; Burchfield

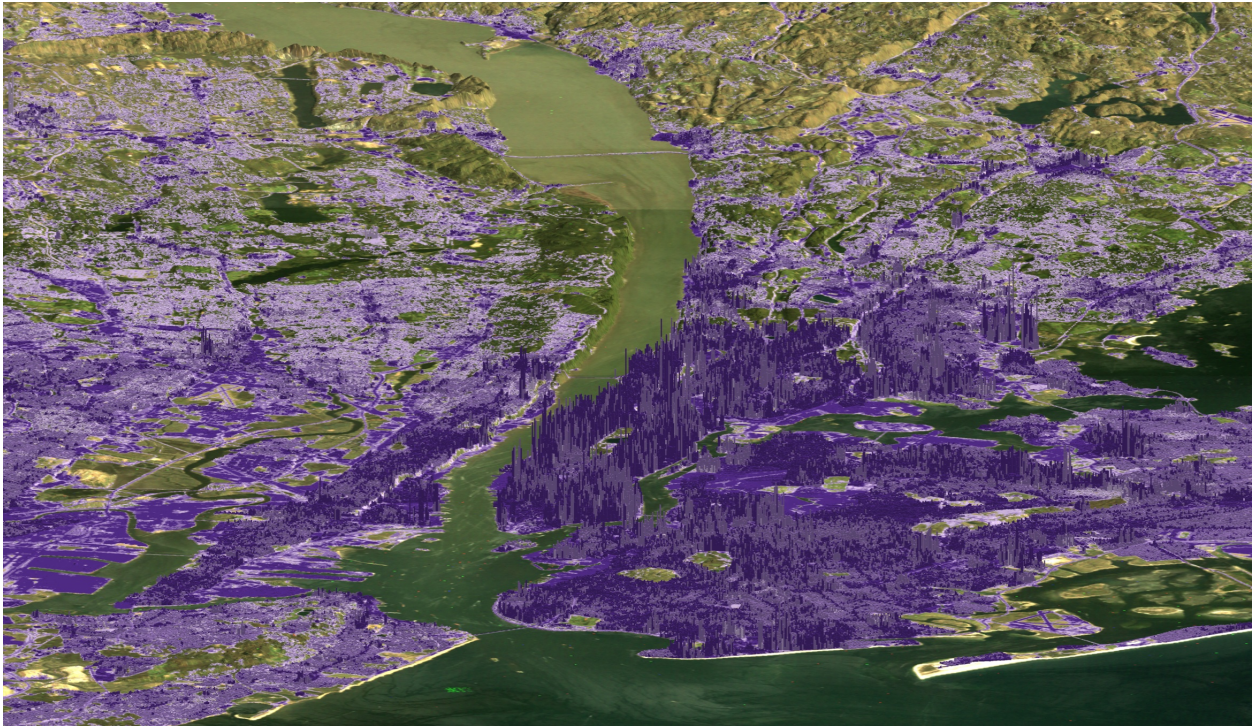
et al., 2006; Saiz, 2010), but until very recently the coverage of the NLCD panel only went back to 2001⁷ (Yang et al., 2018). The new release of the NLCD updated its methodology and expanded coverage back to 1985, so that now it is possible to track the evolution of land use for more than three decades.

It is important to briefly explain how the NLCD is built, because its production method determines what it can and cannot identify reliably. The NLCD is produced by the U.S. Geological Survey from the Landsat Analysis Ready Data (ARD) archive (Earth Resources Observation and Science (EROS) Center, 2017), a multi-decadal stack of multispectral imagery at a common 30×30 meter resolution covering the conterminous United States. The latest release exploits the full time series with machine-learning and temporal-classification methods. As a result, it excels at detecting transitions between natural and artificial cover—that is, distinguishing urbanized (impervious) from non-urban land. However, at 30 meters the spectral and temporal signatures of roofs and other impervious surfaces are too similar for the NLCD to reliably separate building footprints from paved surfaces such as roads, parking lots, or port facilities.

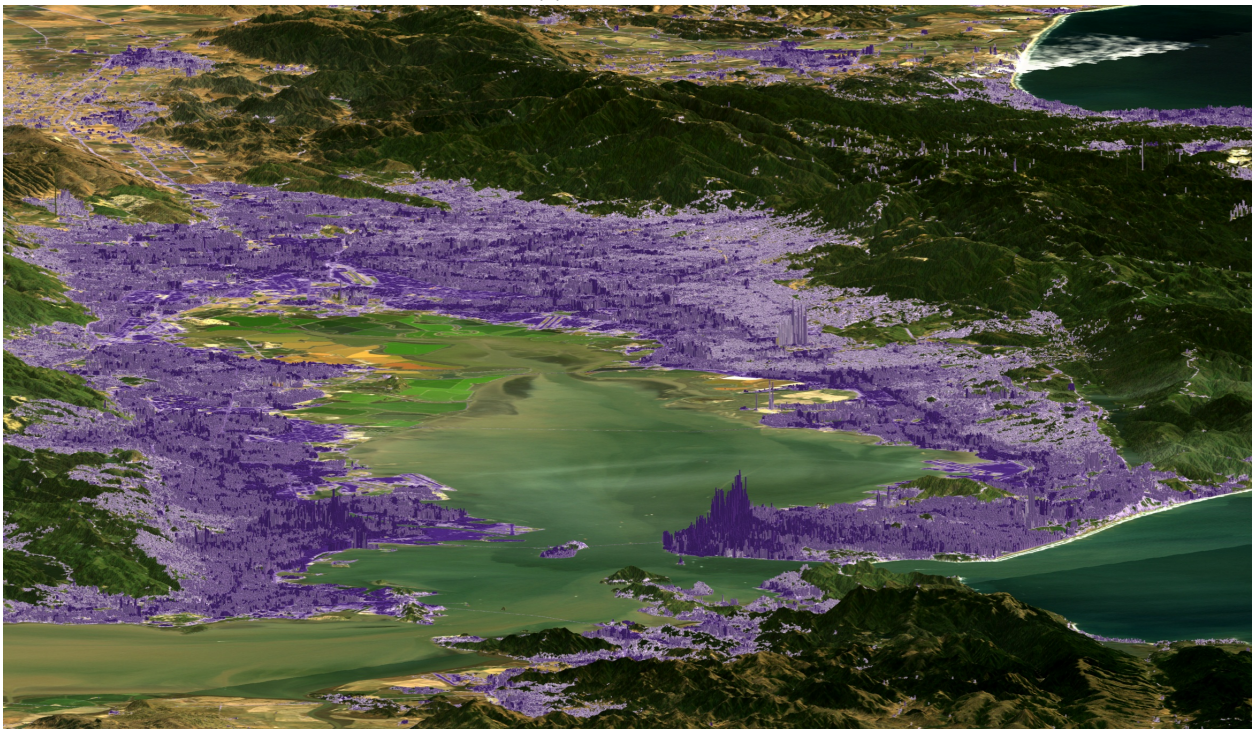
Given this limitation, I construct a panel of *building footprints* by combining the NLCD's time series on imperviousness with a high-coverage cross-sectional map of building polygons circa 2024. Specifically, I use changes over time in the NLCD percentage imperviousness to time the emergence of built cover within each 30×30 meter cell, while using the cross-sectional polygons to identify which impervious areas are buildings. The cross-section comes from Overture Maps and includes footprints for roughly 183 million U.S. buildings, aggregating community-contributed sources (OpenStreetMap, Esri), national and municipal datasets, and machine-learning roofprints from Google and Microsoft. This approach yields a temporally resolved building layer that I use both to delineate urbanized areas (Section 2.2) and to allocate housing units within block groups.

The way I trace the evolution of building footprints builds on two simple ideas. The first is that the percentage of a cell that is covered by artificial surfaces that is *not* covered by buildings is covered by paved surfaces, such as roads or parking lots. The second is that, once a road segment is laid out, it is very unlikely to be destroyed (Barrington-Leigh and Millard-Ball, 2015). For the year 2023, I can obtain the percentage of a 30×30 meter cell that is covered by asphalt by subtracting the percentage of the cell that is covered by buildings from the total area of the cell that is covered by artificial surfaces (*impervious* layer). Then, I trace changes in the asphalt content of each cell backwards in time, using the *percentage imperviousness* layer for each year. If

⁷The original release of the NLCD (Vogelmann et al., 2001) was representative of 1992, but it was a one-off release that was not comparable to later versions of the dataset.



(a) New York MSA



(b) San Francisco MSA

Figure 1: An artistic rendition of the data for the New York and San Francisco Metropolitan Statistical Areas (Panels (a) and (b) respectively). Each panel represents the state of each MSA as they were in 1990. The base color layer is composited from 1990-vintage Landsat satellite imagery. The lilac overlay represents stands for built-up or paved pixels, with darker colors representing more intensity. The vertical “towers” represent the number of housing units in each 30x30 meter cell in 1990, with taller elevation representing more housing units. Terrain elevation is exxaggerated for visual effect.

we assume that asphalt does not get destroyed, then, $\forall t \in \{1990, 2000, 2010, 2020\}$:

$$footprints_t = \max\{impervious_t - roads_{2023}, 0\} \cdot \mathbb{1}(footprints_{2023} > 0)$$

Once this work is done, I have a gridded panel which isolates the percentage of each cell that is covered by buildings. The next step is to use these buildings to pinpoint the location of housing units. I obtain tabulations of housing unit counts from the Census at the level of 2010-vintage Census Block Groups, as well as boundary polygons for those Block Groups from the NHGIS⁸ (Manson et al., 2023). Knowing the number of housing units within the geometric boundaries of the Block Group, and the extent of building footprints for every 30x30 meter cell within those boundaries, I attribute housing units to each cell proportionally to its building footprint.

The resulting dataset is a panel with four layers for each decadal year between 1990 and 2020: land cover categories, percentage imperviousness (percentage of the cell that is covered by buildings or paved surfaces), building footprints (percentage of the cell covered by buildings) and housing units (count of housing units attributed to the cell). Additionally, I download and reproject the National Elevation Map to the same dimensions as the rest of the data, and I compute measures of slope and terrain ruggedness. Figure 1 showcases the merged dataset visually zoomed in to the New York and San Francisco metro areas as they were in 1990.

It is important to note that, while very detailed, the dataset has several layers of imputation and measurement error. At the level of 30x30 meter gridded cells, there will be measurement error coming from the satellite imagery and for each of the imputations I need to perform. However, each of its layers is *at least* as precise as the housing unit tabulations by Census Block group, and being gridded, not tethered to arbitrary boundaries, allows more flexibility for the next steps of the analysis.

2.2 Delineating the growth of urban areas

The next step in the data construction is to trace the *boundaries* of urban areas, that is, to take a stand on what was the extent of their physical footprint at every decade during my period of study. The conventional approach of using fixed administrative boundaries, such as the county-based delineations of Metropolitan Statistical Areas (MSAs), is not satisfactory for my purposes. These county-based definitions tend to not conform to the actual extent of urbanized land, and in

⁸The Census provides housing unit counts at the level of Census Blocks, which are smaller and more detailed than Block Groups. However, the available geometric representations for the boundaries of 1990-vintage Census Blocks in 1990 are of very poor quality. Therefore, I rely on the tabulations interpolated to 2010-vintage Block Groups produced by the NHGIS, which seek to remedy many of these issues.

fact contain vast amounts of rural land, particularly in the Western United States where counties are very large. More importantly, the boundaries of MSAs are fixed over time, so that as cities grow, the amount of rural land within them declines mechanically. To accurately distinguish between extensive growth (sprawl) and intensive growth (redevelopment), my measurement of the urban extent must itself be dynamic.

To avoid these potential pitfalls, I use a method proposed by de Bellefon et al. (2021), which identifies the physical extent of urbanized land based on a statistical criterion on the density of building footprints. The idea is to compare the density of the observed built environment against a counterfactual “unconditional distribution” of buildings where buildings are randomly re-distributed across the the country. Urban areas are areas where the density of buildings is significantly higher than in this counterfactual distribution. In practice, the counterfactual distribution is constructed by bootstrapping the *true* distribution of buildings.

The method works as follows: The *true* distribution of built-up cells is smoothed spatially using a kernel. Then, I create bootstrapped counterfactuals by randomly re-distributing across the country all the cells that are “buildable” , that is, cells that are not covered by bodies of water, not excessively high, and with a slope below a threshold⁹. These bootstrapped counterfactuals are interpreted as draws of an “unconditional” distribution where the buildings are distributed randomly across all “buildable” locations in the country. I then smooth these counterfactual distributions using the same kernel I used with the true distribution. A cell is classified as *urban* when its excess smoothed density is statistically significant at 95% compared to the counterfactual, unconditional distribution of smoothed building densities. For computational reasons, I do not perform this procedure on the full-resolution 30x30 meter dataset, but rather a coarsened version that is aggregated to 210x210 meter cells.

As highlighted by de Bellefon et al. (2021), this method has the advantage of letting the data speak as to what is the threshold of density that defines urban land. However, sensible choices have to be made regarding the choice of kernel and the set of cells that are considered “buildable”. The choice of kernel is important, if I did not smooth the data at all, I would classify as part of the urban areas only the cells that contained buildings themselves, so the classification would be very fragmented. For instance, I would classify as non-urban many cells that do not contain buildings themselves, but that are surrounded by built-up cells, such as urban parks. On the other hand, if I oversmooth the data, I would erase the differences in density across cells.

⁹The original method creates counterfactuals by re-distributing individual buildings across buildable cells. I cannot do this because I only observed the imputed building footprints in earlier periods of my data. Re-distributing entire cells is also less computationally cumbersome. de Bellefon et al. (2021) show how re-distributing cells produces quantitatively similar results to the full method, so that it is sensible when it is unfeasible to implement the full method.

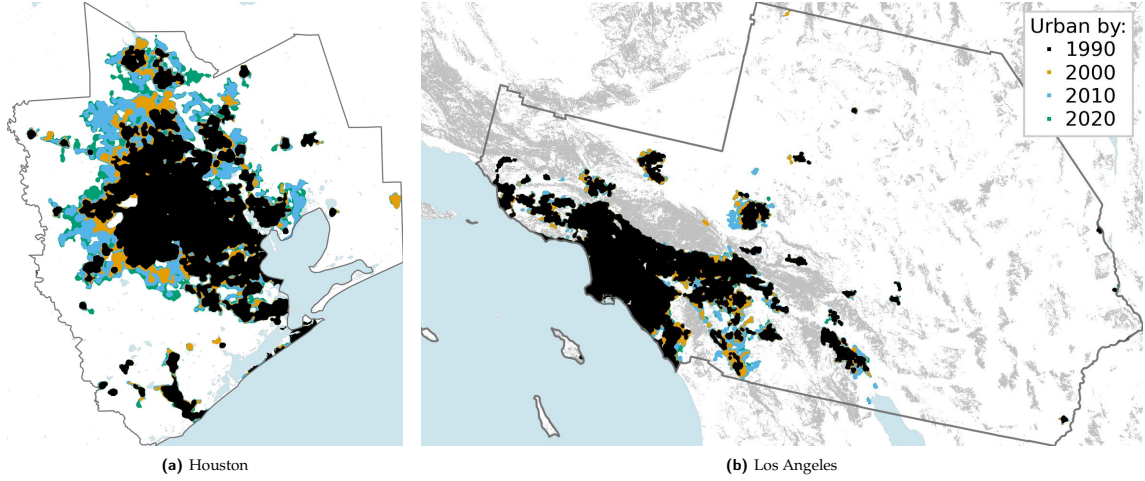


Figure 2: Growth of the urban extent 1990-2020, for Houston (a) and Los Angeles (b) MSAs. The solid shapes show the delineation for each of the years, so that each color after 1990 represents new “urban” land in the metro area. Light gray represents “unbuildable” land (slopes, elevation) and light blue represents water bodies. The gray outline around each map is the boundary of the Metropolitan Statistical Area (MSA). Note how in both metro areas, the amount of new urbanize land in the 2010-2020 decade is very limited compared to the 1990-2000 and 2000-2010 decades.

I define which cells are “buildable” on the basis of their average slope, their elevation, and the presence of permanent bodies of water inside of them. I choose the 99.9th percentile of these values among all cells that contain buildings. “Buildable” cells are thus cells with less than 82% water coverage, less than 18 degrees average slope, and less than 2460 meters of elevation above sea level. Figure A1 in the Appendix has maps of the whole nation outlining these areas. For the choice of smoothing kernel, I follow de Bellefon et al. (2021) in using a bisquare kernel, such that the smoothed density $\hat{z}_j$ of a cell j with coordinates x_j, y_j is given by:

$$\hat{z}_j = \frac{\sum_{i \in N_i} K_h(d_{ij}) z_i}{\sum_{i \in N_i} K_h(d_{ij})}, \quad K_h(d_{ij}) = \left[1 - \left(\frac{d_{ij}^2}{h^2} \right) \right]^2 \text{ if } d_{ij} < h, \text{ and } 0 \text{ otherwise,}$$

where d_{ij} is the cartesian distance between cells i and j , and h is the bandwidth of the kernel. I set the bandwidth to 3000 meters, higher than the 2000 meters used by de Bellefon et al. (2021), to account for the fact that the urban landscape in the Unites States is more sprawling than in France.

I delineate urban areas over all decadal years in my data 1990-2020, drawing the bootstrapped counterfactuals from the *joint* distribution of buildable cells across the four years in the data. Drawing from the joint distribution ensures that the footprint density threshold that defines urban land is the same for all four years of data¹⁰.

Once I have delineated the physical extent of urbanized areas using the density-based method,

¹⁰The alternative would be running the delineation process separately for each year. While this would be appropriate for studying the state of the urban system at very distant points in time, I am concerned with decade-by-decade changes.

I assign them to Metropolitan Statistical Areas (MSAs) for classification and analysis. MSAs are defined by the Census Bureau as groups of counties linked by commuting flows to a central urban core, and they provide a natural way to organize urbanized areas into coherent labor markets. As mentioned before, the MSA boundaries tend to be much larger than the actual boundaries of their urbanized areas, so the urbanized areas I delineate never outgrow their assigned MSA boundaries during my period of study. Figure A10 shows the delineation of urban areas in the conterminous United States in 1990 and 2020, with MSA boundaries plotted as outlines. Note that throughout the paper, whenever I refer to a "metro area," I refer specifically to the areas classified as urbanized within the boundaries of its Census-defined MSA. Also note that my MSA definitions are from the 1999 vintage.

The final step in the data construction is to aggregate the 30x30 meter gridded data into a set of regular hexagonal neighborhoods that serve as the primary geographic units for the descriptive analysis. I create a regular hexagonal grid covering the entire country, with each hexagon having an area of approximately one square kilometer. Hexagons that intersect with the urbanized areas described in section 2.2 are classified as "urban" in a given year. Figure 3 illustrates these neighborhood units using examples from the Chicago metropolitan area, showing how they capture neighborhoods developed at varying intensities—from dense downtown areas to lightly developed fringe zones.

For all statistics reporting where new construction took place, I classify a neighborhood as having experienced new construction in a given decade if it contained at least one cell where the number of housing units increased between the beginning and the end of the decade. This hexagonal aggregation forms the basis of all descriptive results presented in Section 3. For the model calibration in Section 4, I aggregate this data further into the three major neighborhood types $\{Downtown, Residential, Fringe\}$, as described in section 4.4.1.

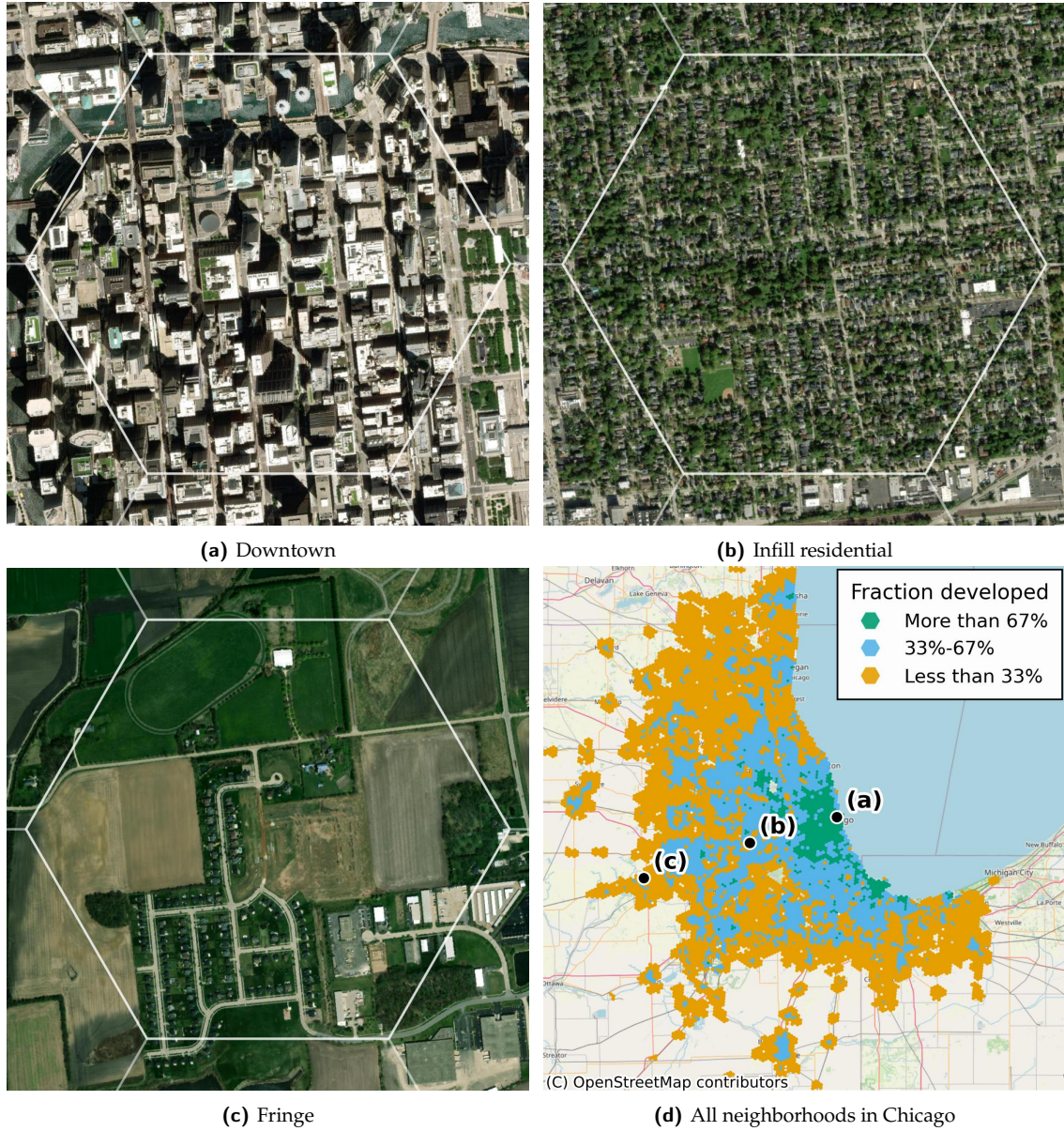


Figure 3: Three examples of neighborhoods developed at different intensities, all in the Chicago MSA. The white hexagon outline in each of the panels (a), (b), (c) represents the outline of one of the neighborhood units which are used for the descriptive analysis. Panel (d) represents all such neighborhoods in the Chicago MSA, classified to discrete bins on the basis of their fraction developed. Panel (a) shows one of the neighborhoods in Chicago's downtown, where 90% of the land is developed. Panel (b) depicts an infill residential neighborhood, with 50% of the land developed. Panel (c) illustrates a fringe neighborhood, where only 15% of the land is developed. The visible light imagery in these maps is sourced from Esri World Imagery.

3 The changing geography of urban growth

I now turn to the main empirical findings of the paper, which can be summarized in three main facts. First, larger cities housed more of their 1990 population in already-dense neighborhoods—a legacy of decades of prior growth—and when they built new housing during the 1990s, they added it disproportionately through infill and redevelopment rather than sprawl. Second, this reliance intensified dramatically over three decades, with big and small cities diverging in where they built new housing. Third, all of the housing cost divergence across metro areas occurred in already-dense neighborhoods; fringe areas showed no price divergence. I begin by characterizing cross-sectional differences in development patterns as of 1990, then trace how these patterns evolved, and finally link this evolution to the spatial distribution of price growth.

Figure 4 reveals striking differences in how cities of different sizes are built. The left panel shows the distribution of neighborhood densities around housing units in 1990—that is, the share of homes located in neighborhoods developed at different intensities. Metro areas are grouped by their 1990 population. While every city has fringe neighborhoods where land is lightly developed, larger metros house a disproportionate share of their population in neighborhoods where land is used very intensively.

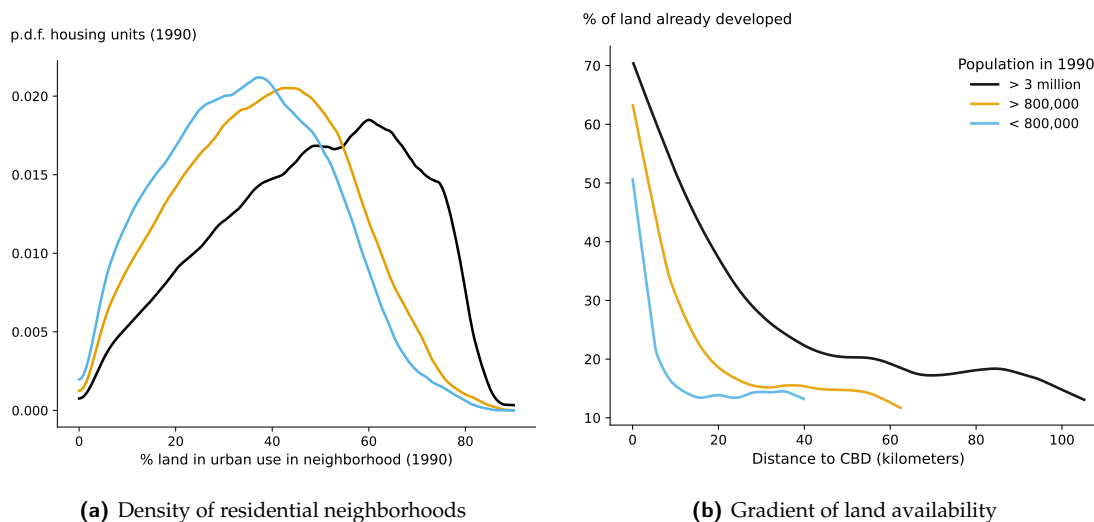


Figure 4: Distribution and gradient of neighborhood densities in 1990, by metro area population. Panel (a) shows the distribution of neighborhood densities, weighted by housing units. Panel (b) shows its gradient as a function of distance in kilometers from the city center. Metro areas are grouped by their 1990 population. Metros with more than 3 million inhabitants in 1990 were: New York, Los Angeles, Chicago, Washington DC, San Francisco, Boston, Philadelphia, Detroit, Houston, Dallas, Miami. Metros with 800,000 to 3 million inhabitants are listed in the Appendix.

The right panel in figure 4 shows how this pattern plays out across space, plotting land use intensity as a function of distance from each city's center. Big cities are more intensively developed throughout their entire urban extent. While fringe areas are developed at similar intensities

across the city size distribution, they lie much further from downtown in larger metros—often more than 100 kilometers out. In other words, the parts of big cities where substantial land availability exists are located far from their urban cores, making them less accessible. Moving toward the center, the differences in development intensity become even starker: big cities tend to have their land much more intensely developed at any given distance from downtown.

These differences in existing housing stocks translate directly into the flows of new construction. Rather than showing the full distribution immediately, I begin with the cross-metro relationship: Figure 5, panel (a), plots 1990 metro population against the average land-use intensity around the typical new housing unit built during the 1990s. Larger metros systematically built new housing in more intensively developed neighborhoods. Intuitively, because big cities already housed a larger share of their population in dense areas, subsequent construction was more likely to occur there as well. The full distribution of where new units were built across neighborhood intensities is presented below in Figure 6.

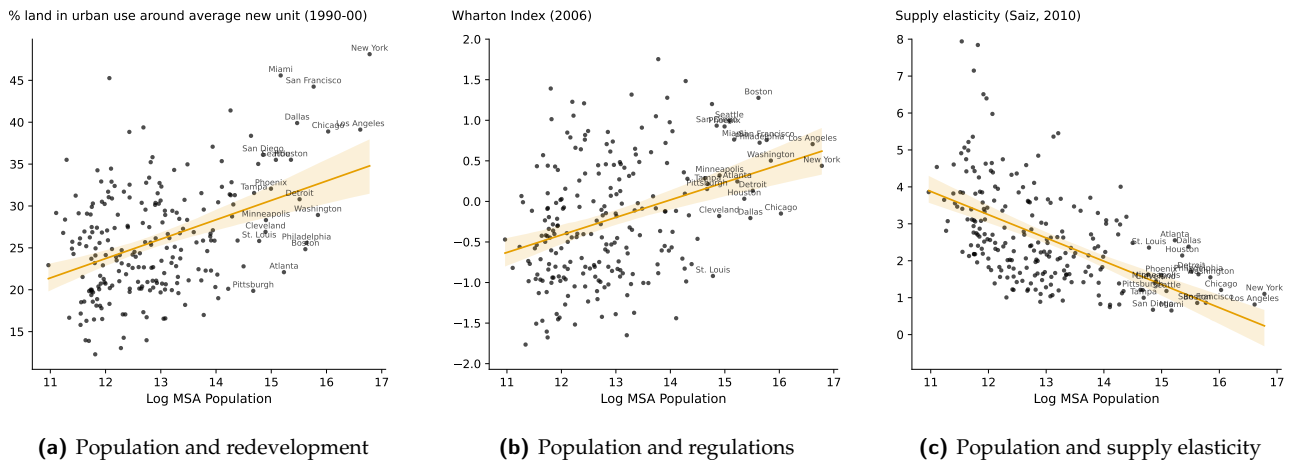


Figure 5: MSA-level scatterplots showing the relationship between the metro’s log population in 1990 and (a) the mean share of developed land around the average new housing unit built during the 1990s (measuring reliance on intensive margins); (b) the Wharton Index of housing regulations from Gyourko et al. (2008); and (c) the Saiz (2010) housing supply elasticity estimates. *Note:* Each point represents one of the 236 Metropolitan Statistical Areas (MSAs) for which both the Wharton Index and Saiz elasticity estimates are available.

The natural implication of this evidence is that any study that focuses only on regulations as a source of supply constraints risks omitting a key confounder: large metros rely more heavily on the costlier margins of construction. Figure 5 illustrates the broader pattern. Panel (b) shows that bigger cities had, on average, more stringent housing regulations (Wharton Index, Gyourko et al. (2008)), and panel (c) shows that they had lower housing supply elasticities (Saiz (2010)). These city-wide measures of constraints are strongly correlated with the degree to which cities rely on intensive development margins—and, as shown in the next subsection, this pattern has only intensified over the decades.

A closer look at supply elasticity schedules reveals that the relationship between regulations and constraints is more nuanced than aggregate measures suggest. Figure A2 in the Appendix shows how supply elasticities vary with distance from the CBD and development intensity across metros with different regulatory environments. More regulated cities do exhibit lower supply elasticities at any given distance from downtown. However, this difference largely reflects the fact that they are more heavily developed at those distances. Heavily developed areas show low supply elasticities regardless of the regulatory environment, while areas where greenfield development is possible exhibit larger differences.

3.1 The divergence by city size since 1990

The cross-sectional patterns documented in the previous subsection reveal that larger cities already relied more heavily on intensive development margins during the 1990s. This subsection documents how these patterns evolved over the subsequent three decades. Big cities became substantially more dependent on infill and redevelopment to build new housing, while small metros continued accommodating growth primarily through expansion. This divergence in development patterns mirrored very closely the divergence in housing costs across. All of the divergence in housing costs across cities happened through differential price growth in already-dense neighborhoods: Fringe prices did not diverge since 1990.

Figure 6 presents how cities further *diverged* by city size in their patterns of new construction. Each of the three panels plots the p.d.f. of local land use intensity around housing units built during each decade 1990-2000, 2000-2010, and 2010-2020, separately by metro size category. Over the three decades, bigger cities only increased their reliance on the intensive margins of development, while smaller cities saw little change. Figure A3 showcases this divergence across metro areas by plotting the metro-level distributions, and the systematic relationship with city size over time. Figure A4 further de-composes this to show that the divergence happened both for geographically constrained and un-constrained metro areas.

It is worthwhile to think about how this divergence happened in terms of distance to the CBD. In figure A5 in the Appendix, I show how, in the 1990s, new construction happened *at the same relative distance from downtown* across the city size distribution, that is, primarily at the urban fringe of each city. Therefore, big cities were tapping more into infill and redevelopment because they were more densely built on average. However, in subsequent decades, big cities experienced a surge of new construction activity in their central cities, while small cities continued to build primarily at their fringes.



Figure 6: Evolution of the share of new housing built in neighborhoods of different densities, by metro area population. Each panel shows the distribution of neighborhood densities weighted by new housing units constructed during each decade (1990-2000, 2000-2010, 2010-2020). Panel (a) shows metros with more than 3 million inhabitants in 1990; panel (b) shows metros with 800,000 to 3 million inhabitants; panel (c) shows metros with fewer than 800,000 inhabitants.

Figure 7 presents the evolution of housing costs¹¹ as a function of neighborhood density across different metro size categories. Average housing prices diverged dramatically across cities, but this divergence occurred only in densely developed neighborhoods. Downtowns (In these plots, neighborhoods to the right of the land use intensity distribution) exploded in price in big cities relative to these cities' own fringe areas, while small city downtowns saw no price growth. Areas at the middle of the distribution — infilled residential neighborhoods — saw a more moderate divergence in prices. Finally, fringe areas saw no divergence at all: living in the urban fringe of a big city in 2020 was no more expensive, relative to the fringe of a small city, than it had been in 1990. Figure A6 in the Appendix shows the same pattern in terms of distance to the CBD.

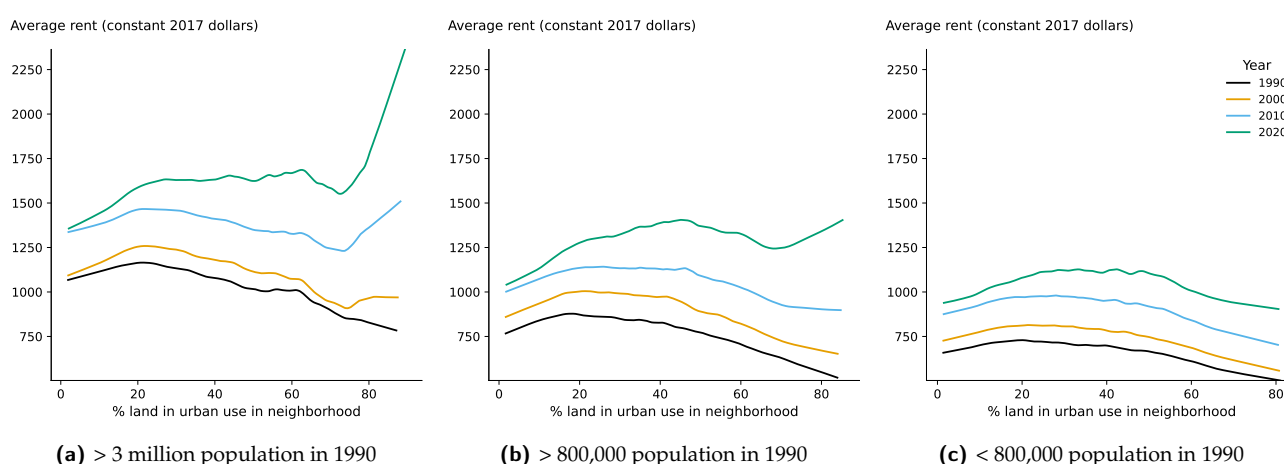


Figure 7: Evolution of housing costs (monthly rent) as a nonparametric function of local intensity of land use. Each panel plots the same variable for each of the metro area size groups used throughout the section.

¹¹Measured as self-reported monthly rent from the Census

These results, together, paint a consistent picture, where within-city supply elasticities have (roughly) remained constant, and it is the demand that has changed sharply within big cities, changing the composition of where new construction takes place, driving up prices in already-dense neighborhoods, and pushing down the *aggregate* supply elasticities of metro areas as a whole due to this demand concentration. In the next section, I develop a model to rationalize these empirical findings.

4 Model

To rationalize the empirical findings in the previous section, I develop a model where the three margins of urban growth—greenfield expansion, infill, and redevelopment—have different costs and respond endogenously to household demand. The model shows how two forces concentrate demand in already-developed neighborhoods: stagnating commuting speeds that raise the value of central locations, and rising urban amenities that make downtown living more attractive. As both forces push construction toward costlier margins, cities become more supply-inelastic, amplifying the impact of demand shocks on prices.

4.1 Household choice and the spatial structure of metro areas

There is a set $\mathcal{M}$ of metro areas in the country, which differ in their geography and productivity. Metro areas are monocentric and circular, with locations being indexed by their distance r to the city center. Only an arc $\alpha_m \in [0, 2\pi]$ is available for development, reflecting how some metros have more geographic barriers to development than others.

Households have a total population of N and are indexed by ω . First, they choose whether to live in one of the metro areas $m \in \mathcal{M}$ or in the rural outside option. Then they choose which metro area to live in, which neighborhood $i \in \{D, R, M\}$, and which distance r within the bounds of their chosen neighborhood. Households living in a metro area must commute to access their jobs. Following Duranton and Puga (2023), the commuting costs faced by a household living at a distance r from the city center are given by:

$$\tau_m(r) = \tau \cdot (N_m)^\phi \cdot r^\gamma \quad (1)$$

Distance traveled by households increases with r with an elasticity $\gamma > 0$. τ is a parameter capturing the baseline speed of travel, and $(N_m)^\phi$ captures the congestion of the metro's roads, which increases with the population of the metro area N_m with an elasticity $\phi > 0$.

Neighborhoods $i \in \{D, R, F\}$ in metro area m differ in their level of residential amenities A_{im} , and households have idiosyncratic preferences over metro areas and neighborhood types. Within each neighborhood, households arbitrage over the price of housing and commuting costs. Thus, the indirect utility of a household ω living in neighborhood i in metro area m at distance r is given by:

$$V_{im}(\omega, r) = (w_m - \tau_m(r) - P_{im}(r)) \cdot A_{im} \cdot v_{im}(\omega) \quad (2)$$

The idiosyncratic shock $v_{im}(\omega)$ is drawn i.i.d. from a generalized extreme value (GEV) distribution:

$$G(\{v_{im}\}) = \exp \left(- \left[\sum_m \left(\sum_{i \in I_m} v_{im}^{-\theta} \right)^{-\psi/\theta} \right] \right) \quad (3)$$

where I_m denotes the set of neighborhoods in metro m , θ governs the within-metro elasticity of substitution, and ψ governs the across-metro substitution. In case they choose to live in the rural outside option, households receive a utility $V_{Rural} = v_{Rural} \cdot v_{Rural}(\omega)$, with $v_{Rural}(\omega)$ being drawn from a GEV distribution with parameter ψ (This implies that, to the households, the choice of the rural outside option is the same as any other metro area).

In order for households to be distributed across all locations in a neighborhood, the bid-rent schedule must ensure households are indifferent as they trade off housing and commuting costs. For any two locations $r' > r$ within a neighborhood, the price of housing $P_{im}(r)$ must therefore perfectly capitalize any change in commuting costs, $\tau_m(r)$:

$$\forall i \in \{D, R, F\} : \quad P_{im}(r) = P_{im}(r') + \tau N_m^\phi (r'^\gamma - r^\gamma) \quad (4)$$

A direct consequence of this arbitrage is that a household's indirect utility is independent of their specific location within the neighborhood:

Proposition 1 (Constant Utility within Neighborhoods). *The indirect utility $V_{im}(\omega, r)$ of a household ω is constant for all radii r within a given neighborhood i in metro m .*

See proof on page A-7.

This result allows us to write the indirect utility for a neighborhood choice without reference to a specific radius. For convenience, let $\hat{P}_{im}$ and $\hat{\tau}_{im}$ be the price and commuting cost evaluated at any arbitrary reference point within neighborhood i . The indirect utility can then be written as:

$$V_{im}(\omega) = (w_m - \hat{\tau}_{im} - \hat{P}_{im}) \cdot A_{im} \cdot v_{im}(\omega) \quad (5)$$

Once the idiosyncratic shocks are drawn, this gives us the familiar conditional logic probabilities across metro areas and neighborhoods (McFadden, 1974), conditional on choosing to live in a metro area:

$$\pi_{im} = \underbrace{\left(\frac{v_{im}^\theta}{\sum_{j \in I_m} v_{jm}^\theta} \right)}_{\pi_{i|m}} \cdot \underbrace{\left(\frac{\left(\sum_{j \in I_m} v_{jm}^\theta \right)^{\psi/\theta}}{\sum_{k \in \mathcal{M}} \left(\sum_{l \in I_k} v_{lk}^\theta \right)^{\psi/\theta}} \right)}_{\pi_m} \quad (6)$$

where $v_{im} = (w_m - \hat{\tau}_{im} - \hat{P}_{im}) \cdot A_{im}$ is the deterministic part of indirect utility. Then, the probability of choosing to live in a metro area is given by:

$$\pi_{Urban} = \frac{\sum_{m \in \mathcal{M}} \left(\sum_{i \in I_m} v_{im}^\theta \right)^{\psi/\theta}}{\sum_{m \in \mathcal{M}} \left(\sum_{i \in I_m} v_{im}^\theta \right)^{\psi/\theta} + v_{Rural}^\psi} \quad (7)$$

4.2 Housing Supply and Land Rents

A competitive developer sector rents land from landowners to rent housing to households. At each location r , competitive developers choose the type of land use—Downtown (D), Residential (R), or Fringe (F)—that maximizes their profit. The boundaries between major neighborhoods are thus determined endogenously at the radii $\{r_{Dm}, r_{Rm}, r_{Fm}\}$ where developers are indifferent between two adjacent uses.

Since the market is perfectly competitive, the rent of a housing unit $P_{im}(r)$ is the sum of land rents $R_m(r)$, which I interpret as the value of a unit of land *before* obtaining construction permits, and a permitting and construction cost per housing unit c_{im} . In the urban fringe, where development occurs on greenfield sites, variations in c_{im} primarily reflect regulatory barriers to converting rural land to urban use (Duranton and Puga, 2023). However, as we move toward denser, more central neighborhoods, c_{im} increasingly reflects the physical costs of infill and redevelopment (such as demolition, site assembly, and building on constrained parcels), in addition to the regulatory costs of obtaining permits in already-developed areas.

The values of c_{im} to the expansion of the city's zones. As each zone expands into the next, the c_{im} wedge rises. In the fringe, this primarily reflects intensified regulatory opposition to converting additional rural land to urban use. In denser neighborhoods, the cost increase reflects both the rising physical difficulty of redevelopment as sites become more constrained, and the escalating regulatory burden as existing residents organize against new construction. The rate at which c_{im} rises as neighborhoods expand is governed by an inverse supply elasticity η_{im} . Locations that are developed at the same intensity but with different regulatory environments have different elasticities. These supply elasticities are obtained from the recent set of estimates by Baum-Snow and

Han (2024).

I will present the supply side in two steps. First, I will describe the developer's problem taking as given the regulatory costs c_{im} . For any set of values c_{im} and a demand schedule, this identifies the extent to which each neighborhood type is developed, the prices of housing, and the land rent gradient. Then, I will describe how the regulatory costs c_{im} evolve over time as neighborhoods grow, based on their local housing supply elasticity. At this point, I will introduce the time subscript in the variables to describe the transition from one decade to the next.

4.2.1 The Developer's Problem

Downtown Development. In the downtown area, developers can fully exhaust the developable lots and build vertically, but face high fixed costs denoted by c_{Dm} . Once they pay the fixed cost, they choose a Floor Area Ratio ($Dens_D$) to build. The variable cost of construction increases with $Dens_D$ according to a Cobb-Douglas technology with elasticity $\eta_{Dm} > 0$ and a cost shock ε_{Dm} . For any location r within the downtown zone, a developer's problem is to choose the $Dens_D$ that maximizes profit:

$$\max_{Dens_D > 0} \left\{ Dens_D \cdot P_{Dm}(r) - \frac{\varepsilon_{Dm}}{\eta_{Dm} + 1} Dens_D^{\eta_{Dm} + 1} - c_{Dm} - R_m(r) \right\} \quad (8)$$

The first-order condition yields the optimal floor area ratio, $Dens_D^*(r) = \left(\frac{P_{Dm}(r)}{\varepsilon_{Dm}} \right)^{1/\eta_{Dm}}$. Substituting this back into the objective function gives the maximized profit (before land rent) for a downtown developer:

$$\Pi_{Dm}(r) = \underbrace{\frac{\eta_{Dm}}{\eta_{Dm} + 1} \varepsilon_{Dm}^{-1/\eta_{Dm}} P_{Dm}(r)^{1+1/\eta_{Dm}}}_{\text{Variable Surplus}} - c_{Dm} \quad (9)$$

Suburban Development. In the suburban areas, developers choose between two fixed-density options. They can develop land at Fringe density $Dens_{Fm}$ for a cost of c_{Fm} per home, or at a higher Residential density $Dens_{Rm} > Dens_{Fm}$ for a higher unit cost of $c_{Rm} > c_{Fm}$ ¹². The profit (before land rent) for these uses at a location r is:

$$\Pi_{im}(r) = Dens_i \cdot (P_{im}(r) - c_{im}), \quad \text{for } i \in \{R, F\} \quad (10)$$

¹²Implicitly, this assumes that R neighborhoods infill to a fraction of developed lots $Fracdev_R > Fracdev_F$

4.2.2 Endogenous Boundaries, Prices, and the Land Rent Gradient

The spatial structure of each metro area is determined by the competition for land. At every location r , developers choose the land use that yields the highest return. In a competitive market, they bid up the price of land, $R_m(r)$, until it equals the full profit (net of non-land costs) of the winning use. The equilibrium land rent schedule is therefore the upper envelope of the potential profits from all uses:

$$R_m(r) = \max \{0, \Pi_{Fm}(r), \Pi_{Rm}(r), \Pi_{Dm}(r)\} \quad (11)$$

The boundaries between neighborhoods emerge at the specific radii where the profit curves of two adjacent land uses intersect, making developers indifferent between them. We can characterize the entire price and rent structure of the city by solving for the prices at these boundaries, starting from the city's edge and moving inwards. Proposition 2 summarizes these results.

Proposition 2 (Equilibrium Prices and Land Values at Endogenous Boundaries). *In equilibrium, housing prices at the three endogenous boundaries $\{r_{Fm}, r_{Rm}, r_{Dm}\}$ reflect two components: (i) the land value derived from capitalized commuting savings, and (ii) the wedge c_{im} that governs neighborhood expansion. Given model parameters, boundaries, and the values of the wedges c_{im} , the housing costs at each boundary de-compose in the following way:*

- (i) **Fringe-Rural Boundary (r_{Fm}):** At the urban edge, the value of land is that of the rural outside option, which is normalized to zero. Thus, the equilibrium price equals the marginal cost of fringe development, revealing the wedge:

$$P_{Fm}(r_{Fm}) = c_{Fm} \quad (12)$$

- (ii) **Residential-Fringe Boundary (r_{Rm}):** At the Residential-Fringe boundary (r_{Rm}), the price covers the residential construction cost plus the capitalized value of commuting savings relative to the fringe:

$$P_{Rm}(r_{Rm}) = \underbrace{c_{Rm}}_{\text{wedge}} + \underbrace{\frac{\text{Dens}_{Fm}}{\text{Dens}_{Rm}} \tau N_m^\phi (r_{Fm}^\gamma - r_{Rm}^\gamma)}_{\text{Land value from commuting savings}} \quad (13)$$

- (iii) **Downtown-Residential Boundary (r_{Dm}):** At the Downtown-Residential boundary (r_{Dm}), the price is determined by the full land value, which includes capitalized commuting savings from both the Residential and Fringe zones:

$$P_{Dm}(r_{Dm}) = \left[\left(\frac{\eta_{Dm} + 1}{\eta_{Dm}} \right) \varepsilon_{Dm}^{1/\eta_{Dm}} \cdot (c_{Dm} + \tau N_m^\phi [\text{Dens}_{Rm}(r_{Rm}^\gamma - r_{Dm}^\gamma) + \text{Dens}_{Fm}(r_{Fm}^\gamma - r_{Rm}^\gamma)]) \right]^{\frac{\eta_{Dm}}{\eta_{Dm} + 1}} \quad (14)$$

These expressions allow us to separate the pre-permitting land values (determined by commuting savings)

and the wedges c_{im} .

See proof on page A-7. Note that these boundary prices, and the commuting costs at the boundaries can be used as the representative $\hat{P}_{im}$ and $\hat{\tau}_{im}$ values to solve equilibria. That is, we can set $\hat{P}_{im} = P_{im}(r_{im})$, and similarly for commuting costs, $\hat{\tau}_{im} = \tau N_m^\phi r_{im}^\gamma$ for $i \in \{F, R, D\}$, without having to compute *average* prices or commuting costs within each ring.

4.2.3 Total housing supply in each ring

The total housing supply in each ring is obtained by integrating over the available land and built density within its radii. Since F and R ring are built at fixed densities, the total housing supply in each of these neighborhoods has a simple closed form:

$$\forall i \in \{R, F\} : H_{im} = \text{Dens}_{im} \cdot \alpha_m \cdot \frac{r_{i-1,m}^2 - r_{im}^2}{2} \quad (15)$$

In the case of the *Downtown*, we need to integrate over the variable FAR from the CBD all the way to the r_{Dm} boundary, which comes from the bid-rent condition in equation 4. The total housing supply in the downtown neighborhood is given by:

$$H_{Dm} = \alpha_m \int_0^{r_{Dm}} \left(\frac{P_{Dm}(r_{Dm}) + \tau N_m^\phi (r_{Dm}^\gamma - r^\gamma)}{\varepsilon_{Dm}} \right)^{1/\eta_{Dm}} r dr \quad (16)$$

Where $P_{Dm}(r_{Dm})$ is given by equation 14.

4.2.4 The response of c_{im} to neighborhood growth

The construction wedges c_{im} react to the expansion of each neighborhood type. As it grows, local opposition to development ratchets up, raising the regulatory cost c_{im} . The update is given by the following rule:

$$c_{im,t'} = c_{im,t} \cdot \exp(\eta_{im} \Delta \log \mathcal{A}_{im,t'} + \varepsilon_{im,t'}) \quad (17)$$

Where $\Delta \log \mathcal{A}_{im,t'}$ measures the growth in the physical area $\mathcal{A}_{im}$ of neighborhood i between periods t and t' :

$$\Delta \log \mathcal{A}_{im,t'} = \log(\mathcal{A}_{im,t'}) - \log(\mathcal{A}_{im,t}), \quad \text{where } \mathcal{A}_{im,t} = \pi(r_{im,t}^2 - r_{i-1,m,t}^2) \quad (18)$$

4.3 Spatial Equilibrium

A spatial equilibrium at time t is fully characterized by the vector of endogenous neighborhood boundaries $\mathbf{r}_t = \{r_{Dm,t}, r_{Rm,t}, r_{Fm,t}\}_{m \in \mathcal{M}}$ across all metro areas. Given the exogenous fundamentals—wages $\{w_{m,t}\}$, amenities $\{A_{im,t}\}$, construction cost wedges $\{c_{im,t}\}$, and the intensive margin parameters for downtown development—the boundaries are determined by market clearing conditions in every urban neighborhood simultaneously. In the transition from one decade to the next, the growth of each zone endogenously updates the construction wedges $\{c_{im,t}\}$ according to equation 17, so that the evolution of the urban system has a recursive structure decade by decade.

4.3.1 Spatial equilibrium given $\{c_{im,t}\}$

Formally, an equilibrium is a vector of boundaries $\mathbf{r}_t^*$ that clears all housing markets: For every zone $i \in \{D, R, F\}$ in every metro area $m \in \mathcal{M}$:

$$H_{im}(\mathbf{r}_t^*) = N_{im}^{\text{demand}}(\mathbf{r}_t^*) \quad (19)$$

The left-hand side, $H_{im}(\mathbf{r}_t^*)$, is the housing supply determined by the profit-maximizing decisions of competitive developers, as detailed in section 4.2.3. The right-hand side, $N_{im}^{\text{demand}}(\mathbf{r}_t^*)$, is the number of households choosing to live in neighborhood i of metro m , derived from the nested logit structure in equation 6. Crucially, both sides depend on the full vector of boundaries $\mathbf{r}_t^*$: the boundaries determine housing supply directly through geometry, and indirectly determine demand through their effect on housing prices and commuting costs, which in turn affect the deterministic utility v_{im} that households use to make location choices.

4.3.2 Equilibrium Transition

The model is solved recursively in decadal time steps. The state of the economy at time t is summarized by the vector of boundaries $\mathbf{r}_t$ and construction wedges $\mathbf{c}_t = \{c_{im,t}\}$ across all metro-zone pairs. At the transition to period t' , the economy experiences a sequence of exogenous shocks. Wages evolve to their new values $\{w_{m,t'}\}$ and amenities to $\{A_{im,t'}\}$, reflecting changes in productivity and residential quality across metros and zones. The national population grows to $N_{t'}$, the baseline commuting cost parameter updates to $\tau_{t'}$, and the rural outside option utility becomes $v_{\text{rural},t'}$. Finally, idiosyncratic supply shocks $\{\varepsilon_{im,t'}\}$ are realized for each zone in each metro, capturing unobserved regulatory or political changes that affect construction costs independently of neighborhood expansion.

Given these shocks and the previous state $(\mathbf{r}_t, \mathbf{c}_t)$, the construction wedges mechanically respond to the growth or contraction of each zone according to equation 17 based on the change in their physical area and the local housing supply elasticity η_{im} .

The new equilibrium at time t' is then a vector of boundaries $\mathbf{r}_{t'}^*$ such that housing markets clear in every zone of every metro area. That is, for every zone $i \in \{D, R, F\}$ in every metro area $m \in \mathcal{M}$, the equilibrium boundaries must satisfy:

$$H_{im}(\mathbf{r}_{t'}^*; \mathbf{c}_{t'}) = N_{im}^{\text{demand}}(\mathbf{r}_{t'}^*; \{w_{m,t'}\}, \{A_{im,t'}\}, \tau_{t'}, v_{\text{rural},t'}, N_{t'})$$

In section D.2 in the Appendix I describe the numerical solution to solve the equilibrium. In order to solve the model, I define a fixed-point mapping between housing quantities and their corresponding boundaries, which allows me to solve in population space via a pseudo-Newton method.

4.4 Calibration of the Model

In this subsection I describe the calibration of the model's parameters and exogenous fundamentals, how I construct the spatial structure of each metro area, and the inversion procedure to recover residential amenities and unobserved supply shocks. Table 1 summarizes the parameters used in the model and their sources.

4.4.1 Metro spatial structure

The first step in the calibration is to aggregate my dataset to the stylized structure of the model, while respecting the spatial scales of metro areas. Matching the spatial scale is crucial, since commuting costs in the model depend directly on it.

I aggregate the hexagon-level data described in section to the major neighborhood types $i \in \{D, R, F\}$ by binning the hexagonal grid cells on the basis of their land use intensity. In principle, the threshold to classify a hexagon *Downtown* is that it is 67% developed or more, which is the point where we observe a sizeable increase in multi-family buildings, as well as a sharp differential increase in prices and building activity. Since smaller metros do not tend to have much housing in such high-density areas, I add an additional criterion: Whenever a metro did not have at least 15% of its housing in hexagons with at least 67% development in 1990, the *Downtown* threshold is lowered to the density of the 85th percentile in that metro.

I follow a similar approach with the *Residential* threshold, which I set to 35% development,

that is, roughly the density of the 45th percentile home nationally in 1990. Whenever a metro area does not have at least 45% of its housing in hexagons with at least 35% development, I lower the threshold to the density of the 45th percentile in that metro. All urban hexagons below the *Residential* threshold are classified as *Fringe*.

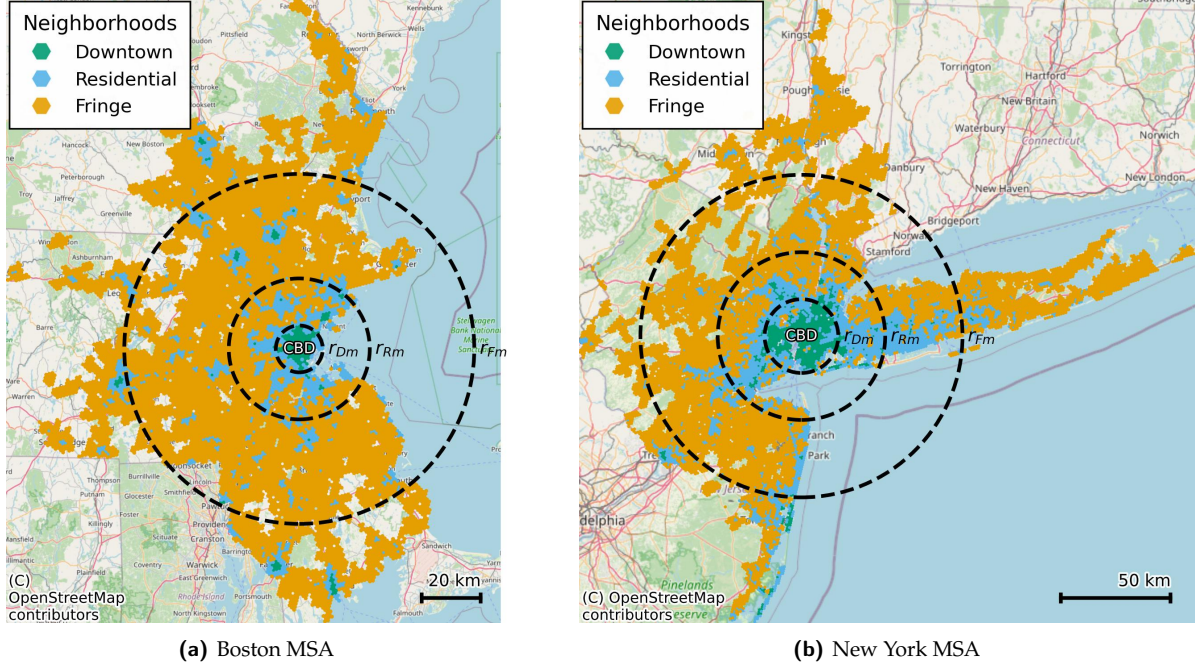


Figure 8: Fit of the stylized monocentric boundaries to the observed neighborhood classifications for the Boston (a) and New York (b) metro areas in 2020. The colored areas represent the actual classification of hexagonal neighborhoods into Downtown (green), Residential (blue), and Fringe (orange) based on their land use intensity. The dashed circles represent the model’s inverted radii (r_{Dm}, r_{Rm}, r_{Fm}), which define the boundaries of the three stylized zones in the monocentric model.

Once the hexagons are classified, I use their local housing densities to compute the average housing density of each neighborhood type in each metro area, $Dens_{im}$. I also compute the arc of developable land α_m in each metro area by computing the amount of developable land in a 50 kilometer radius around the CBD, similar to Saiz (2010). The combination of these two measures and the housing stocks H_{imt} allows me to invert the radii r_{imt} . Figure 8 shows the fit of the stylized boundaries to the true geography of the city for Boston and New York metro areas in 2020. Figure A11 shows the fit of the model to average distance to the CBD across all metro areas and years.

4.4.2 Demand elasticities

I parametrize the demand elasticities θ and ψ using parameter estimates from similar settings in the literature. For the across-metro elasticity of substitution ψ , I set a value of 3, following (Hsieh and Moretti, 2019). For the elasticity of substitution between neighborhoods within a metro, the size of the neighborhoods matter: Couture et al. (2024) estimate an elasticity of 3 between “downtown” (the 10% most central neighborhoods of a metro) and the rest of the metro in large

Parameter	Value	Source
<i>Demand Parameters</i>		
θ	4	Couture et al. (2024), Ahlfeldt et al. (2015)
ψ	3	Hsieh and Moretti (2019)
<i>Commuting Technology Parameters</i>		
τ	Between 112 and 181	Own calibration
γ	0.07	Duranton and Puga (2023)
ϕ	0.04	
<i>Housing Supply Elasticities (By metro and major neighborhood)</i>		
η_{Dm}	See Figure 9	Baum-Snow and Han (2024)
$\eta_{R \rightarrow Dm}$		
$\eta_{F \rightarrow Rm}$		
η_{Fm}		

Table 1: This table summarizes all parameters used in the quantitative spatial model. “Own calibration” refers to parameters constructed directly from the dataset described in Section 2. Supply elasticity parameters are imputed from tract-level estimates by averaging over relevant neighborhood types within each metro area. Residential amenities A_{im} are recovered through the model inversion procedure detailed in the Appendix. The baseline commuting cost parameter τ is calibrated separately for each year to match average commuting times in urban areas.

US metro areas, while Ahlfeldt et al. (2015) find an elasticity of 6.8 between circa 3000 blocks in Berlin. Therefore, I set a value of θ of 4, somewhat more elastic than the substitution between metros.

4.4.3 Commuting structure

I parametrize the commuting cost parameters γ and ϕ in equation 1 using estimates from Duranton and Puga (2023). They estimate γ , the elasticity of driving distance with respect to the residential distance from the CBD, from confidential microdata from the National Household Travel Survey. They find a value of 0.07, that is, on average, households increase their driving distance by about 0.7% for every 1 kilometer increase in their distance from the CBD. For parameter ϕ , that is, the elasticity of traffic congestion with respect to metro population, they produce estimates from two separate sources. One is the same confidential NHTS data they use for the γ parameter, while the other is as a dataset of millions of predictions of travel times, sampled from Google Maps by Akbar et al. (2023). Both methods yield an estimate of 0.04, which I use in my parametrization.

The baseline commuting cost parameter τ scales the total dollar value of commuting, and can be interpreted as a composite of a technology parameter that determines speed, and the dis-taste for time spent commuting. I calibrate it separately for each year in the following way. Given values for all other parameters, and the true equilibrium outcomes, each value of τ maps monotonically to the total dollar cost of all the commuting that takes place in the urban system. I

attribute to each hour spent commuting half the value of the average hourly wage (Couture et al., 2018), and then I set the value of τ so that the average commuting time in the model matches the average commuting time within urban areas in that year, as measured in tract-level Census and ACS data. Figure A12 in the Appendix shows the fit of this untargted moment.

4.4.4 Housing supply elasticities

I calibrate the supply elasticities of each metro-neighborhood using the Census Tract-level estimates produced by Baum-Snow and Han (2024). These estimates are produced by running a regression of the kind:

$$\Delta \log(H_{im}) = \theta_m + X_{im}\delta + \eta_{im} \cdot \Delta \log(P_{im}) + \epsilon_{im}$$

where η_{im} is parametrized to depend on tract observables, namely topography, urban land use intensity, and zoning regulations. Identification comes by instrumenting the observed price changes with a shift-share instrument (Bartik, 1991) of each tract's *Resident Commuter Market Access*, a commute-weighted metric of access to employment opportunities developed by Tsivanidis (2023).

These estimates were created using 2000-2010 data, and I employ them in my model under the assumption that *observationally similar neighborhoods in metro area m had the same supply elasticities throughout the period*. I impute each of the $\{\eta_{Dm}, \eta_{R \rightarrow Dm}, \eta_{F \rightarrow Rm}, \eta_{Fm}\}$ by averaging the elasticity estimates for the relevant areas: η_{Dm} , the intensive margin elasticity of downtowns, is the average for the areas classified as *Downtown* in 2000. $\eta_{R \rightarrow Dm}$ drives the costs of downtown expansion into the *Residential* ring, so it is the average elasticity of each metro's neighborhoods classified as residential. The same logic applies with the $\eta_{F \rightarrow Rm}$ that governs expansion into the fringe. Finally, for η_{Fm} I draw an expansion buffer of 15% around each urban area in 2000, and take the average of the intersecting tracts.

Figure 9 shows the distribution of housing supply elasticity estimates aggregated to metro-neighborhood level. As documented by Baum-Snow and Han (2024), there is substantial heterogeneity in supply elasticities both across and within metro areas. Each curve represents the p.d.f function of the η_{im} values of each ring across the 275 metro areas. Variations in supply elasticities across metros are captured by looking at the same curve. Note that the variation across metros is capturing both regulatory differences and the fact that I use lower density thresholds to classify the $\{D, R, F\}$ zones in smaller metros, which mechanically leads to higher supply elasticities for each zone.

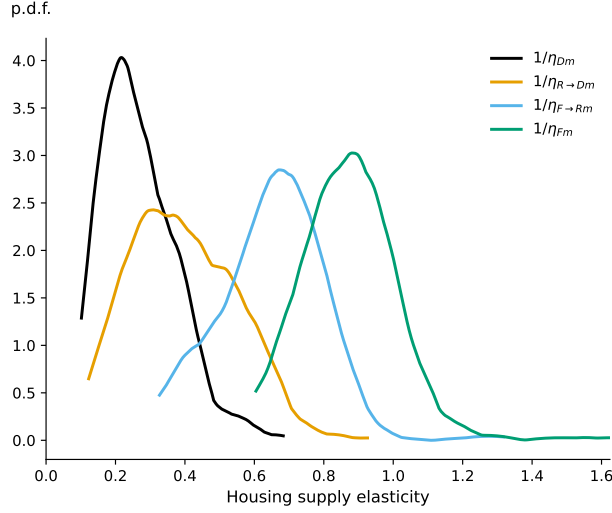


Figure 9: Density function of metro-major neighborhood type housing supply elasticities. Each curve is the kernel density of the average supply elasticities for neighborhoods of a type, at the level of metro areas. For example, the "Downtown" curve is the kernel density of $\{\eta_{Dm}\}_{m \in M}$ for each of the 275 metro areas.

4.5 Model inversion: Amenities, and supply shocks

Given parameter values and the observed equilibrium outcomes, the model can be inverted to recover the unobserved amenities A_{imt} and costs c_{imt} that rationalize, respectively, the observed population allocations and prices at each of the neighborhood boundaries. Once the costs c_{imt} have been recovered, the local housing supply elasticities and the growth of each neighborhood identify the unobserved supply shocks ε_{imt} .

The inversion of amenities is given by expanding the choice probability equation 6 to solve for A_{im} . For any given set of parameters, fundamentals, and observed outcomes, there is a unique vector of amenities that rationalizes the observed population allocations. The details of the inversion procedure are given in section D.1 of the Appendix. Similarly, the inversion of the costs c_{imt} is given by expanding the boundary pricing conditions in Proposition 2 to solve for c_{im} .

Once the costs c_{imt} have been recovered, I can use the growth of each neighborhood and the local supply elasticities to back out the unobserved supply shocks ε_{imt} from equation 17. Whenever I run any counterfactuals or decomposition exercises, I keep the history of these unobserved supply shocks.

5 Model results

With the calibrated model in hand, I now produce a set of results that shed light on the facts that I outline in section 3, and which allow me to measure the importance of cities' internal geographies

for housing affordability. First, I will use the demand fundamentals recovered from the model to understand why households pay such a high premium to live in a relatively reduced set of prime locations, and why this premium has increased so much since 1990. As I have reiterated before, this is the combination of several forces coming together to concentrate housing demand in a reduced set of very developed, supply-inelastic areas.

Second, I will produce a set of counterfactual exercises that isolate the role of the main mechanism in this paper. The first such counterfactual isolates the contribution of housing costs of the *additional* demand concentration since 1990, while the second attempts to fully shut of the costs of infill and redevelopment in big cities observe the population reallocations in such an economy.

5.1 Characterizing the concentration of demand

The first exercise that the model is useful for is to characterize the demand fundamentals, in their 1990 levels and how they changed since. For context throughout this section, I will report the evolution of each variable by grouping over main zones. For example, figure 10 the 1990-2020 divergence of housing costs for *all downtowns*, *all residential areas*, and *all fringe areas*, respectively, by plotting the log rent deviation from the national mean, weighted by the number of households living in each zone.

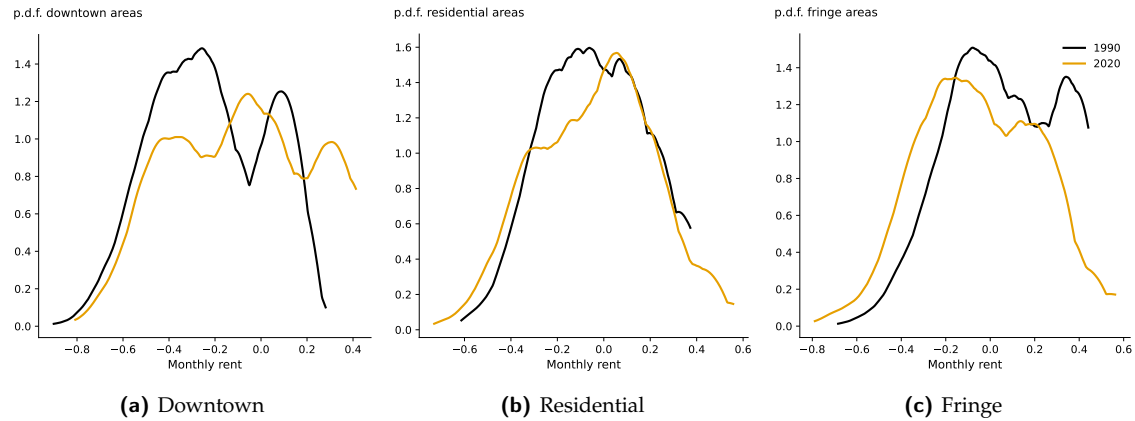


Figure 10: Rents Log rent deviation from national mean by neighborhood type, weighted by number of households. Each panel shows the evolution of relative housing costs for a different neighborhood type: (a) Downtown, (b) Residential, and (c) Fringe.

I start with the residential amenities A_{im} . Being the residual of the inversion procedure, they capture all the benefits of living in any given location that are not explained by wages and commuting costs. That is, the value of local public goods, such as schools, safety, environmental quality, but also urban consumption amenities, such as access to cultural institutions, restaurants, and social interactions¹³. Figure 11 showcases the distribution of recovered amenities by main

¹³This is the key insight of the Rosen-Roback model (Roback, 1982; Rosen, 1979)

zone (All downtowns, all residential areas, all fringe areas), in 1990 and 2020. The plots show the de-meaned log amenities, weighted by the number of households living in each zone.

The changes in the spatial distribution of amenities already partially explain why housing costs increased so much downtown. The right side of the distribution of downtown area amenities — corresponding to big cities — gets substantially fatter between 1990 and 2020, just as fringe amenities in big cities get compressed towards the mean. In other words, downtown areas in big cities became relatively more attractive in terms of amenities relative to their own fringe and to the rest of the country, and this is part of the reason why big cities turned to redevelopment of their downtowns.

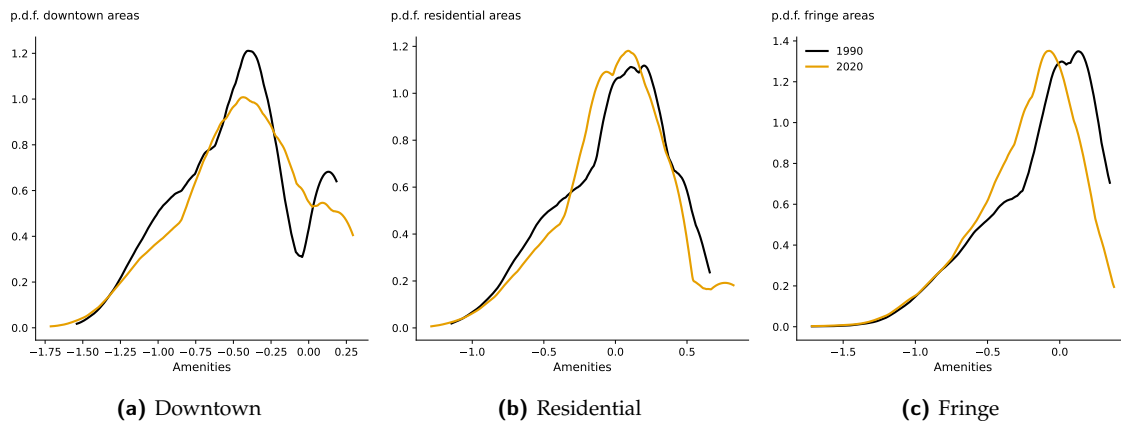


Figure 11: Amenities by neighborhood type, weighted by number of households. Each panel shows the evolution of relative housing costs for a different neighborhood type: (a) Downtown, (b) Residential, and (c) Fringe.

It is important to note that this finding is consistent with a large body of work in the *urban revival* literature. Starting with Baum-Snow (2020), Couture and Handbury (2020), Couture and Handbury (2023), and Diamond (2016), a large literature has documented the spatial concentration of residential amenities in central areas of big cities. While not the focus of this paper, this literature finds that this is an equilibrium outcome from the changing income sorting patterns of households: As young, high-income households sort into central areas of big cities, they create the demand for the provision of urban amenities that further fuels the attractiveness of these locations.

In turn, this resorting has deeper structural roots. Diamond (2016) documents how the differential wage growth for high skill workers in big cities was the spark for re-sorting across cities. In turn, Couture et al. (2024) shows how rising income inequality *within* those big cities allowed high-income households to outbid lower-income households for central locations, effectively creating *high-quality* central neighborhoods that did not exist before. Moreno-Maldonado and Santamaría (2025) document how delayed childbearing among high-income households extended

their period of valuing urban amenities, further concentrating demand in central neighborhoods.

For the purposes of this paper, the key takeaway is that the recovered amenities show a clear pattern of demand concentration in the downtowns of big cities, and that their suburban fringes became relatively less attractive over time. Given the surveyed literature, I interpret these recovered amenities as the byproduct of this broader structural shift in what makes cities attractive to households, and I take my recovered amenities as a history of exogenous fundamentals.

I then look at the evolution of commuting times. Figure 12 shows the evolution and distribution of model-implied two-way daily commuting times in minutes by main zone and city. Just as in the real data, shown in figure A7 in the appendix, commuting times vary more across metros than within, due to the vast differences in spatial scale and congestion levels across cities: bigger cities have longer commuting times. Because cities generally got bigger, commuting times increased across the board. Because urban growth creates more potential commuting savings, as outlined in Proposition 2, this generally concentrated demand in more central areas.

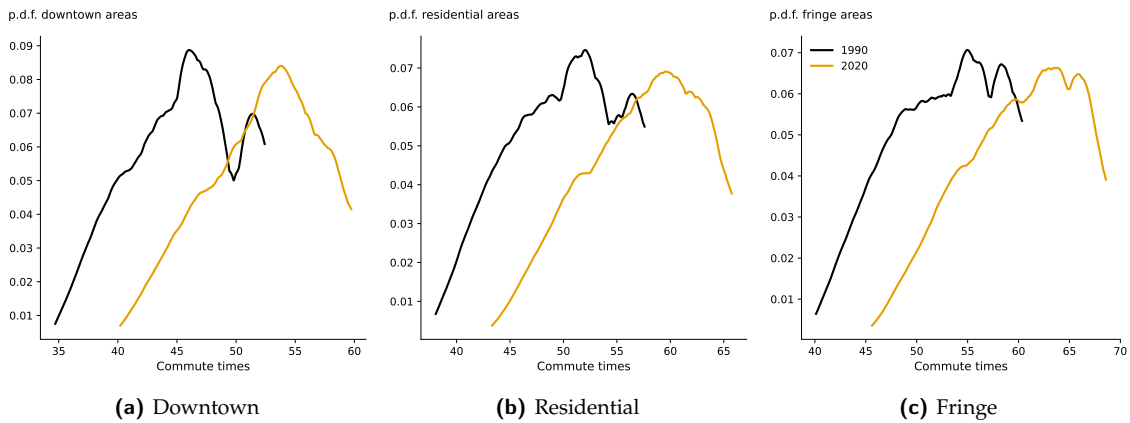


Figure 12: Evolution of commuting costs implied by the model by neighborhood type, weighted by number of households. Each panel shows the evolution of model-implied commuting costs for a different neighborhood type: (a) Downtown, (b) Residential, and (c) Fringe.

Why do households pay such a premium to live in big cities though? The answer lies in both wages and amenities. As was the insight of Glaeser et al. (2001), big cities attract population because they offer both higher wages and better amenities (See figure A8 in the appendix). And while the spatial distribution of amenities did not change much when considering metro-level averages, wages did substantially. Figure 13 shows the evolution of average wages and population across metro areas between 1990 and 2020. The biggest metro areas saw their average wages increase substantially more than smaller metros, fueling a strong concentration of population in big cities. Because the superstar metro areas were already very congested, the population increasingly grew in a “second tier” of big cities such as Houston, Dallas, Atlanta, Austin, or Miami.

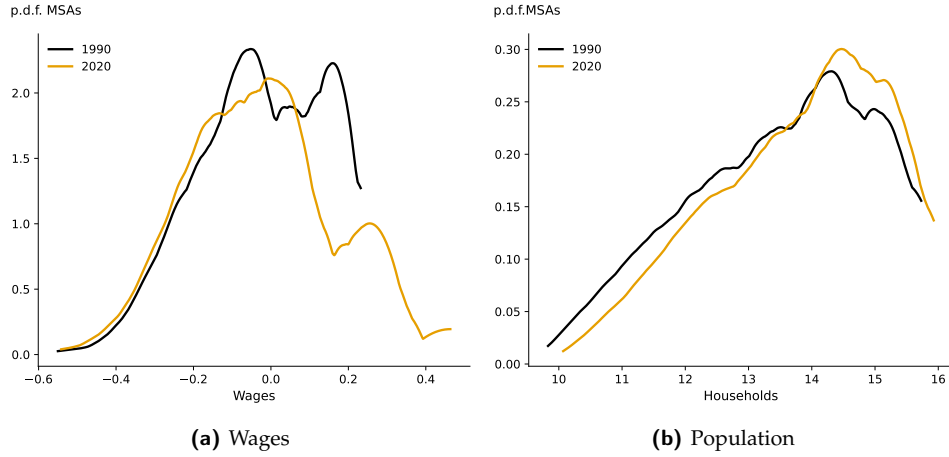


Figure 13: Evolution of wages and population across metro areas, 1990-2020.

5.2 Demand concentration and the rise in housing expenditures since 1990

The fact that housing demand *generally* concentrated in supply-inelastic areas since 1990 poses a natural question: how much of the rise in housing costs can be attributed to this demand concentration, and what fraction comes from concentration within big cities versus concentration towards big cities overall? To answer this, I construct a set of counterfactual histories of the economy that shut off some of the changes in the spatial distribution of the demand fundamentals — amenities, wages, commuting technology — while allowing the rest of the economy to evolve as observed. Comparing the outcomes from these counterfactual economies to the baseline allows me to isolate the contribution of each of these demand-side forces to the rise in housing costs since 1990.

The first such counterfactual isolates the contribution of shifting demand towards infill and redevelopment within metros. Part of this happened because metros simply grew in size and had their commuting times increase, while another part was because of *urban revival*, i.e. the within-city changes in the spatial distribution of amenities. Therefore, in this *within-city demand* counterfactual, I freeze the within-metro distribution of amenities at 1990 levels, and I construct a set of alternative commuting cost parameters τ that prevent the increase of commuting times. This way, metros' overall attractiveness (average metro-level amenities, metro-level wage distributions) evolves as observed, but the additional demand concentration within metros is shut off.

The second counterfactual, which I term the *across-city demand* counterfactual, isolates the orthogonal channel: I let the within-metro distribution of amenities and commuting costs evolve as observed, but I freeze the metro-level average amenities and wage distributions at 1990 levels (This mostly affects the wage distribution). This way, the additional demand concentration within metros is allowed to evolve as observed, but big cities do not become differentially more produc-

tive over time. Finally, in the third counterfactual, I shut off all demand concentration by freezing both the within-metro distribution of amenities and commuting costs, as well as the metro-level average amenities and wage distributions at 1990 levels. This serves as a benchmark to compare the two previous counterfactuals, and offers a view into an economy where the U.S. urban system grew in a relatively more uniform way.

As an outcome for these counterfactuals, I focus on a measure of housing expenditures, that is, what fraction of household income is spent on housing, which I define as the ratio of average monthly rent to average monthly income, in each metro area and nationally. Housing expenditures increased substantially between 1990 and 2020, from 28.42 % to 31.35 % of wage income nationally, and the standard deviation across metro areas increased by 45.16 %. Figure A9 in the Appendix shows the metro-level distribution. For this exercise, I keep the total urban population fixed, by adjusting the value of the rural outside option $v_{\text{rural},t}$ so that the changes in expenditures are only driven by the re-sorting of population within the urban system and not migration from rural areas.

	Housing expenditure national growth (p.p.)	Share of total mean change	Across-city std. dev. change	Share of total std. change
True 1990–2020 changes	10.31% (2.93 p.p.)	100%	45.16%	100%
<i>Contribution by:</i>				
Changes in demand within cities	2.33% (0.66 p.p.)	22.57%	-1.91%	−4.22%
Changes in demand across cities	1.49% (0.42 p.p.)	14.47%	16.48%	36.50%
All changes in demand	3.60% (1.02 p.p.)	34.88%	7.07%	15.66%

Table 2: Decomposition of housing cost growth, 1990-2020

Table 2 summarizes the results from these counterfactual exercises. Demand concentration since 1990 contributed to of the rise in national housing expenditures, of which more than half corresponds to the concentration of demand within cities. However, even if the entire rise in demand concentration was shut off, housing expenditures would have risen substantially. In the next subsection, I explore more extreme counterfactuals to take into account the full role of demand for infill and redevelopment, as well as regulations, in driving up housing costs.

5.3 The (limited) effects of de-regulation

Given the results found in this paper so far how much additional population could major cities *actually* accommodate under more permissive zoning, and at what prices? I address this question by simulating regulatory reform in seven cities with strict housing regulations: New York, Los

Angeles, Washington DC, San Francisco, Boston, Seattle, and San Diego¹⁴.

I construct counterfactual supply elasticities by exploiting the tract-level estimates from Baum-Snow and Han (2024), which parametrize how local observables—development intensity, topography, and regulations—map into supply constraints. For each of the seven cities, I recalculate what their tract-level elasticities would be if they had the supply elasticity schedule of the 75th percentile city¹⁵. I then aggregate these adjusted tract-level estimates into the four neighborhood-specific elasticities $\{\eta_{Dm}, \eta_{R \rightarrow Dm}, \eta_{F \rightarrow Rm}, \eta_{Fm}\}$ following the procedure in Section 4.4.4, and simulate the full urban system under this deregulation scenario. Figure A14 in the Appendix shows the adjusted supply elasticity schedules for these seven cities as a function of neighborhood fraction developed.

Table 3 summarizes the results from the de-regulation counterfactual. The results show that relaxing zoning would have modest effects in allowing these cities to accommodate more population at lower prices. The 1990-2020 population growth rate in these seven cities would have been 12% higher than baseline, allowing them to accommodate 2.7% more population by 2020. Housing costs grow 5% less than baseline. These results are significantly smaller than those in studies such as Hsieh and Moretti (2019) and Duranton and Puga (2023), because most of the demand in these big cities is concentrated in the *Downtown* and *Residential* areas, where the Baum-Snow and Han (2024) estimates predict that supply elasticities would remain low under deregulation.

An easy way to see this is to consider a naive model, more akin to Hsieh and Moretti (2019), which abstracts from the internal geography of cities altogether. I create one such model where each metro area is treated as a homogeneous unit with a single supply elasticity, amenity level, and commuting cost. In this naive model, detailed in Section F of the Appendix, de-regulation is simulated not by adjusting the supply elasticity schedule, but simply by setting the seven cities' elasticities to the national 75th percentile. A naive counterfactual like this predicts much larger effects from deregulation, just like what the literature usually predicts. The naive model predicts that deregulation would increase the 1990-2020 population growth rate in these cities by 39% and reduce price growth by 16%—effects roughly thrice as large as those from my full model.

These findings suggest that the de-regulation that is usually proposed by the literature cannot fully solve the housing challenges of big cities. While relaxing zoning constraints would help at

¹⁴I focus on these cities to facilitate comparison with Duranton and Puga (2023).

¹⁵Baum-Snow and Han (2024) introduce heterogeneity in the supply elasticity schedules across cities by estimating a Finite Mixture Model (FMM) with two classes. Cities have different probabilities of belonging to one class or another depending on their share developed 50 kilometers from the CBD, their geographic features, and their regulatory environment as measured by the Wharton Index. Since all these three elements are in fact highly correlated, I simply set the probability of belonging to the more elastic group to the 75th percentile nationally, in order to get an *upper bound* of the effect of regulations.

(a) Effect of deregulation on housing quantities and population

	Observed		De-regulation	
	Households by 1990 (thousands)	Households by 2020 (thousands)	Additional growth 1990-2020	Households by 2020 (thousands)
New York	6,786	8,270	6.6%	8,368
Los Angeles	5,178	6,503	27.9%	6,872
Washington	2,414	3,581	8.8%	3,684
San Francisco	2,404	3,020	9.4%	3,078
Boston	2,129	2,686	3.5%	2,706
Seattle	1,135	1,871	10.6%	1,949
San Diego	910	1,201	6.9%	1,221
Rest of cities	50,289	76,628	-1.2%	76,308
Rural areas	30,391	35,867	0.0%	35,867

(b) Effect of deregulation on housing costs

	Observed		De-regulation	
	Avg. rent by 1990 (2017 \$)	Avg. rent by 2020 (2017 \$)	Additional growth 1990-2020	Avg. rent by 2020 (2017 \$)
New York	1,070	1,711	-4.8%	1,681
Los Angeles	1,171	1,932	-6.8%	1,880
Washington	1,094	1,748	-7.5%	1,699
San Francisco	1,245	2,375	-3.8%	2,331
Boston	1,051	1,567	-4.5%	1,544
Seattle	888	1,802	-6.6%	1,742
San Diego	1,131	2,097	-2.6%	2,072
Rest of cities	770	1,197	-1.4%	1,191
Rural areas	—	—	—	—

Table 3: Counterfactual results from relaxing zoning in seven cities with strict housing regulations. Zoning relaxation is simulated by adjusting tract-level probabilities of belonging to high or low supply elasticity groups in the Baum-Snow and Han (2024) finite mixture model estimates to the 75th percentile nationally, then re-aggregating to neighborhood-level elasticities. Panel (a) shows the effect on housing quantities, while panel (b) shows the effect on housing costs.

the margin by making fringe development easier, the estimates from Baum-Snow and Han (2024) reveal that no city in the United States, regardless of its regulatory environment, has high supply elasticities in areas where redevelopment is necessary to accommodate additional demand. Recent work by Rollet (2025) micro-founds this in the context of New York City, showing that even in the complete absence of zoning, supply elasticities would be very low because of the high fixed costs of redevelopment. The implications of this will be discussed further in the conclusion.

6 Conclusion

This paper demonstrates that the internal geography of cities —both the demand and the supply— fundamentally shapes housing affordability and cities’ capacity for growth. In the context of U.S. cities, I show how the rising costs of housing in America reflect not just regulatory constraints, but a more fundamental form of congestion in where households want to live and how costly it is to build housing to accommodate them there. This challenge extends well beyond U.S. borders. Housing prices have surged in big cities in virtually every developed country. As economic activity continues to concentrate in major urban centers, this trend is likely to continue. Understanding this issue is therefore crucial for designing effective policy responses.

What policy responses might address this challenge? The exercises in this paper do suggest that de-regulation could help somewhat, but we need more detailed data and micro-founded models of redevelopment to really understand this. In this direction, a big advance has been made recently: Rollet (2025) shows how a complete de-regulation of zoning in New York City would help increase supply in the city, because some areas of face so much demand pressure that redevelopment would be profitable. However, even under complete de-regulation, the inherent costs to redevelopment would prevail: supply elasticities would still be low compared to virtually any area in the country, and housing costs would remain high.

However, an even more important question is, *what could reduce the concentration of demand in the first place?* Just as the railway (Heblich et al., 2020) and the automobile (Baum-Snow, 2007) once enabled access to urban opportunities without crowding, modern telecommunications may offer similar promise. Delventhal and Parkhomenko (2024) and Delventhal et al. (2022) show early signs that remote work can facilitate population reallocations that reduce pressure on the most expensive locations. Beyond such technological shifts, one could consider strengthening alternatives to superstar cities, something that has been happening organically in the U.S. as second-tier cities grow rapidly.

The path forward likely requires a combination of these strategies: smarter regulation that

reduces artificial barriers while acknowledging the real costs of dense construction, technological and organizational changes that reduce the spatial concentration of economic activity, and policies that strengthen alternatives to superstar cities. Each of these approaches deserves careful study to help solve what has become one of the most pressing policy challenges in developed economies.

References

- Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6), 2127–2189. <https://doi.org/10.3982/ecta10876>
- Akbar, P., Couture, V., Duranton, G., & Storeygard, A. (2023). The fast, the slow, and the congested.
- Alonso, W. (1964). *Location and land use: Toward a general theory of land rent*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674730854>
- Barrington-Leigh, C., & Millard-Ball, A. (2015). A century of sprawl in the united states. *Proceedings of the National Academy of Sciences of the United States of America*, 112(27), 8244–8249. <https://doi.org/10.1073/pnas.1504033112>
- Bartik, T. J. (1991). *Who benefits from state and local economic development policies?* W.E. Upjohn Institute. Retrieved April 25, 2025, from <http://www.jstor.org/stable/j.ctvh4zh1q>
- Baum-Snow, N. (2007). Did highways cause suburbanization? [Publisher: Oxford University Press]. *The Quarterly Journal of Economics*, 122(2), 775–805. Retrieved July 24, 2023, from <https://www.jstor.org/stable/25098858>
- Baum-Snow, N. (2020). Urban transport expansions and changes in the spatial structure of u.s. cities: Implications for productivity and welfare. *The Review of Economics and Statistics*, 102(5), 929–945. https://doi.org/10.1162/rest_a_00855
- Baum-Snow, N., & Han, L. (2024). The microgeography of housing supply [Publisher: The University of Chicago Press]. *Journal of Political Economy*, 132(6), 1897–1946. <https://doi.org/10.1086/728110>
- Baum-Snow, N., & Hartley, D. (2020). Accounting for central neighborhood change, 1980–2010. *Journal of Urban Economics*, 117, 103228. <https://doi.org/10.1016/j.jue.2019.103228>
- Boustan, L. P. (2010). Was postwar suburbanization "white flight"? evidence from the black migration. *Quarterly Journal of Economics*, 125(1), 417–443. <https://doi.org/10.1162/qjec.2010.125.1.417>
- Burchfield, M., Overman, H. G., Puga, D., & Turner, M. A. (2006). Causes of sprawl: A portrait from space. *Quarterly Journal of Economics*, 121(2), 587–633. <https://doi.org/10.1162/qjec.2006.121.2.587>
- Combes, P.-P., Duranton, G., & Gobillon, L. (2019). The costs of agglomeration: House and land prices in french cities. *The Review of Economic Studies*, 86(4), 1556–1589. <https://doi.org/10.1093/restud/rdy063>
- Couture, V., Duranton, G., & Turner, M. A. (2018). Speed. *The Review of Economics and Statistics*, 100(4), 725–739. https://doi.org/10.1162/rest_a_00744

- Couture, V., Gaubert, C., Handbury, J., & Hurst, E. (2024). Income growth and the distributional effects of urban spatial sorting. *The Review of Economic Studies*, 91(2), 858–898. <https://doi.org/10.1093/restud/rdad048>
- Couture, V., & Handbury, J. (2020). Urban revival in america. *Journal of Urban Economics*, 119, 103267. <https://doi.org/10.1016/j.jue.2020.103267>
- Couture, V., & Handbury, J. (2023). Neighborhood change, gentrification, and the urbanization of college graduates. *Journal of Economic Perspectives*, 37(2), 29–52. <https://doi.org/10.1257/jep.37.2.29>
- de Bellefon, M.-P., Combes, P.-P., Duranton, G., Gobillon, L., & Gorin, C. (2021). Delineating urban areas using building density. *Journal of Urban Economics*, 125, 103226. <https://doi.org/10.1016/j.jue.2019.103226>
- Delventhal, M., & Parkhomenko, A. (2024). Spatial implications of telecommuting.
- Delventhal, M. J., Kwon, E., & Parkhomenko, A. (2022). JUE insight: How do cities change when we work from home? *Journal of Urban Economics*, 127, 103331. <https://doi.org/10.1016/j.jue.2021.103331>
- Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3), 479–524. <https://doi.org/10.1257/aer.20131706>
- Donaldson, D., & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4), 171–198. <https://doi.org/10.1257/jep.30.4.171>
- Duranton, G., & Puga, D. (2023). Urban growth and its aggregate implications [eprint: <https://onlinelibrary.wiley.com/doi/10.3982/ECTA17936>]. *Econometrica*, 91(6), 2219–2259. <https://doi.org/10.3982/ECTA17936>
- Earth Resources Observation and Science (EROS) Center. (2017). Landsat level-2 united states analysis ready data, collection 1. <https://doi.org/10.5066/F7319TSJ>
- Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the u.s. declined? [Publisher: Elsevier Inc.]. *Journal of Urban Economics*, 102, 76–90. <https://doi.org/10.1016/j.jue.2017.07.002>
- Glaeser, E. L., & Gyourko, J. (2025). America's housing affordability crisis and the decline of housing supply. *BPEA Conference Draft, Spring*.
- Glaeser, E. L., Gyourko, J., & Saks, R. E. (2005). Why have housing prices gone up? *American Economic Review*, 95(2), 329–333. <https://doi.org/10.1257/000282805774669961>
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities [Publisher: University of Chicago Press]. *Journal of Political Economy*, 100(6), 1126–1152. Retrieved October 9, 2025, from <http://www.jstor.org/stable/2138829>

- Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography*, 1(1), 27–50. <https://doi.org/10.1093/jeg/1.1.27>
- Gyourko, J., Saiz, A., & Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index [Publisher: Temporary Publisher]. *Urban Studies*, 45(3), 693–729. <http://www.jstor.org/stable/43197779>
- Harari, M. (2020). Cities in bad shape: Urban geometry in india. *American Economic Review*, 110(8), 2377–2421. <https://doi.org/10.1257/aer.20171673>
- Heblich, S., Redding, S. J., & Sturm, D. M. (2020). The making of the modern metropolis: Evidence from london*. *The Quarterly Journal of Economics*, 135(4), 2059–2133. <https://doi.org/10.1093/qje/qjaa014>
- Hsieh, C. T., & Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2), 1–39. <https://doi.org/10.1257/MAC.20170388>
- Manson, S., Schroeder, J., Van Riper, D., Knowles, K., Kugler, T., Roberts, F., & Ruggles, S. (2023). National historical geographic information system: Version 18.0. <https://doi.org/10.18128/D050.V18.0>
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior [Publisher: Academic Press]. *Frontiers in Econometrics*.
- Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area [Publisher: American Economic Association]. *The American Economic Review*, 57(2), 197–210. Retrieved March 23, 2023, from <https://www.jstor.org/stable/1821621>
- Moreno-Maldonado, A., & Santamaría, C. (2025). *Delayed childbearing and urban revival: A structural approach*.
- Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.
- Muth, R. F. (1969). *Cities and housing: The spatial pattern of urban residential land use*. Univ. of Chicago Pr.
- Redding, S. J., & Rossi-Hansberg, E. (2017). Quantitative spatial economics [Publisher: Annual Reviews]. *Annual Review of Economics*, 9, 21–58. <https://doi.org/10.1146/annurev-economics-063016-103713>
- Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6), 1257–1278. <https://doi.org/10.1086/261120>
- Rollet, V. (2025). Can we rebuild a city? the dynamics of urban redevelopment.
- Rosen, S. (1979). Wage-based indexes of urban quality of life. In P. Mieszkowski & M. Straszheim (Eds.), *Current issues in urban economics* (pp. 74–104). Johns Hopkins University Press.
- Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics*, 125(3), 1253–1296. <https://doi.org/10.1162/qjec.2010.125.3.1253>

- Tsivanidis, N. (2023). Evaluating the impact of urban transit infrastructure: Evidence from Bogotá's TransMilenio. Retrieved October 12, 2023, from https://static1.squarespace.com/static/55bb98e2e4b0ba843f39599e/t/64c98e1aa0fcf82d1bda8d52/1690930717404/TsivanidisTransMillenio_8.2023.pdf
- U.S. Geological Survey. (2025). *Annual NLCD (national land cover database)— the next generation of land cover mapping*. <https://doi.org/10.3133/fs20253001>
- Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., & Van Driel, J. N. (2001). Completion of the 1990s national land cover data set for the conterminous united states from landsat thematic mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, 67(6), 650–662. <https://pubs.usgs.gov/publication/70159374>
- Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G. C., Rigge, M., & Xian, G. (2018). A new generation of the united states national land cover database: Requirements, research priorities, design, and implementation strategies [Publisher: Elsevier]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 108–123. <https://doi.org/10.1016/j.isprsjprs.2018.09.006>

Appendices

Appendix A Data appendix

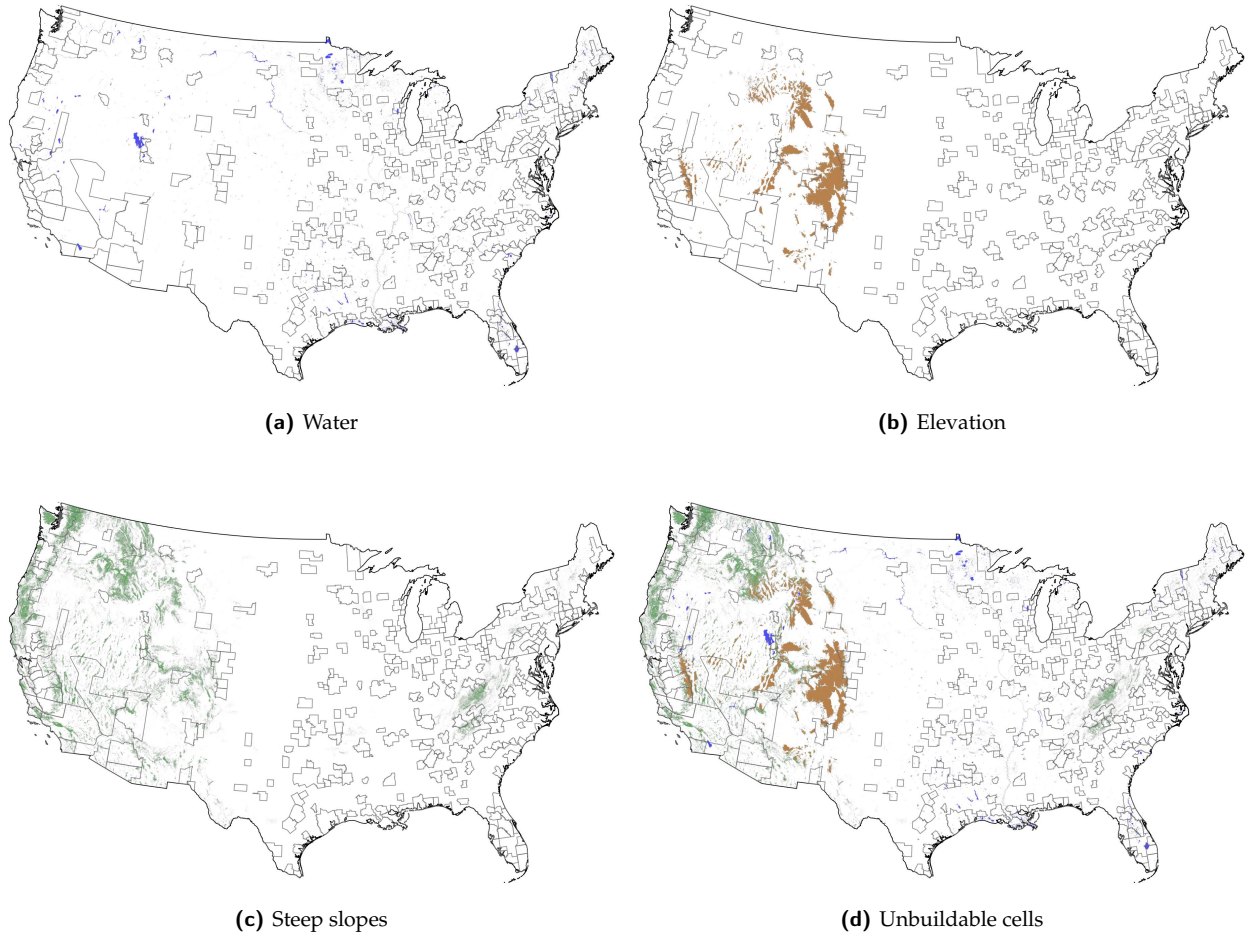


Figure A1: Unbuildable cells across the United States, calculated from the NLCD and the national elevation map. For reference, I include the outlines of the 275 Metropolitan Statistical Areas I use in my analysis.

Appendix B Additional results

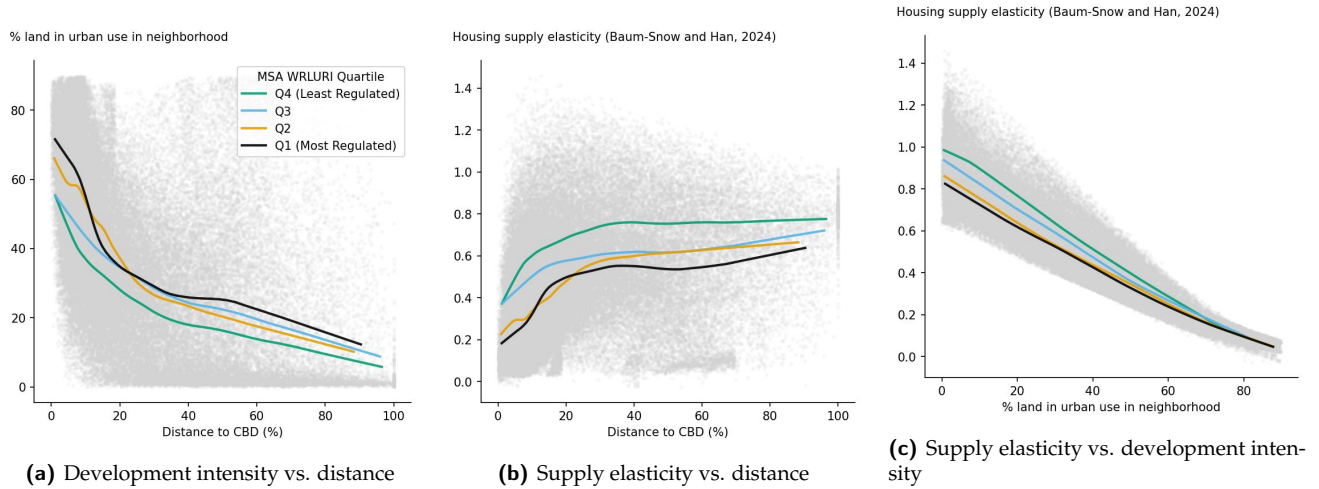


Figure A2: Relationship between development intensity, distance from CBD, and housing supply elasticities, by regulatory environment. Metro areas are grouped by their Wharton Regulatory Land Use Index (WRLURI).

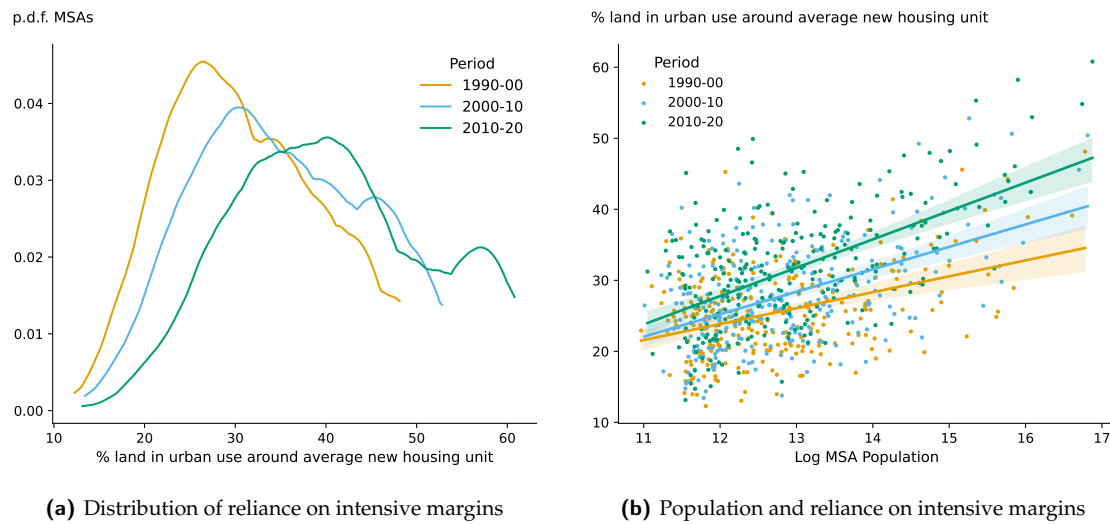


Figure A3: Divergence in reliance on intensive margins of development across metro areas. Panel (a) shows the distribution of MSA-level average land use intensity around new housing units built during each decade 1990-2000, 2000-2010, and 2010-2020. Panel (b) plots the relationship between log metro population in 1990 and the mean share of developed land around new housing units built during the 1990s, revealing that larger metros systematically rely more heavily on infill and redevelopment.

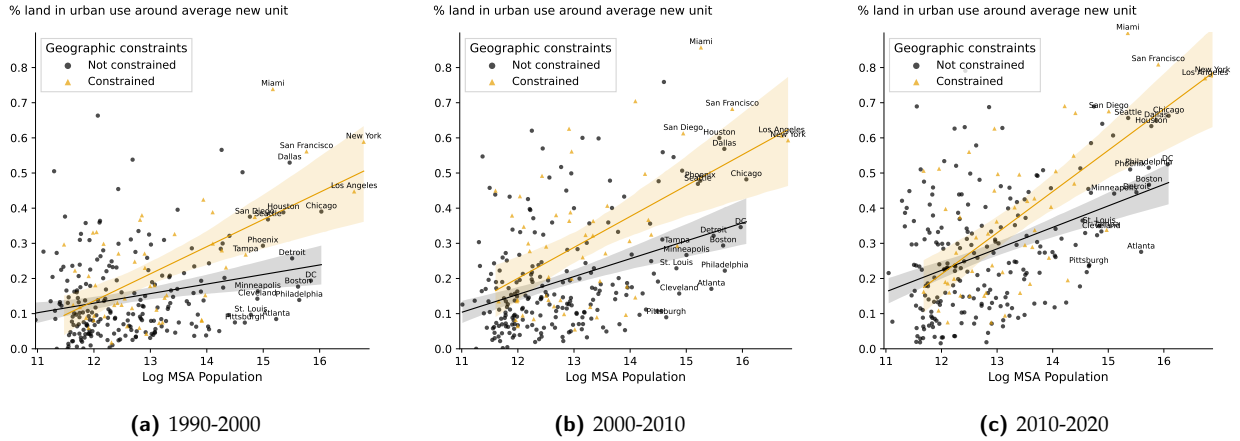


Figure A4: Redevelopment by geographically constrained and unconstrained metro areas. Metro areas are split by whether they are geographically constrained to development. Geographic constraints are calculated similarly to Saiz (2010), by drawing a 50 kilometer radius around the CBD and calculating the share of developable land. I set as “geographically constrained” those where at least 20% of their landmass is not developable due to elevation, steep slopes, or bodies of water. Note how, on average, big cities with geographic constraints *do* redevelop more than those without, but that big cities tend to redevelop more on average. If we pay a bit of attention to the cities in the plot, it can also be argued that most of the big, geographically constrained cities are also the ones that had historical downtowns. All of this is addressed in the quantitative model.

Note: Some geographically constrained cities are New York, Los Angeles, San Francisco, Miami, San Diego, Tampa ; some unconstrained cities are Chicago, DC, Philadelphia, Boston, Detroit, Dallas, Houston, Seattle, Atlanta, Cleveland, Minneapolis, St. Louis, Pittsburgh, Phoenix



Figure A5: Evolution of the share of new housing built in neighborhoods at different distances from the CBD, measured as the fraction of total distance to the urban edge (weighted by housing units). Each panel shows the distribution for a different metro size group: (a) metros with more than 3 million inhabitants in 1990; (b) metros with 800,000 to 3 million inhabitants; (c) metros with fewer than 800,000 inhabitants. Larger metros increasingly concentrated new construction closer to downtown, while smaller metros continued to build primarily at the urban fringe.

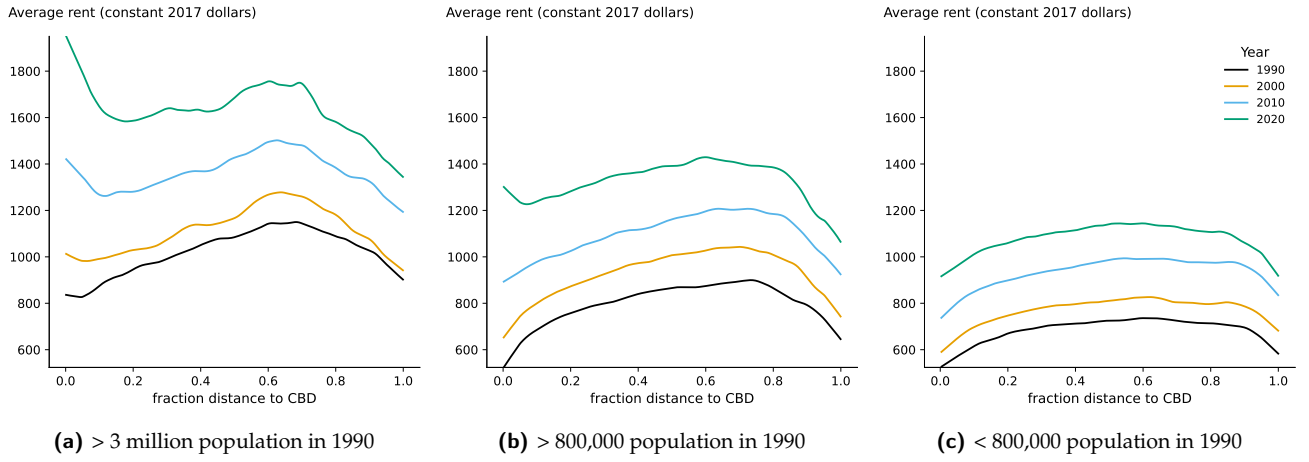


Figure A6: Evolution of housing costs (monthly rent) as a nonparametric function of distance from the CBD, measured as fraction of total distance to the urban edge (weighted by housing units). Each panel shows the relationship for a different metro size group: (a) metros with more than 3 million inhabitants in 1990; (b) metros with 800,000 to 3 million inhabitants; (c) metros with fewer than 800,000 inhabitants.

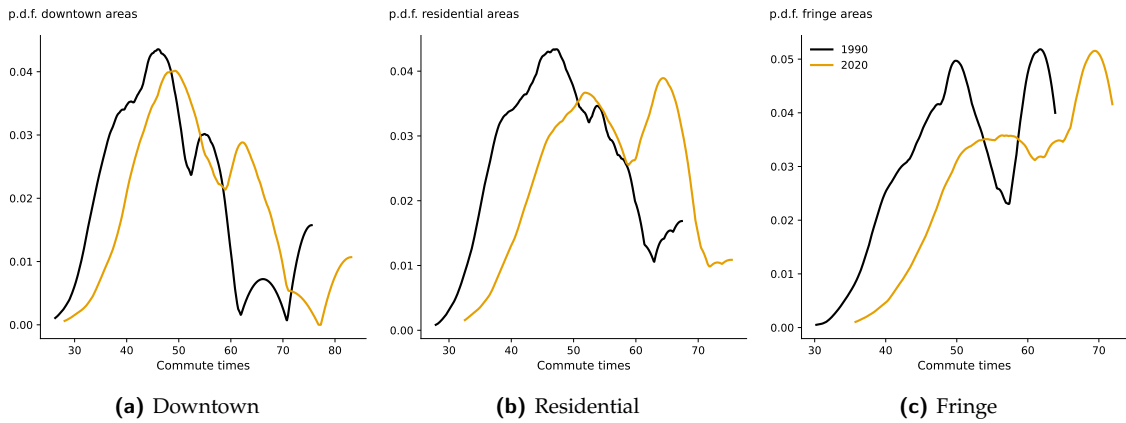


Figure A7: Evolution of actual commuting costs from Census data by neighborhood type, weighted by number of households. Each panel shows the evolution of observed commuting costs for a different neighborhood type: (a) Downtown, (b) Residential, and (c) Fringe.

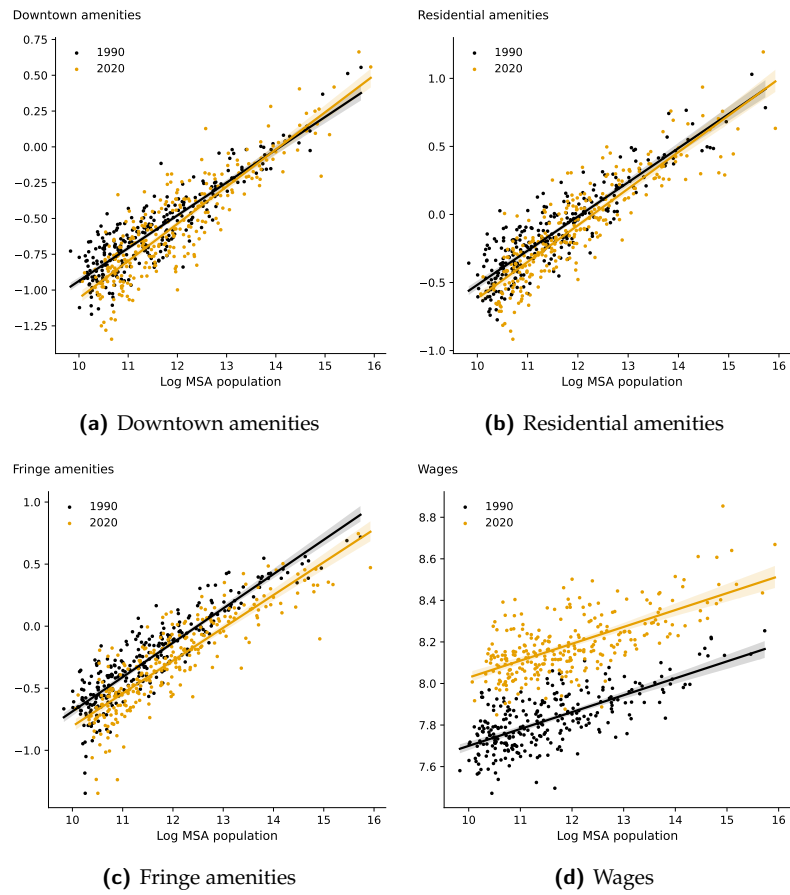


Figure A8: Relationship between metro population and demand fundamentals: amenities in each zone, and wages, for 1990 and 2020. Note how it can be observed that downtown amenities became more attractive in big cities over time.

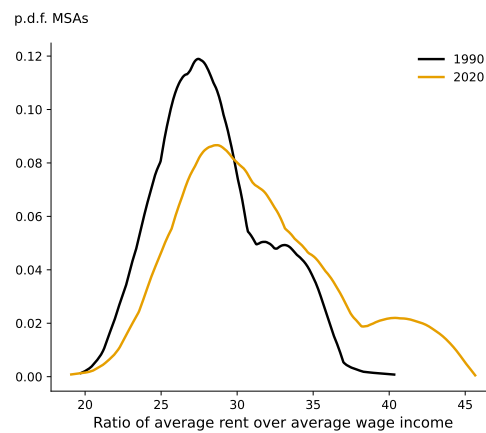


Figure A9: Distribution of MSA-level housing expenditure shares (Average rent over average post-tax wage income) in 1990 and 2020.

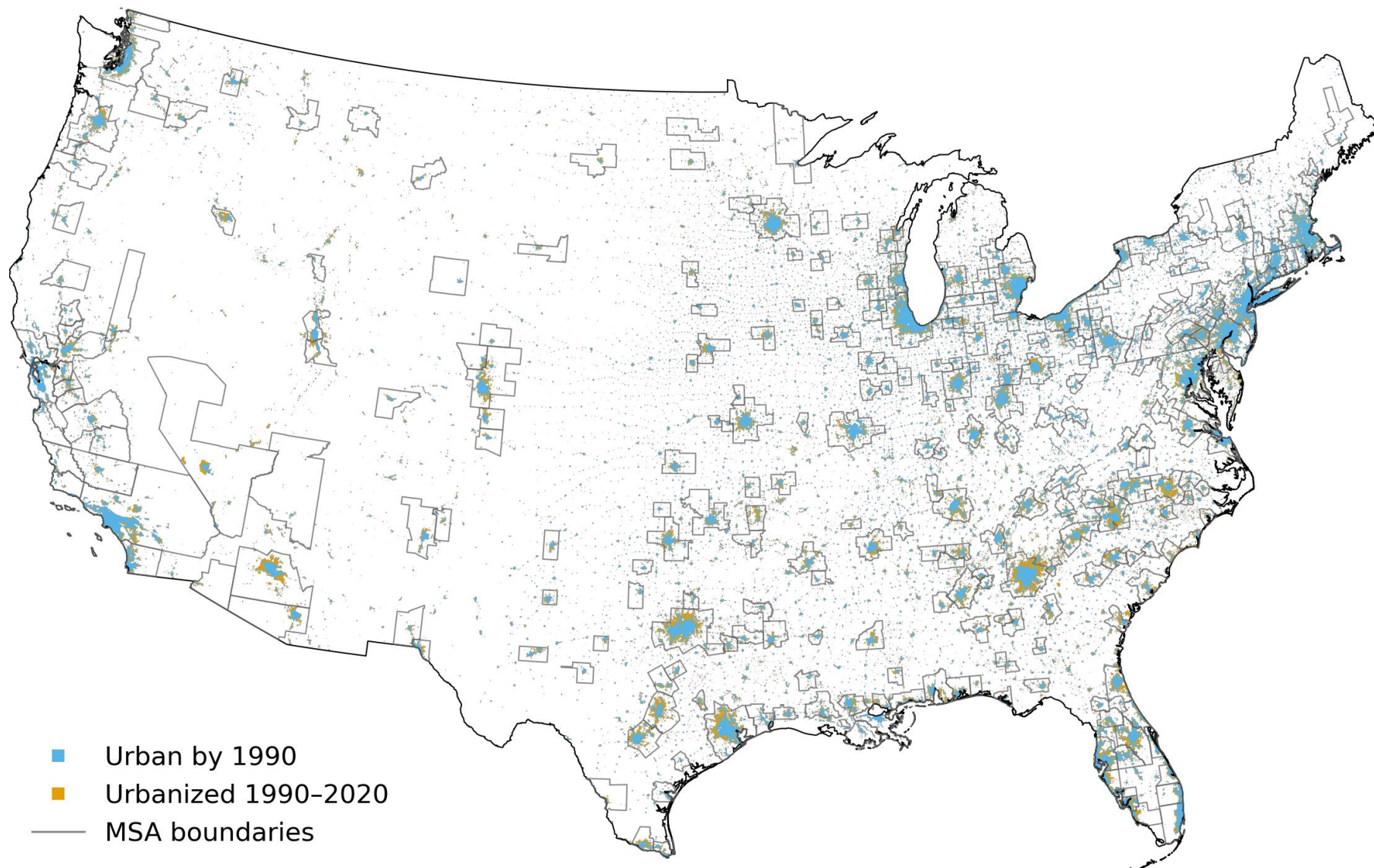


Figure A10: The growth of urban areas in the United States, 1990-2020

Appendix C Derivations from the model

Proof of Proposition 1. For any two locations r and r' within the same neighborhood i , we have:

$$V_{im}(\omega, r) = \left(w_m - \tau N_m^\phi r^\gamma - P_{im}(r) \right) \cdot A_{im} \cdot v_{im}(\omega) \quad (20)$$

Substituting the bid-rent condition $P_{im}(r) = P_{im}(r') + \tau N_m^\phi (r'^\gamma - r^\gamma)$:

$$\begin{aligned} V_{im}(\omega, r) &= \left(w_m - \tau N_m^\phi r^\gamma - \left(P_{im}(r') + \tau N_m^\phi (r'^\gamma - r^\gamma) \right) \right) \cdot A_{im} \cdot v_{im}(\omega) \\ &= \left(w_m - \tau N_m^\phi r^\gamma - P_{im}(r') - \tau N_m^\phi r'^\gamma + \tau N_m^\phi r^\gamma \right) \cdot A_{im} \cdot v_{im}(\omega) \\ &= \left(w_m - P_{im}(r') - \tau N_m^\phi r'^\gamma \right) \cdot A_{im} \cdot v_{im}(\omega) \\ &= V_{im}(\omega, r') \end{aligned}$$

Therefore, utility is constant across all locations within the neighborhood. $\square$

Proof of Proposition 2. (i) The Fringe-Rural Boundary (r_{Fm}): At the city's edge, land is converted from rural use, which has an opportunity cost normalized to zero. For developers to be willing to build, the profit from fringe development must be non-negative. The boundary r_{Fm} is where profit is exactly zero: $\Pi_{Fm}(r_{Fm}) = \text{Dens}_{Fm}(P_{Fm}(r_{Fm}) - c_{Fm}) = 0$. Since $\text{Dens}_{Fm} > 0$, this implies $P_{Fm}(r_{Fm}) = c_{Fm}$.

(ii) The Residential-Fringe Boundary (r_{Rm}): Developers are indifferent between residential and fringe use at r_{Rm} , so their profits are equal:

$$\text{Dens}_{Rm}(P_{Rm}(r_{Rm}) - c_{Rm}) = \text{Dens}_{Fm}(P_{Fm}(r_{Rm}) - c_{Fm})$$

Rearranging to solve for $P_{Rm}(r_{Rm})$ gives:

$$P_{Rm}(r_{Rm}) = c_{Rm} + \frac{\text{Dens}_{Fm}}{\text{Dens}_{Rm}} (P_{Fm}(r_{Rm}) - c_{Fm})$$

From the bid-rent condition (equation 4), we can express the fringe price at r_{Rm} in terms of the price at r_{Fm} :

$$P_{Fm}(r_{Rm}) = P_{Fm}(r_{Fm}) + \tau N_m^\phi (r_{Fm}^\gamma - r_{Rm}^\gamma)$$

Substitute $P_{Fm}(r_{Fm}) = c_{Fm}$ from part (i) yields:

$$P_{Fm}(r_{Rm}) = c_{Fm} + \tau N_m^\phi (r_{Fm}^\gamma - r_{Rm}^\gamma)$$

Plugging this back into the expression for $P_{Rm}(r_{Rm})$:

$$P_{Rm}(r_{Rm}) = c_{Rm} + \frac{Dens_{Fm}}{Dens_{Rm}} \left(\left[c_{Fm} + \tau N_m^\phi (r_{Fm}^\gamma - r_{Rm}^\gamma) \right] - c_{Fm} \right)$$

The fringe construction costs, c_{Fm} , cancel, yielding the expression in the proposition.

(iii) The Downtown-Residential Boundary (r_{Dm}): At r_{Dm} , the profit from downtown development equals the profit from residential development: $\Pi_{Dm}(r_{Dm}) = \Pi_{Rm}(r_{Dm})$. Substituting the profit functions gives:

$$\frac{\eta_{Dm}}{\eta_{Dm} + 1} \varepsilon_{Dm}^{-1/\eta_{Dm}} P_{Dm}(r_{Dm})^{1+1/\eta_{Dm}} - c_{Dm} = Dens_{Rm} (P_{Rm}(r_{Dm}) - c_{Rm})$$

The right-hand side, $\Pi_{Rm}(r_{Dm})$, represents the value of land for residential use at this boundary. We can express it in terms of model primitives by linking prices across boundaries. First, using the bid-rent condition (equation 4), we link $P_{Rm}(r_{Dm})$ to $P_{Rm}(r_{Rm})$:

$$P_{Rm}(r_{Dm}) = P_{Rm}(r_{Rm}) + \tau N_m^\phi (r_{Rm}^\gamma - r_{Dm}^\gamma)$$

Substitute the expression for $P_{Rm}(r_{Rm})$ from part (ii):

$$P_{Rm}(r_{Dm}) = \left[c_{Rm} + \frac{Dens_{Fm}}{Dens_{Rm}} \tau N_m^\phi (r_{Fm}^\gamma - r_{Rm}^\gamma) \right] + \tau N_m^\phi (r_{Rm}^\gamma - r_{Dm}^\gamma)$$

Now substitute this into the residential profit term $\Pi_{Rm}(r_{Dm}) = Dens_{Rm}(P_{Rm}(r_{Dm}) - c_{Rm})$:

$$\Pi_{Rm}(r_{Dm}) = Dens_{Rm} \left(\left[c_{Rm} + \frac{Dens_{Fm}}{Dens_{Rm}} \tau N_m^\phi (r_{Fm}^\gamma - r_{Rm}^\gamma) + \tau N_m^\phi (r_{Rm}^\gamma - r_{Dm}^\gamma) \right] - c_{Rm} \right)$$

The residential construction costs, c_{Rm} , cancel. Distributing $Dens_{Rm}$ gives the total land value at r_{Dm} :

$$\Pi_{Rm}(r_{Dm}) = Dens_{Fm} \tau N_m^\phi (r_{Fm}^\gamma - r_{Rm}^\gamma) + Dens_{Rm} \tau N_m^\phi (r_{Rm}^\gamma - r_{Dm}^\gamma)$$

Substituting this back into the indifference condition:

$$\frac{\eta_{Dm}}{\eta_{Dm} + 1} \varepsilon_{Dm}^{-1/\eta_{Dm}} P_{Dm}(r_{Dm})^{\frac{\eta_{Dm}+1}{\eta_{Dm}}} - c_{Dm} = \tau N_m^\phi [Dens_{Fm} (r_{Fm}^\gamma - r_{Rm}^\gamma) + Dens_{Rm} (r_{Rm}^\gamma - r_{Dm}^\gamma)]$$

Finally, solving for $P_{Dm}(r_{Dm})$ by isolating the term and raising both sides to the appropriate power gives the expression in the proposition. $\square$

Appendix D Computational Algorithms

D.1 Model Inversion to Recover Amenities

To recover the unobserved neighborhood-level amenities, we invert the model. This procedure takes as given the full set of model parameters $\{\tau, \phi, \gamma, \theta, \psi, c_{Dm}, c_{Rm}, c_{Fm}, \text{Dens}_{Rm}, \text{Dens}_{Fm}, \varepsilon_{Dm}, \eta_{Dm}\}$, the observed data on population allocations $\{N_{im}^{\text{data}}\}$ and neighborhood boundaries $\{r_{Dm}^{\text{data}}, r_{Rm}^{\text{data}}, r_{Fm}^{\text{data}}\}$, and the metro-level fundamentals, specifically the set of wages $\{w_m\}$ and developable land arcs $\{\alpha_m\}$. The algorithm finds the unique vector of amenities, $\{A_{im}\}$, that ensures the model's predicted household location choices are perfectly consistent with the observed choices in the data. This is accomplished using a contraction mapping algorithm that iteratively adjusts an initial guess for the amenities until the model's implied population shares converge to the observed population shares.

The algorithm proceeds as follows:

1. **Initialization.** We start with the given inputs. From the observed populations N_{im}^{data} , we compute the observed population shares, $\pi_{im}^{\text{data}} = N_{im}^{\text{data}} / \sum_{j,n} N_{jn}^{\text{data}}$, and the observed metro-level populations, $N_m^{\text{data}} = \sum_i N_{im}^{\text{data}}$. We make an initial guess for the vector of amenities, $A_{im}^{(0)}$, typically by setting $A_{im}^{(0)} = 1$ for all i, m . A tolerance level ϵ and a damping parameter $\lambda \in (0, 1]$ are set for the iterative procedure.
2. **Iterative Loop.** For each iteration $k = 0, 1, 2, \dots$:
 - (a) Using the *observed* boundaries r_{im}^{data} and metro populations N_m^{data} , we compute the reference housing prices $\hat{p}_{im}^{\text{data}}$ and commuting costs $\hat{\tau}_{im}^{\text{data}}$ for each neighborhood, following the characterization in Proposition 2. These values are treated as fixed throughout the inversion.
 - (b) We combine these values with the current amenity guess $A_{im}^{(k)}$ to calculate the mean utility component for each choice:

$$\delta_{im}^{(k)} = \left(w_m - \hat{\tau}_{im}^{\text{data}} - \hat{p}_{im}^{\text{data}} \right) \cdot A_{im}^{(k)}$$

- (c) Using the vector of mean utilities $\{\delta_{im}^{(k)}\}$, we compute the model's implied choice probabilities, $\pi_{im}^{(k)}$, according to the GEV structure in Equation 6.
- (d) We check for convergence by comparing the model-implied shares with the data. If $\max_{i,m} |\pi_{im}^{(k)} - \pi_{im}^{\text{data}}| < \epsilon$, the algorithm has converged, and the solution is $A_{im}^* = A_{im}^{(k)}$.

(e) If convergence is not achieved, we update the amenity guess using the following rule:

$$A_{im}^{(k+1)} = A_{im}^{(k)} \cdot \left(\frac{\pi_{im}^{\text{data}}}{\pi_{im}^{(k)}} \right)^\lambda$$

This update increases the amenity guess for locations that are more popular in the data than predicted by the model (and vice-versa), ensuring that the algorithm moves towards the true values. The loop then proceeds to iteration $k + 1$.

D.2 Numerical Solution Algorithm

I solve the spatial equilibrium system using a trust-region quasi-Newton method operating in log-population space. The key innovation is exploiting a monotonic mapping between housing quantities and neighborhood boundaries that allows me to reformulate the equilibrium problem in terms of populations rather than radii directly.

D.2.1 Population-to-Boundary Inversion

The model's supply-side conditions establish a one-to-one mapping between housing quantities $\{H_{Dm}, H_{Rm}, H_{Fm}\}$ and neighborhood radii $\{r_{Dm}, r_{Rm}, r_{Fm}\}$ for each metro area m . Given target populations, I invert this mapping through the following procedure:

For the fixed-density Residential and Fringe rings, the relationship between housing stock and radii is direct. The Residential ring occupies the annulus $[r_{Dm}, r_{Rm}]$, so:

$$H_{Rm} = \text{Dens}_{Rm} \cdot \frac{\alpha_m}{2} (r_{Rm}^2 - r_{Dm}^2) \quad (21)$$

Given H_{Rm}^{target} and a trial value for r_{Dm} , I can solve directly for r_{Rm} :

$$r_{Rm} = \sqrt{r_{Dm}^2 + \frac{2H_{Rm}^{\text{target}}}{\text{Dens}_{Rm} \cdot \alpha_m}} \quad (22)$$

and analogously for r_{Fm} given r_{Rm} .

The Downtown radius r_{Dm} is determined by matching the Downtown housing target H_{Dm}^{target} through bisection. For each trial r_{Dm} :

1. Compute r_{Rm} and r_{Fm} from the ring area formulas above
2. Evaluate the implied Downtown housing stock by solving an inner fixed-point problem for

metro population, which yields:

$$H_{Dm}(r_{Dm}) = \alpha_m \int_0^{r_{Dm}} \left(\frac{P_{Dm}(r)}{\varepsilon_{Dm}} \right)^{1/\eta_{Dm}} r dr \quad (23)$$

where $P_{Dm}(r) = P_{Dm}(r_{Dm}) + \tau N_m^\phi (r_{Dm}^\gamma - r^\gamma)$ from the bid-rent condition in equation (4), and $P_{Dm}(r_{Dm})$ is given by equation (14)

3. Update the bisection bounds on r_{Dm} based on whether $H_{Dm}(r_{Dm})$ exceeds or falls short of H_{Dm}^{target}

This procedure converges rapidly because the mapping from r_{Dm} to H_{Dm} is monotonic: larger Downtown radii necessarily produce more Downtown housing.

D.2.2 Quasi-Newton Solver in Log-Population Space

I reformulate the equilibrium conditions as a rootfinding problem in the log-population vector $z \in \mathbb{R}^{3M}$, where $z = \log(H)$ stacks $[\log H_{D1}, \log H_{R1}, \log H_{F1}, \dots, \log H_{DM}, \log H_{RM}, \log H_{FM}]$. The residual function is:

$$F(z) = \log H^{\text{supply}}(z) - \log H^{\text{demand}}(z) \quad (24)$$

where $H^{\text{supply}}(z)$ is obtained by (i) inverting $\exp(z)$ to radii via the procedure above, (ii) computing supply at those radii through equations (14)–(??) and the integral in the second step above, and $H^{\text{demand}}(z)$ comes from the nested logit choice probabilities in equation (6) multiplied by total population N .

The solver minimizes $\|F(z)\|$ using a trust-region method with rank-one Jacobian updates. The initial Jacobian is computed via parallel finite differences, exploiting the structure that perturbing metro m 's populations only requires re-solving that metro's boundary inversion. Between Jacobian rebuilds, I update the approximate Jacobian using a quasi-Newton formula:

$$J_{k+1} = J_k + \frac{(F_{k+1} - F_k - J_k s_k) s_k^T}{s_k^T s_k} \quad (25)$$

where $s_k = z_{k+1} - z_k$ is the accepted step. The dogleg trust-region method combines a steepest descent direction with a Newton step, with the trust region radius adjusted based on the ratio of actual to predicted reduction in $\|F\|^2$.

Working in log-population space rather than log-radius space offers two advantages: (1) the population targets directly correspond to the equilibrium conditions we wish to satisfy, avoiding the need to differentiate through the complex supply integral, and (2) the monotonic population-

boundary mapping ensures that the inversion step is always well-defined, preventing the numerical instabilities that can arise when radii become mis-ordered.

Appendix E Model Fit

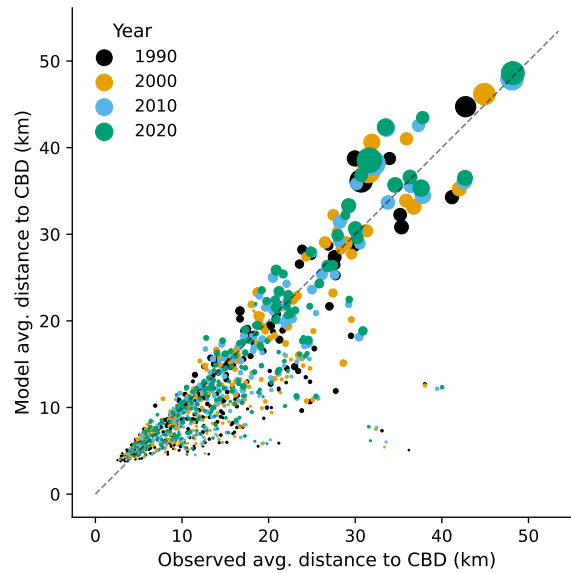


Figure A11: Model fit for average distance to CBD. Each point represents a metro area, comparing the observed average distance to the central business district with the model-predicted value. The size of each point is proportional to the metro's number of households. Some small metros have a bad fit because the MSA contains multiple small urban areas that are not well captured by the monocentric model.

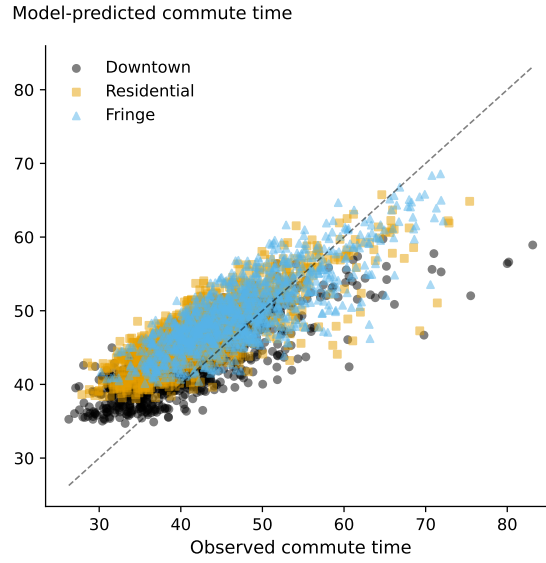


Figure A12: Fit between model-predicted and observed commuting times. Each point represents a metro-zone-year observation. The model's predicted commuting times (based on the the zones' boundaries, calibrated τ parameter and metro-specific congestion) are plotted against actual commuting times as measured in tract-level Census and ACS data.

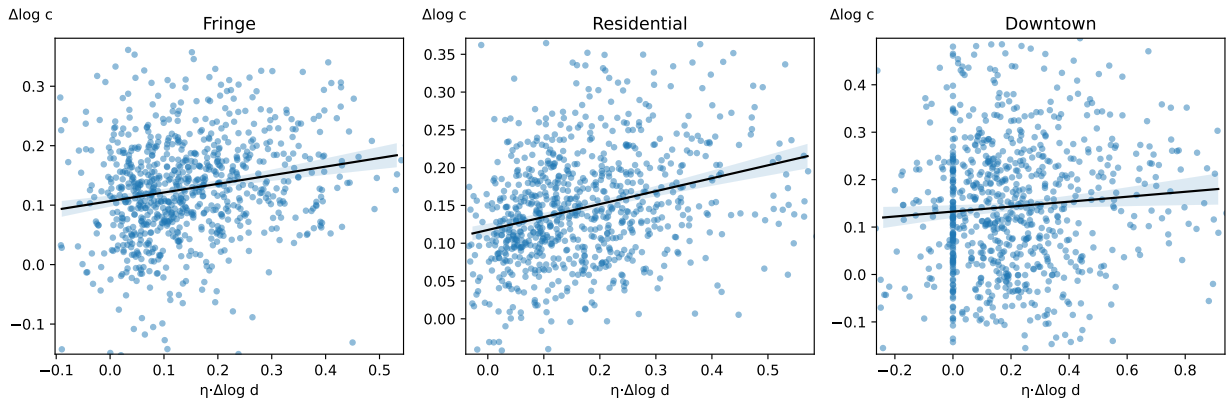


Figure A13: Fit between model-implied changes in c_{im} due to the growth in neighborhood boundaries, and the true growth in c_{im} as inverted from the data. Each point is a metro-neighborhood-year observation. As expected, there is a high correlation, but the slope is less than one: neighborhoods with high levels of growth tend to have positive unobserved supply shocks.

Appendix F De-regulation counterfactuals

F.1 Reduced-form model without within-metro geography

This appendix presents a reduced-form version of the model without within-metro geography. Each metro $m \in \mathcal{M}$ is a single alternative with one housing market and one price p_m . Households choose across metros and a rural outside option; housing is supplied by competitive developers with an inverse supply elasticity η_m .

F.1.1 Households

There is a population N of identical households indexed by ω , who choose a location among the set of metros $\mathcal{M}$ or a rural outside option. The deterministic component of indirect utility for metro m is

$$v_m(p_m) = (w_m - \tau_m - p_m) A_m, \quad (26)$$

where w_m is the wage, τ_m is a fixed commuting cost, p_m is the metro housing price, and A_m is an amenity shifter.

Idiosyncratic tastes are drawn from a GEV structure that delivers across-metro substitution parameter $\psi > 0$. Let v_R denote the deterministic value of the rural option. The resulting choice probabilities are

$$\pi_{m|Urban} = \frac{v_m(p_m)^\psi}{\sum_{k \in \mathcal{M}} v_k(p_k)^\psi}, \quad (27)$$

$$\pi_{Urban} = \frac{\sum_{k \in \mathcal{M}} v_k(p_k)^\psi}{\sum_{k \in \mathcal{M}} v_k(p_k)^\psi + v_R^\psi}, \quad (28)$$

$$\pi_m = \pi_{m|Urban} \pi_{Urban} = \frac{v_m(p_m)^\psi}{\sum_{k \in \mathcal{M}} v_k(p_k)^\psi + v_R^\psi}. \quad (29)$$

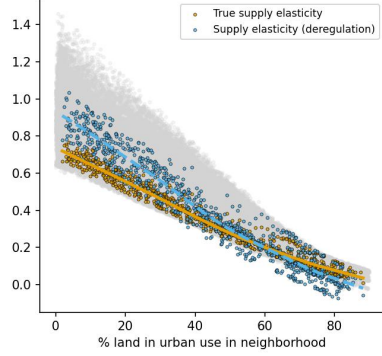
Demand in metro m is then $N_m^d = N \pi_m$.

F.1.2 Housing Supply in Each Metro

Housing is supplied by competitive developers. Around the previous-period baseline $(q_{m,t}, p_{m,t})$, the inverse supply is log-linear with (inverse) elasticity $\eta_m > 0$ and a supply shock ε_m :

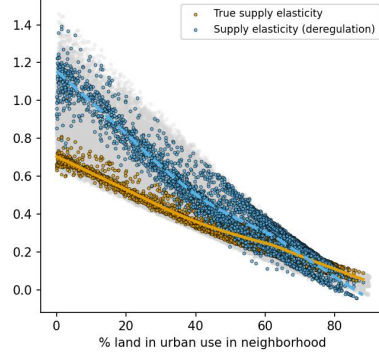
$$p_m(q_m) = p_{m,t} \exp\left(\eta_m [\log q_m - \log q_{m,t}] + \varepsilon_m\right). \quad (30)$$

Housing supply elasticity (Baum-Snow and Han, 2024)



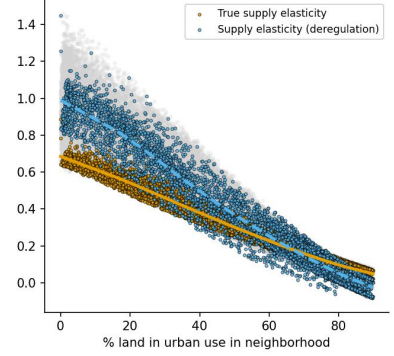
(a) Boston

Housing supply elasticity (Baum-Snow and Han, 2024)



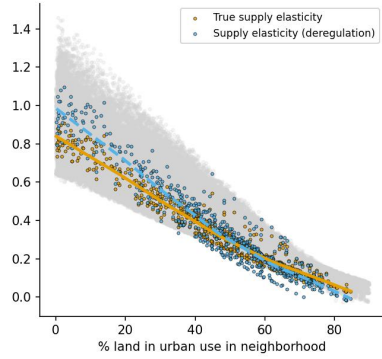
(b) Los Angeles

Housing supply elasticity (Baum-Snow and Han, 2024)



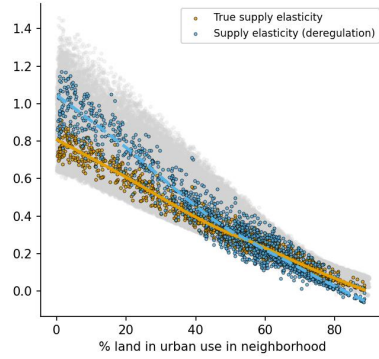
(c) New York

Housing supply elasticity (Baum-Snow and Han, 2024)



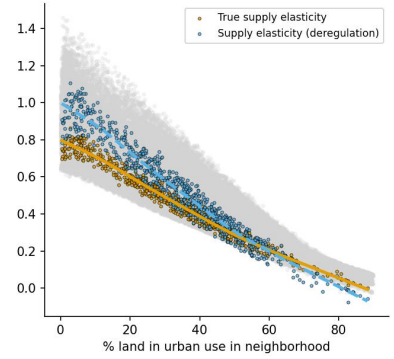
(d) San Diego

Housing supply elasticity (Baum-Snow and Han, 2024)



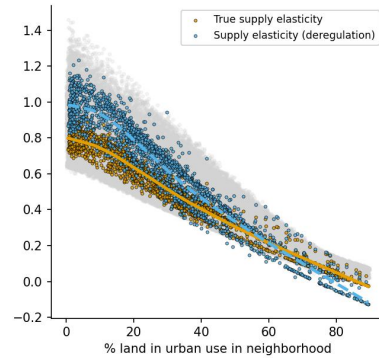
(e) San Francisco

Housing supply elasticity (Baum-Snow and Han, 2024)



(f) Seattle

Housing supply elasticity (Baum-Snow and Han, 2024)



(g) Washington DC

Figure A14: Housing supply elasticity schedules for the seven metro areas I de-regulate in counterfactual simulations, as a function of neighborhood fraction developed. Each panel shows the *true* elasticity schedule, and the counterfactual schedule under de-regulation.

Equivalently, in levels,

$$p_m(q_m) = \kappa_{m,t} e^{\varepsilon_m} q_m^{\eta_m}, \quad \kappa_{m,t} \equiv p_{m,t} q_{m,t}^{-\eta_m}. \quad (31)$$

F.1.3 Equilibrium

Given fundamentals $\{w_m, \tau_m, A_m\}_{m \in \mathcal{M}}$, the rural value v_R , population N , the previous-period anchors $\{q_{m,t}, p_{m,t}\}$, and shocks $\{\varepsilon_m\}$, a (spatial) equilibrium is a vector $\{q_m, p_m\}_{m \in \mathcal{M}}$ such that, for every metro m :

$$\text{(Supply)} \quad p_m = p_{m,t} \exp\left(\eta_m [\log q_m - \log q_{m,t}] + \varepsilon_m\right), \quad (32)$$

$$\text{(Demand)} \quad q_m = N \frac{v_m(p_m)^\psi}{\sum_{k \in \mathcal{M}} v_k(p_k)^\psi + v_R^\psi}, \quad (33)$$

with $v_m(p_m) = (w_m - \tau_m - p_m)A_m$. Equations (32)–(33) jointly determine prices $\{p_m\}$ and quantities $\{q_m\}$ across all metros.

(a) Effect of deregulation on housing quantities and population

	Observed		De-regulation	
	Households by 1990 (thousands)	Households by 2020 (thousands)	Additional growth 1990-2020	Households by 2020 (thousands)
New York	6,786	8,270	-6.2%	8,178
Los Angeles	5,178	6,503	83.9%	7,615
Washington	2,414	3,581	11.7%	3,718
San Francisco	2,404	3,020	85.4%	3,546
Boston	2,129	2,686	-4.0%	2,664
Seattle	1,135	1,871	67.0%	2,364
San Diego	910	1,201	88.4%	1,458
Rest of cities	50,289	76,628	-4.0%	75,569
Rural areas	30,391	35,867	-24.7%	34,515

(b) Effect of deregulation on housing costs

	Observed		De-regulation	
	Avg. rent by 1990 (2017 \$)	Avg. rent by 2020 (2017 \$)	Additional growth 1990-2020	Avg. rent by 2020 (2017 \$)
New York	1,070	1,711	-5.2%	1,678
Los Angeles	1,171	1,932	-19.9%	1,781
Washington	1,094	1,748	-13.6%	1,659
San Francisco	1,245	2,375	-25.6%	2,086
Boston	1,051	1,567	-6.8%	1,532
Seattle	888	1,802	-34.1%	1,490
San Diego	1,131	2,097	-20.7%	1,897
Rest of cities	770	1,197	-4.6%	1,177
Rural areas	—	—	—	—

Table 4: Counterfactual results from relaxing zoning in seven cities with strict housing regulations in the naive model. Zoning relaxation is simulated by setting metro-level supply elasticities to the 75th percentile nationally, using aggregate supply elasticity estimates from Baum-Snow and Han (2024). Panel (a) shows the effect on housing quantities, while panel (b) shows the effect on housing costs.