The Effects of Residential Zoning in U.S. Housing Markets

Jaehee Song*

April 18, 2024

Abstract

I construct a new nationwide dataset to measure the stringency of residential zoning in the U.S. and use it to examine the effects of zoning on housing prices and demographic sorting. First, I develop and apply a structural break detection algorithm to infer minimum lot size regulations from property tax records. These minimum lot size estimates cover 16,217 local jurisdictions and are geographically detailed, capturing both across-jurisdiction and within-jurisdiction variations in zoning stringency. I then use this constructed data and a spatial discontinuity design at municipality borders to evaluate the impact of the regulations in housing markets. I find that a larger minimum lot size increases home prices and rents. For example, doubling the minimum lot size increases sales prices by 10 percent and rents by 6 percent. I also find that neighborhoods with restrictive zoning disproportionately attract high-income white homeowners, intensifying residential segregation.

JEL Classification: R0, H7, C8, K2

Keywords: Zoning laws, minimum lot size, housing affordability, segregation

---

*University of Colorado, Leeds School of Business (email: jaehee.song@colorado.edu). I gratefully acknowledge financial support from the Tobin Center for Economic Policy at Yale University. I thank Steven Berry, John Eric Humphries, and Costas Meghir for continuous support and helpful comments throughout the project. I thank Labor/Public Finance and Industrial Organization groups at Yale University. I also thank seminar and conference participants at AREUEA, BFI Women in Empirical Microeconomics, NBER Summer Institute, UEA, Arizona State University, Federal Housing Finance Agency, Federal Reserve Bank of Cleveland, Federal Reserve Bank of Minneapolis, Harvard Kennedy School, MIT Center for Real Estate, NYU Furman Center, Seoul National University, Syracuse University, University of British Columbia, University of Colorado Boulder, University of Toronto, Wharton Real Estate, and Yonsei University.

1 Introduction

Local governments in the U.S. regulate housing supply through residential zoning, which restricts the quantity and type of housing constructed. A large and growing body of empirical literature finds that local zoning laws increase housing prices and intensify segregation (Glaeser and Ward, 2009; Kok et al., 2014; Trounstine, 2020). Accordingly, recent shortages of affordable housing have triggered nationwide debates on the relaxation of local zoning laws.[1] However, existing research is based on either case studies or surveys with limited geographic coverage, and there is a dearth of comprehensive evidence of the stringency of zoning laws and their effects on housing markets across the nation.

Zoning data is often maintained by local zoning authorities, and there are no nationwide datasets. Hence, existing empirical studies on zoning laws rely on land use survey data (Turner et al., 2014; Trounstine, 2020; Gyourko and Krimmel, 2021, for example) or manual collection of local zoning data (Bronin, 2021; Sahn, 2021, for example). These approaches are very costly and almost implausible at the national level, with thirty thousand local zoning authorities. For example, Wharton Land Use Survey, the largest and most commonly used data source in the literature, only includes 2,450 municipalities among more than two thousand of them across the nation (Gyourko et al., 2021). As such, existing empirical studies of zoning laws have sparse geographic coverage.

This paper addresses this gap in the literature by providing the first nationwide examination of minimum lot size regulations, one of the most common types of density restrictions in residential zoning. I construct a new geographically-detailed dataset on minimum lot sizes covering 52 million single-family homes and 16,217 municipalities in 825 Core-Based Statistical Areas in the United States, which covers more than 60% of households. Using this data and adopting a spatial discontinuity design, I analyze the current restrictiveness of the regulations and their impact on housing prices and residential sorting.

In the first part of the paper, I develop and apply an algorithm that detects minimum lot sizes

1. Some states have already passed statewide relaxation of zoning, including Oregon in 2019 (House Bill 2001) and California in 2021 (Senate Bill 9). In 2021, the federal government included a proposal in the infrastructure plan to provide incentives to local communities to eliminate local exclusionary zoning practices.

from observed building characteristics in property tax records. The key idea is that minimum lot size restrictions make it substantially more difficult to build a house on a lot smaller than the cutoff.[2] If the restriction is a meaningful constraint, then lot sizes of new constructions will bunch right at the cutoff. I thus detect the point of bunching in the empirical distribution of newly constructed lot sizes using a structural break detection algorithm and define it as the minimum lot size.[3] I implement this approach at the zoning district level, which I observe in property tax records in some municipalities and proxy with Census block groups otherwise. I then validate the minimum lot size estimates by comparing them with MAPC Zoning Atlas. I find that the estimated minimum lot sizes reflect the current minimum lot sizes very precisely, with a median error of 11.7%.

The constructed minimum lot size data measures the stringency of housing density restrictions at zoning district levels in 16,217 municipalities, covering almost 80% of local governments in Core-Based Statistical Areas.[4] This is much more geographically granular, capturing within-municipality variations in zoning, and has broader coverage than any other existing zoning datasets, capturing across-municipality variations.[5] I use the constructed dataset to provide an empirical description of the current state of residential zoning across the U.S. First, minimum lot size restrictions are generally stringent, with a mean of 16,000 square feet. Second, 75% of all municipalities impose minimum lot sizes of 1 acre or larger in at least one district in their jurisdiction. Finally, about 15% of homes constructed since 1940 bunch within 5% of minimum lot sizes. The high bunching rate suggests that the minimum lot size restriction is a binding constraint in lot size decisions.

In the second part of the paper, I estimate the effects of minimum lot size restrictions on home prices and rents by using a boundary discontinuity design at municipality borders. To do

2. Under perfect compliance, minimum lot size would ban construction on a smaller lot than the regulated size. In practice, however, some municipalities allow zoning variances to circumvent existing zoning laws, although they are often an arduous process.

3. The algorithm takes a similar approach to OLS-based structural break detection in Andrews (1993) and Zeileis (2005).

4. To define local municipalities with zoning authority, I construct a map of about 32,000 functioning local governments in the U.S. based on the 2010 Census Guide to State and Local Census Geography.

5. Most land use survey data is at the municipality level covering at most a couple of thousand municipalities. The largest manual compilation of local zoning codebooks is Connecticut Zoning Atlas, which covers 2,260 zoning districts, and there are ongoing efforts to build similar statewide datasets.

so, I compile housing market data from CoreLogic property tax history, CoreLogic deed records, and CoreLogic Multiple Listings Service data and merge the deed data with Home Mortgage Disclosure Act (HMDA) data. This rich data allows me to study the near universe of single-family homes, observing their sales records, rental listings, and homeowner demographics at the transaction level. I adopt a boundary discontinuity design at municipal borders and control for pre-zoning neighborhood characteristics collected from 1940 Full-Count Census data and tax records to address the endogeneity concern in zoning. The controls include the population, demographic compositions, homeownership rate, home values, and rents in 1940. I also control whether the community was developed during the postwar suburbanization, during which zoning laws were most actively adopted, proxied by its existence in the geocoded 1940 Full-Count Census. Finally, I show that my empirical findings are robust to the inclusion of municipality-by-school district fixed effects.

Baseline price effect estimates from the municipal border analysis indicate that doubling the minimum lot size increases sales prices of a home by 10 percent and rents by 6 percent. The price premium of stringent zoning may arise from the following mechanisms. First, a larger minimum lot size requires homes to be bigger, and therefore, housing becomes more expensive due to affected building characteristics. I call this the direct effect. Second, zoning affects neighborhood characteristics and, thus, housing prices. For example, stringent zoning may decrease neighborhood congestion, increase the local tax collection, and change neighborhood demographics. I call this the neighborhood effect. I find that 84 percent of the price effect and 66 percent of the rent effect are attributed to the direct effect and the rest to the neighborhood effect. The statistically significant and positive neighborhood effect indicates that restrictive zoning is positively priced in the markets overall.

Additionally, I analyze how local zoning regulations affect demographic sorting. I repeat the border analysis with homeowner race, ethnicity, and income from CoreLogic-HMDA data as outcome variables. I find that doubling the minimum lot size increases the probability of homeowners being non-Hispanic white by 2.3 percentage points and homeowners' income at mortgage applications by 9.5 percent. That is, zoning laws play a substantial role in shifting neighborhood homeowner characteristics.

This paper contributes to several bodies of literature. First, this paper contributes to the literature on measuring the stringency of zoning laws. One popular data collection method is to conduct surveys. Works that use survey data either focus on individual states such as California (Quigley et al., 2008; Jackson, 2018; Mawhorter and Reid, 2018), or use data that surveys small numbers of jurisdictions across the United States, such as the Wharton Land Use Survey (Gyourko et al., 2008; Gyourko et al., 2021) or the Brookings Survey (Pendall et al., 2006), which respectively surveyed 2,450 and 1,844 jurisdictions. Another way to measure the stringency of land use regulations is to manually collect and aggregate local zoning ordinances. Notably, MAPC Zoning Atlas (2020), Menendian et al. (2020), and Bronin (2021) each built zoning datasets respectively covering 101 jurisdictions in Greater Boston, 100 jurisdictions in the Bay Area, and 180 jurisdictions in Connecticut. Finally, this paper is most closely related to a handful of studies developing a data-driven approach to compile or infer local zoning laws. Zabel and Dalton (2011) use structural break detection to infer changes in minimum lot size regulations in the greater Boston area. Nechamkin and MacDonald (2019) apply a random forest algorithm to predict zoning codes in Washington D.C. Mleczko and Desmond (2023) apply natural language processing to process local zoning codes in 2639 municipalities. Cui (2023) develops a structural break detection algorithm to detect the timing of the adoption of zoning laws. In comparison, I focus on inferring minimum lot sizes at the zoning district level to construct a nationwide dataset with unique geographic coverage and details.

Second, this paper contributes to the literature studying the impact of land use regulations on housing markets. Glaeser and Gyourko (2002), Glaeser and Ward (2009), and Kok et al. (2014) show the positive association between stricter regulations and higher housing prices. Other papers study the positive association of stricter regulations with less land development (Wu and Cho, 2007), higher land prices (Gyourko and Krimmel, 2021), loss in economic output (Hsieh and Moretti, 2019), and welfare loss (Brueckner and Sridhar, 2012; Albouy and Ehrlich, 2018). This paper is especially related to works that use a boundary discontinuity design to address the endogeneity of land use regulations (Kahn et al., 2010; Turner et al., 2014; Kulka, 2019; Anagol et al., 2021; Resseger, 2022; Kulka et al., 2023). I contribute to this literature by compiling a nationwide map of local governments to apply a boundary discontinuity design at municipality borders to estimate the impacts of minimum lot size regulations. Taking advantage of the comprehensiveness of the zoning data, I also show that my empirical results are robust to the inclusion of municipality-by-school district fixed effects.

Finally, this paper contributes to the literature on the relationship between density restrictions in residential zoning and neighborhood demographics (Rothwell and Massey, 2009; Trounstine, 2020; Freemark, 2023). This paper estimates the extent to which neighborhood demographics, including race and income, are shifted by zoning laws using transaction-level homeowner demographics from HMDA. Also, I embed the spatial discontinuity design to address the endogeneity of zoning.

The remainder of the paper is organized as follows. Section 2 describes the algorithm that estimates minimum lot areas from property characteristics data and characterizes the current state of zoning from the constructed dataset. In Section 3, I conduct municipal border analyses on housing prices and demographic sorting. Section 4 concludes.

2 New Nationwide Dataset of Minimum Lot Size Restrictions

Existing zoning datasets mostly rely on survey methods or manual compilation of local zoning laws, but they have several limitations. First, they have narrow and/or granular coverage of jurisdictions.[6] Second, land use surveys omit details of zoning ordinances, making them inadequate for studying specific zoning reforms.[7] Finally, land use surveys often suffer from low response rates and measurement errors (Mleczko and Desmond, 2023). These limitations in zoning datasets have been the biggest challenge in the growing empirical literature on zoning, especially when researchers aim to study a broad geographic region.

To overcome this challenge, I develop a scalable algorithm to detect zoning regulations

6. Data scarcity is especially relevant when data is manually collected; manual compilation of local zoning laws is very expensive and implausible at the national level, as the United States has over 30,000 zoning jurisdictions. For instance, Bronin (2021) assembles zoning data in 180 jurisdictions in Connecticut, taking more than four months with 20 research assistants. Some recent works, such as Mleczko and Desmond (2023) and Bartik et al. (2023), develop natural language processing methods to collect and clean zoning details more efficiently.

7. For example, the Density Restriction Index (DRI) in the Wharton Residential Land Use Regulation Index indicates whether a community has any minimum lot size requirement and, if it does, whether the largest minimum lot size is smaller than 0.5 acres, between 0.5 and 1 acre, 1 to 2 acres, or larger than 2 acres. As such, it is common to only report the "strictest" or the "most typical" regulation levels in intervals.

from observed property characteristics to construct a nationwide dataset on neighborhood-level stringency of residential zoning. In particular, I estimate minimum lot sizes applied to single-family homes by detecting structural breaks in constructed lot sizes by zoning district and Census Block Group levels using CoreLogic Tax Assessor data. This geographically-detailed zoning dataset covers 52 million single-family homes in 16,217 municipalities in 825 CBSAs across the United States.[8] Finally, I validate these estimates to MAPC Zoning Atlas with publicly available information on minimum lot size regulations in 105 municipalities in California.

Throughout this paper, I focus on minimum lot size regulation as it is the most important dimension of zoning laws in single-family home construction for the following reasons. First, minimum lot size regulation is the most common type of zoning law. For instance, 91% of jurisdictions in the Wharton Land Use Regulatory Index survey and 96% in the Terner Center California Land Use Survey reply that they impose minimum lot size restrictions. Second, minimum lot size is one of the most distortionary zoning laws in single-family construction. To illustrate, I compute the bunching rate of single-family home construction around minimum lot sizes and maximum floor-area ratio, another type of zoning law often studied in the literature. In CoreLogic data merged with MAPC Zoning Atlas, 18 percent of single-family construction after 1940 bunch at minimum lot sizes while only 5 percent of the construction bunch at maximum floor-area ratios. As such, although residential zoning regulates multiple dimensions of building characteristics, minimum lot size regulation is one of its most critical components, especially in single-family home construction.

2.1 Constructing the Minimum Lot Size Dataset

In this subsection, I describe the procedure of the minimum lot size data construction in greater detail. I begin by describing the key data input, property tax records from CoreLogic Tax Assessor data. Then I explain how I use the geolocation, construction year, and lot size data of single-family home construction to apply a structural break detection algorithm to estimate minimum lot size regulations.

8. To compare, there are 86 million single-family homes in CBSAs according to 2020 American Community Survey. Note that the minimum lot size data exclude single-family homes without functioning local government entities. As such, the minimum lot size data covers *at least* 60% of single-family homes subject to local zoning laws.

Data Used in Minimum Lot Size Estimation

The primary dataset I use for zoning data construction is CoreLogic Tax Assessor data. CoreLogic tax data includes property tax records from 2009 to 2019 for residential and non-residential properties.[9] Each tax record includes property characteristics, such as building type, lot area, number of rooms, and construction year, as well as tax-assessed values and tax amounts in each tax year. I collect and geocode single-family home tax records to focus on min lot size restrictions applied to single-family homes.

I merge single-family home tax records with Census County Subdivision maps and Census Place maps. I use the geographic information to group the parcels into municipalities with functioning local governments that could set zoning laws, according to 2010 Census Guide to State and Local Census Geography. See Appendix Table A1 for more detail about the definition of municipalities. As a result, I obtain the property tax records of 67 million single-family homes that were constructed after 1940 within 830 Core-Based Statistical Areas.[10]

Construct Neighborhood $\times$ Construction Period Grids

Local governments divide their land into zoning districts and set zoning laws in each district. For example, they tend to apply tighter zoning restrictions in residential-only neighborhoods in the periphery compared to mixed-use neighborhoods in city centers.[11] As such, zoning laws vary across zoning districts, even within a municipality. Hence, the stringency of zoning laws should be measured at the zoning district level.

To proxy zoning districts to capture the within-municipality variation in zoning, I partly rely on zoning district information in CoreLogic tax data. This zoning district data is not always clean and has limited coverage.[12] Therefore, in the municipalities where the zoning district information is missing, I use Census block groups to define a neighborhood. I choose the

9. CoreLogic collects this data from local government entities responsible for levying property taxes, which are typically county governments.

10. I exclude observations that do not belong to any Core-Based Statistical Areas to exclude rural areas throughout the paper.

11. For example, many cities have R1 district and R2 districts, where both are residential districts, with R2 allowing denser housing development. It is common to have multiple residential districts within a municipality.

12. 46 % of single-family home parcels in the sample have CoreLogic zoning data field filled. When the data field is filled, it has an unclean zoning district name, for example, R1, R-1, or R2.

Census block group for the spatial unit of interest because its boundary partly reflects local legal boundaries, and its map is available nationwide in a shapefile format.

Since existing structures would be grandfathered in when zoning laws are newly adopted or changed, effective min lot size restrictions may vary by construction years. Unfortunately, I do not observe the year of adoption of zoning laws by the municipality, and this information is extremely hard to collect (Cui, 2023). For my baseline min lot size estimates, therefore, I use all single-family home construction built after 1940, given that most municipalities first adopted local zoning in the mid-to-late 1900s.

Minimum Lot Size Detection

New construction must comply with minimum lot size regulations unless they meet exception criteria or apply for zoning variances or rezoning. Such non-compliance is expensive and uncommon. Therefore, new construction that aims for small lot sizes is forced to construct at the minimum lot size or not construct at all. This creates a discontinuity or kink in the distribution of residential lots at the minimum lot size. Figure 1 illustrates such a structural break in observed lot sizes in CoreLogic tax data merged with MAPC Zoning Atlas. It depicts that single-family homes constructed after 1940 substantially bunch at minimum lot sizes, and the structural break in the distribution function is evident at the lot size requirement.

I detect kinks in the distribution of lot sizes motivated by the structural break detection literature (Andrews, 1993; Zeileis, 2005) and define these breakpoints as the minimum lot sizes. Specifically, I solve the following minimization problem:

$$\min_{\bar{x}_h} \left[\min_{F_{h,1}, F_{h,2} \in \mathcal{F}} \frac{1}{N_h} \sum_{x_{i,h}} (\hat{F}_h(x_{i,h}) - F_{h,1}(x_{i,h}) \mathbb{1}(x_{i,h} < \bar{x}_h) - F_{h,2}(x_{i,h}) \mathbb{1}(x_{i,h} \geq \bar{x}_h))^2 \right]$$

where h is a neighborhood (either zoning district or Census Block Group), i is a single-family home, $x_{i,h}$'s are the observed lot sizes, N_h is the # of single-family homes in the neighborhood, and $\hat{F}_h(\cdot)$ is the empirical distribution of single-family home lot sizes in the neighborhood. $F_{h,1}$ and $F_{h,2}$ are the (estimated) cumulative distribution functions for single-family homes smaller than the minimum lot size and for single-family homes bigger than or equal to the minimum lot size, respectively. $F_{h,1}$ and $F_{h,2}$ are chosen from a family of smooth functions

($\mathcal{F}$), which I set to be the family of seventh-order polynomials. I identify the value $\bar{x}_h$ that minimizes the sum of squared residuals with a single breakpoint at each neighborhood with more than 50 observed single-family home lot sizes. I use this estimate as my minimum lot size estimate. To illustrate, Figure 2 depicts the actual minimum lot sizes and detected structural breaks (estimated minimum lot sizes) in four example zoning districts. I apply the structural break detection algorithm nationwide.

2.2 Validation of Minimum Lot Size Estimates

In this subsection, I validate the minimum lot size estimates using MAPC Zoning Atlas. MAPC Zoning Atlas compiles zoning information in 101 cities and towns of Metropolitan Boston and includes minimum lot size data in each zoning district. It is one of the few public data sets that include actual minimum lot sizes for multiple municipalities.[13] I merge my parcel-level minimum lot size estimates with the Zoning Atlas map and compare the estimates (estimated MLS) with the minimum lot sizes provided in MAPC (actual MLS).[14] The median error rate ($= \frac{\text{estimated MLS - actual MLS}}{\text{actual MLS}}$) is 11.7%, and the correlation between estimated MLS and actual MLS in the validation data is 0.81. Panel A of Table 1 reports further statistics of the error rates in the validation dataset.

To provide benchmarks to interpret the error rates quantitatively, I report error rates when minimum lot sizes were proxied by simpler statistics of observed lot sizes. For example, if there were no zoning variances and no changes in minimum lot sizes since 1940, then simply taking the minimum of constructed lot sizes would capture the regulation precisely. Table 2 reports the benchmark error rates when minimum lot sizes were proxied by the 1st percentile or 10th percentile of lot sizes of single-family homes constructed since 1940. The median error rate is 73.8% and 42.1%, respectively, far larger than the 11.7% median error in my estimates. Even

13. Most other existing zoning datasets provide limited and censored information on minimum lot sizes. For example, similar to Wharton Residential Land Use Regulation Index reporting the largest minimum lot sizes in each municipality in categories, Connecticut Zoning Atlas includes information on whether minimum lot size regulation exists, and if it does, whether it is smaller than 0.47 acre, between 0.47 and 0.91 acre, between 0.92 and 1.84 acres, or larger than 1.84 acres.

14. I use all past 11 years of single-family home tax records as of 2019 and merge estimated MLS based on parcel location. All statistics, for example, percentiles, are weighted by the number of property tax records.

the 1st decline error appears to be 50.1% and 14.3%, respectively, indicating that these simple proxies are not valid in many locations.

Additionally, I compare Wharton Residential Land Use Regulatory Index data with MAPC Zoning Atlas data to illustrate discrepancies in existing zoning data sets. Wharton Land Use Regulatory Index data is the most popular measure of zoning stringency, and the data is collected by surveying local municipalities. The Density Restriction Index (DRI) in the survey reflects whether the community has any minimum lot size regulations, and if it does, whether their *largest* minimum lot size is no larger than 0.5 acres, between 0.5 and 1 acres, 1 and 2 acres, or more than 2 acres. I link the 2018 Wharton survey with MAPC Zoning Atlas and compare minimum lot size information in the two datasets.

Table 3 reports the comparison between Wharton DRI and the largest minimum lot sizes in MAPC Zoning Atlas in 30 municipalities that appear in both datasets. I find that only 17 of the 30 municipalities have matching minimum lot size information in the Wharton survey and MAPC Zoning Atlas, while 6 of them underreported their largest minimum lot sizes, and the rest 7 municipalities overreported their largest minimum lot sizes. This discrepancy in the zoning data is potentially due to survey measurement errors (Mleczko and Desmond, 2023) or mismatch between the study time periods.

Wharton land use survey has a strong advantage in characterizing the restrictiveness of overall local regulatory environments, including administrative delays and implicit hurdles in new residential development. In comparison, my zoning data offers more geographically granular and precise measures of density restriction but only captures minimum lot sizes applied to single-family home construction.

2.3 Descriptive Statistics: The State of Local Zoning Laws in the United States

The constructed minimum lot size data allows me to provide the first nationwide descriptions of the regulations across the United States. I document the following three facts that characterize the current stringency and restrictiveness of zoning laws.

First, I find that minimum lot size restrictions are stringent overall: the average single-family home is subject to 16,000 square feet min lot size, and the median is subject to 9,200 square feet.[15] About half of residential land in the U.S. is subject to min lot size restrictions of 19,000 square feet or larger. See Table A2 and A3 for detailed summary statistics by state.

Second, most municipalities impose extremely stringent minimum lot size regulations on at least some part of their land. For example, 74.6 % of all municipalities have at least one zoning district with 1 acre or larger minimum lot sizes. 44.0 % of all municipalities have at least one zoning district with 2 acres or larger minimum lot sizes.

Third, min lot size regulations distort housing characteristics: about 15.4% of single-family homes built after 1940 are bunching within 5% of regulated lot sizes. This bunching rate proxies the restrictiveness of min lot size regulations. Figure 3 depicts the stringency, measured by the median min lot size, and the restrictiveness, measured by the bunching rate, by state. New Jersey has the most restrictive zoning with a bunching rate of 23%, followed by Florida and California with 22% and 19% bunching, respectively.

3 Estimating the Effects of Zoning in Housing Markets

Minimum lot size regulations are a type of density restriction in residential zoning. They restrict housing supply, shape neighborhoods, and, in turn, affect housing prices. In this section, I estimate the effects of minimum lot size regulations on home sales prices and rents and investigate the mechanisms of the price effects using the novel zoning dataset. In addition, I estimate their effects on residential sorting on the dimensions of homeowner race and income. I start by explaining underlying assumptions in the boundary discontinuity design and datasets used in the border analysis. I then discuss the results.

15. The mean and median are calculated by weighting min lot size estimates by the number of parcels in each district.

3.1 Empirical Strategy

A major challenge to understanding the causal effects of zoning on housing market outcomes is that the stringency of zoning regulations may be correlated with unobserved location amenities. For example, suburban places with higher-quality residential environments may set stricter zoning to prevent low-income households from entering. In contrast, populated cities with well-developed infrastructures may allow denser development with relaxed zoning. As such, minimum lot sizes may be positively (in the former case) or negatively (in the latter case) correlated with location amenities. As location amenities affect housing market outcomes, such as prices and demographics, the OLS coefficients of minimum lot size on these outcomes may be biased.

To address this concern, I adopt a boundary discontinuity design developed by Black (1999). In my setting, I compare housing prices at municipality borders, such as city, town, and township borders, where the stringency of zoning varies. Additionally, I control for pre-zoning neighborhood attributes, which I construct by geocoding the 1940 full-count Census and aggregating at the border region levels. Finally, I include municipality (city, town, township, and county with functioning local governments)-by-school district fixed effects to address the concern about potential differences in local services, such as trash collection and snow removal, and school district quality that may be correlated with zoning stringency. This is possible due to the granularity of minimum lot size estimates with variations within municipalities in addition to across-municipality variations. The key identification assumption is that unobserved amenities are as good as random conditional on observable characteristics within some pre-specified small region around the border.

3.2 Data Used in the Border Analysis

I compile the following datasets to construct transaction-level home sales prices, rents, and homeowner demographics, as well as neighborhood characteristics. I use them for the municipal border analysis in addition to the minimum lot size dataset I describe in Section 2.

Corelogic Deed-Home Mortgage Disclosure Act Data

CoreLogic deed data, available since the 1980s, includes deed and mortgage transactions. I collect sale transactions of single-family homes from the deed data from 2000 to 2018 and merge property characteristics from CoreLogic Tax Assessor data, described in 2, including lot size, building square footage, (effective) construction year, number of bedrooms, and number of bathrooms. I then link them with Home Mortgage Disclosure Act (HMDA) loan application register data, following the approach of Bayer et al. (2007), which involves matching Census tract, mortgage year, and lender name.[16] The match rate is 65.7% for CoreLogic deed records of residential properties with mortgage information.[17] The resulting dataset provides transaction-level prices and demographics (race, ethnicity, and income) of mortgage applicants of single-family home sales.

I geocode CoreLogic-HMDA data and merge it with Census County Subdivision maps and Census Place maps to identify single-family homes in municipality borders.[18] I use 0.1 km (0.06 miles) to 1.0 km (0.6 miles) to define border regions and choose 0.5km as a baseline. Standard choices of the border distance in spatial boundary discontinuity design range from 0.24km (0.15 miles) to 0.56km (0.35 miles) (Bayer et al., 2007; Turner et al., 2014). I restrict my analysis to within-county municipality borders with at least 25 single-family home parcels on both sides of borders.

For price analyses, I use CoreLogic deed data before merging HMDA to include cash sales and unmatched transactions. The sample consists of 10.6 million transactions in 10,315 border regions. See Panel A of Table 4 for sample statistics. For demographics analyses, I use CoreLogic-HMDA data where I observe mortgage lenders' race and income. The sample consists of 4.6 million transactions in 9,594 border regions. See Panel B of Table 4 for sample statistics.

16. Other papers that have implemented this merging procedure include Avenancio-Leon and Howard (2019) and Diamond and McQuade (2019).

17. This match rate is comparable to other papers that use the same approach. For example, Bayer et al. (2007) have a match rate of 60% in the Bay Area. Bayer et al. (2007) also illustrate the representativeness of the matched sample.

18. See Appendix Table A1 for more detail about the definition of municipality boundaries.

Corelogic Multiple Listings Service Data

CoreLogic listings data includes sale and rental listings collected from 154 multiple listings services (MLS).[19] Each MLS maintains its own database of property listings and often covers specific local markets. CoreLogic aggregates these databases into a single dataset. The dataset includes the list date and price and, if the listing is closed, the closing date and price. Since I already have sale information from CoreLogic deed data, I focus on rental listings in the MLS data and use closing prices of single-family home rental listings in the analysis. Although rental listings count for only small part of the data, I obtain 300 thousand geocoded single-family home rental listings with closing prices in 3,023 border regions for analysis. See Panel C of Table 4 for sample statistics.

1940 Census Full-Count Data

The set of pre-zoning neighborhood characteristics come from 1940 Ancestry.com and IPUMS complete-count Census restricted data (Ruggles et al., 2017). I geocode the street addresses of individual-level 1940 Census data and merge them with the 2019 municipality boundaries to obtain historical neighborhood characteristics of border regions. Note that this neighborhood data is sparse because most of the municipalities were created during postwar suburbanization. Thus, I observe whether the municipality existed as of 1940 and, if it did, the neighborhood characteristics of each border region. The characteristics include the total population, household size/structure, mobility, demographics such as race and income, homeownership, average home value, and average rent. I use these pre-zoning neighborhood characteristics as control variables.

2019 School District Boundaries

I merge the geocoded CoreLogic deed, deed-HMDA, and MLS datasets with 2019 school district boundaries. By doing so, I observe the unified school district, elementary school district, and secondary school district that each transacted home was located in. I construct school district fixed effects for each pair of a unified school district, elementary school district, and secondary school district and include them as controls.

19. There are 597 multiple listing services as of 2020, according to the Real Estate Standards Organization.

3.3 Effects on Home Prices and Rents

My baseline transaction-level regressions take the form

$$\log p_{it} = \beta_{MLS} \log MLS_i + X_i\beta_X + \lambda_{b(i)} + \lambda_t + \lambda_{m\times s} + \varepsilon_{ibmt}, \tag{1}$$

where i denotes a single-family home transaction and t is the housing market, defined as county-by-transaction year. $\log p$ is log of sales price or rent, and $\log MLS_i$ is log of minimum lot size (MLS) applied to the single-family home. In this specification, I assume the treatment effects of minimum lot size restrictions are linear in log of MLS. In Appendix Figure A1, I present alternative binscatter regressions, which show consistent results. X_i is a set of historic neighborhood-level controls. See Appendix Table A4 for the full list of control variables. λ_b is border region fixed effects, defined by an unordered pair of municipalities, λ_t is county × transaction year fixed effects, and $\lambda_{m\times s}$ is local government (city, town, township, or county) × school district fixed effects. The inclusion of border region fixed effects implements the spatial discontinuity design.

β_{MLS} in Equation 1 captures the overall long-term price effects of MLS. Note that price outcomes here are for a single-family home instead of per square foot. This way, β_{MLS} captures *direct effect*; each housing unit is more expensive as it is required to be bigger. In addition, β_{MLS} includes neighborhood effects; restrictive zoning shapes neighborhood amenities that may be (either positively or negatively) valued by households. Neighborhood effects may include effects on neighborhood ambiance and demographics. However, when including school district and municipality fixed effects, β_{MLS} does not capture some of the municipality-wide effects. For example, zoning may affect local tax bases or peer demographics of school districts, but these channels are not reflected in the coefficient estimates. Finally, β_{MLS} does not include the supply effect; restrictive zoning limits housing supply, increasing housing prices. This is because the border discontinuity design is based on the assumption that neighborhoods on both sides of borders are comparable, controlling for observables. Under the assumption, local housing supply conditions should affect both sides of the borders. Therefore, the supply effect is captured by fixed effects instead of being reflected on β_{MLS} in my empirical setting.

Table 5 reports the coefficient estimates with the baseline border distance (0.5km). The full specification result in Column (3) indicates that restrictive zoning increases home sales prices: $\hat{\beta}_{MLS} = 0.1347$, with p-value < 0.01. That is, doubling MLS would increase the sales price of a home by about 10 percent. The full specification result in Column (3) indicates that restrictive zoning increases home sales prices: $\hat{\beta}_{MLS} = 0.0881$, with p-value < 0.01. That is, doubling MLS would increase the rent of a home by about 6 percent.

3.4 Price Effect Mechanisms

As discussed in the previous subsection, the baseline coefficient estimates reflect the overall price effects, including the direct effect and neighborhood effect. I further investigate the baseline price effect estimates by the mechanisms. To do so, I run border discontinuity regressions in Equation 1 with building-level characteristics as additional control variables. In discussing the results, I consider minimum lot size as a proxy for the overall stringency of zoning, including restrictions on other building dimensions, such as floor-area-ratio and height. Alternatively, one may say that minimum lot size regulations only shift the lot size while they do not affect other building characteristics. An alternative interpretation based on this assumption can be found in Appendix Table A6 where I consider building-level characteristics other than lot size as additional control variables instead of endogenous variables.[20]

Table 6 reports the regression results with price effect mechanisms. β_{MLS}, the price effect coefficient for the sales price, controlling for observable building-level characteristics in Column (2) is 0.0220, decreased from 0.1382 in the baseline specification. Note that Column (2) controls for building square feet, age, number of bedrooms, and number of bathrooms in addition to lot size, which is directly affected by minimum lot size regulations. About 84 percent of the overall price effect of zoning is caused by shifting building characteristics, which is the direct effect. Similarly, 66 percent of the overall rent effect may be attributed to the direct effect.

Appendix Figure A2 illustrates how each of the building characteristics is affected by mini-

20. Other building characteristics than lot size may be affected by minimum lot size regulations due to preferences of homeowners and builders. For example, homeowners and builders may prefer larger square footage in a larger lot. As such, if building characteristics interact with lot size in their utility function, a regulation on lot size affects overall building dimensions. In this case, this alternative interpretation is not valid, and my baseline interpretation is preferable.

mum lot sizes. Stricter (larger) minimum lot sizes increase lot size, square footage, # of bedrooms, and # of bathrooms. This may be due to correlated zoning laws of other types, such as maximum floor-area-ratio or setback requirements, but also due to market preference. Homeowners and builders may prefer having larger square footage and more rooms when the lot is larger. Stricter minimum lot sizes increase building ages as well, indicating that new construction is more limited when zoning is stricter. Overall, these shifts in building characteristics by zoning laws account for the majority of their price and rent effects.

The remaining effect of zoning laws on sales prices, which I consider as the neighborhood effect, is 0.0220 in Column (2), and the remaining effect on rents is 0.0302 in Column (4). These neighborhood effects are attributed to neighborhood environments shaped by stringent zoning that are capitalized in, including the following examples. First, low-density neighborhoods may generate a quieter and more livable ambiance of the neighborhood. Second, higher property values of neighboring homes increase local property tax collection, thus increasing housing prices in the neighborhood. Third, the neighborhood effect on sales prices reflects the insurance value of zoning; stringent zoning may insure against negative changes in the neighborhood, protecting property values in the future. Finally, It is worth mentioning that I cannot rule out potential biases from unobserved building quality correlated with MLS in β_{MLS}. For example, if homes in large MLS neighborhoods tend to have luxury interiors, it would be captured in the neighborhood effect estimates.

The magnitude of estimated neighborhood effects of stringent zoning is fairly small. For example, if neighborhood zoning is twice as stringent with building characteristics fixed, the home price would be 1.5 percent higher, and rent would be 2 percent higher. Yet, the positively estimated neighborhood price effect of stringent zoning implies that existing homeowners have an economic incentive to make zoning more restrictive. When neighborhood zoning becomes more stringent, existing properties are grandfathered in, and their building characteristics do not have to change. Even then, existing homeowners enjoy the neighborhood price premium of stringent zoning through the neighborhood effect.

Figure 4 illustrates how the overall and neighborhood effect estimates of zoning on sales price and rent vary by distance to municipality borders. The top panel indicates that both direct effects and neighborhood effects on sales prices increase as the border region is defined further away from municipality borderlines. Increasing direct effects indicate that zoning regulations are more binding closer to the municipality centers, affecting building characteristics more substantially and thus increasing housing prices. Furthermore, neighborhood effects of zoning increase as the border distance increases. This may be because neighborhood amenities from low-density are maximized when the property is surrounded by larger similar-density environments instead of being right off the border. The subplot on the rent effect in the bottom panel shows a similar trend, especially for the direct effect, although the estimates are much noisier due to the smaller sample size.

3.5 Effects on Residential Sorting

In this section, I examine whether zoning laws induce demographic sorting. To do so, I repeat the border analysis with Equation 1 with homeowner race and income from HMDA data as outcome variables. Table 7 reports the β_{MLS} estimates for homeowner race and income. In the baseline specifications in Columns (2) and (5), I find that more restrictive zoning (larger minimum lot sizes) increases non-Hispanic white homeowners and higher-income homeowners in the neighborhood. For example, when the minimum lot size is doubled, the probability of non-Hispanic white homeowners increases by 2.3 percent points, and homeowner income increases by 9.5 percent. These estimates are robust to the inclusion of municipality-by-school district fixed effects in Columns (3) and (6). That is, zoning regulations explain not only across-municipality segregation but also within-municipality segregation.

Demographic sorting with respect to zoning restrictiveness is important in understanding exclusionary zoning. Exclusionary zoning is a discriminatory practice in which communities impose restrictive zoning laws in order to exclude racial minorities and low-income households. My analysis of demographic sorting due to zoning restrictiveness finds that stringent zoning disproportionately attracts white and high-income homeowners. That is, zoning laws may have been used for implicit exclusion of racial minorities and low-income households.

4 Conclusion

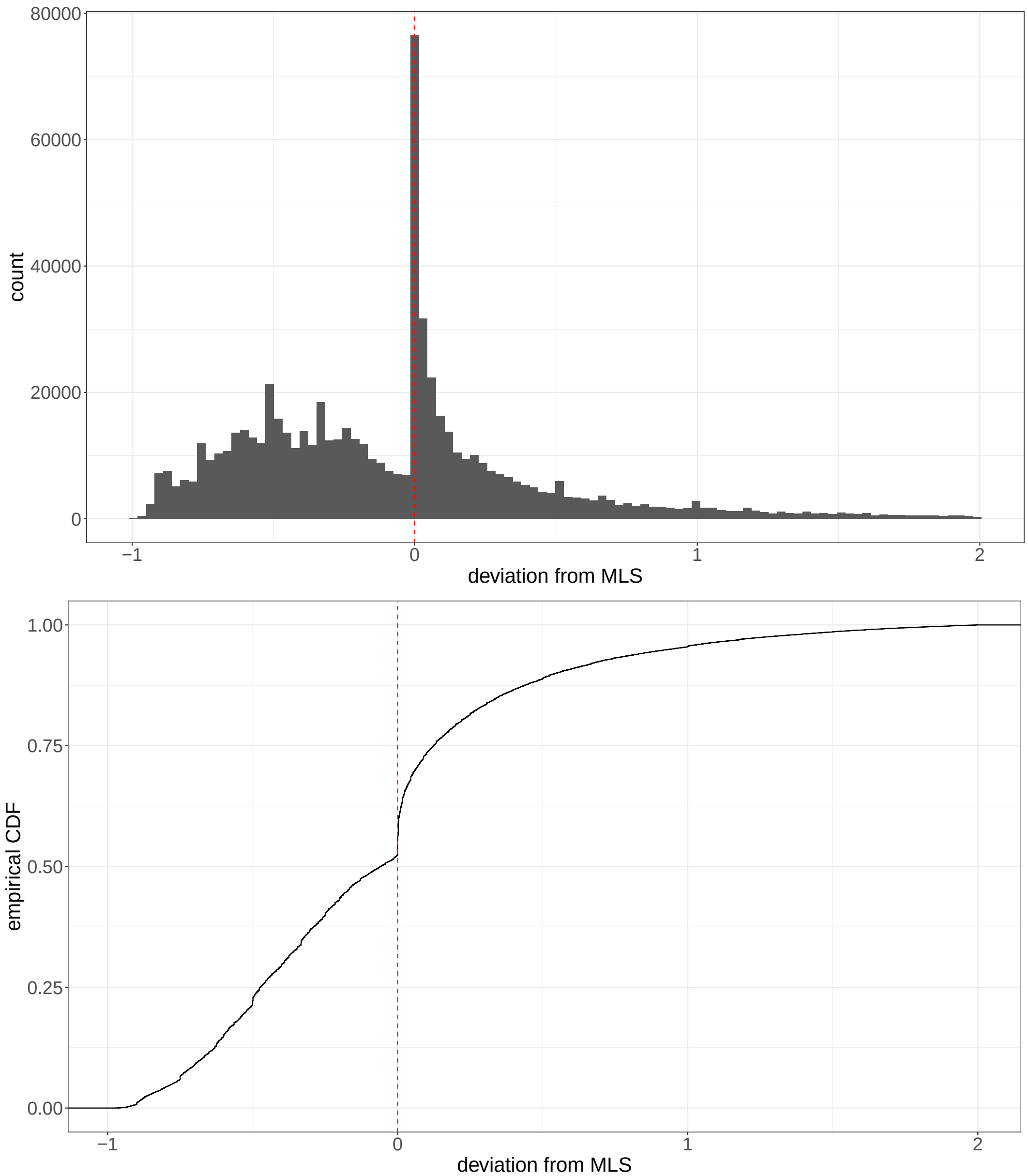
Residential zoning has been accused of limiting housing supply and contributing to the housing affordability crisis. The empirical literature on zoning laws has been growing rapidly, but limited geographic coverage of existing zoning data is still a major challenge in the literature. In this paper, I build a nationwide dataset on minimum lot size regulations and study its impact on housing markets. I find that minimum lot size restrictions play significant roles in increasing housing prices, primarily by shifting the building characteristics. I also find that the regulations intensify segregation by disproportionately attract high-income white homeowners.

My analysis begins with constructing a nationwide dataset of neighborhood-level minimum lot size estimates. I propose a new approach to estimate dimensional requirements by detecting structural breaks in observed property characteristics in property tax records. I apply a structural break detection algorithm to estimate minimum lot sizes across the nation and build a dataset on zoning district-level minimum lot size estimates with substantially broader coverage than existing zoning datasets. This dataset allows me to conduct the first nationwide examination of minimum lot size regulations.

I leverage a spatial discontinuity design in the stringency of minimum lot size restrictions at municipal borders to estimate their effects on housing markets. Adopting the spatial discontinuity design helps address the endogeneity of zoning. I additionally control for location characteristics in 1940 that may be related to the stringency of zoning. I find that a larger minimum lot size increases both sales prices and rents. For example, doubling minimum lot sizes would increase home sales prices by 10 percent and rents by 6 percent. I also find that zoning laws intensify residential segregation; neighborhoods with restrictive zoning disproportionately attract high-income white homeowners due to the lack of affordable housing options and heterogeneous preference for neighborhood amenities shaped by zoning.

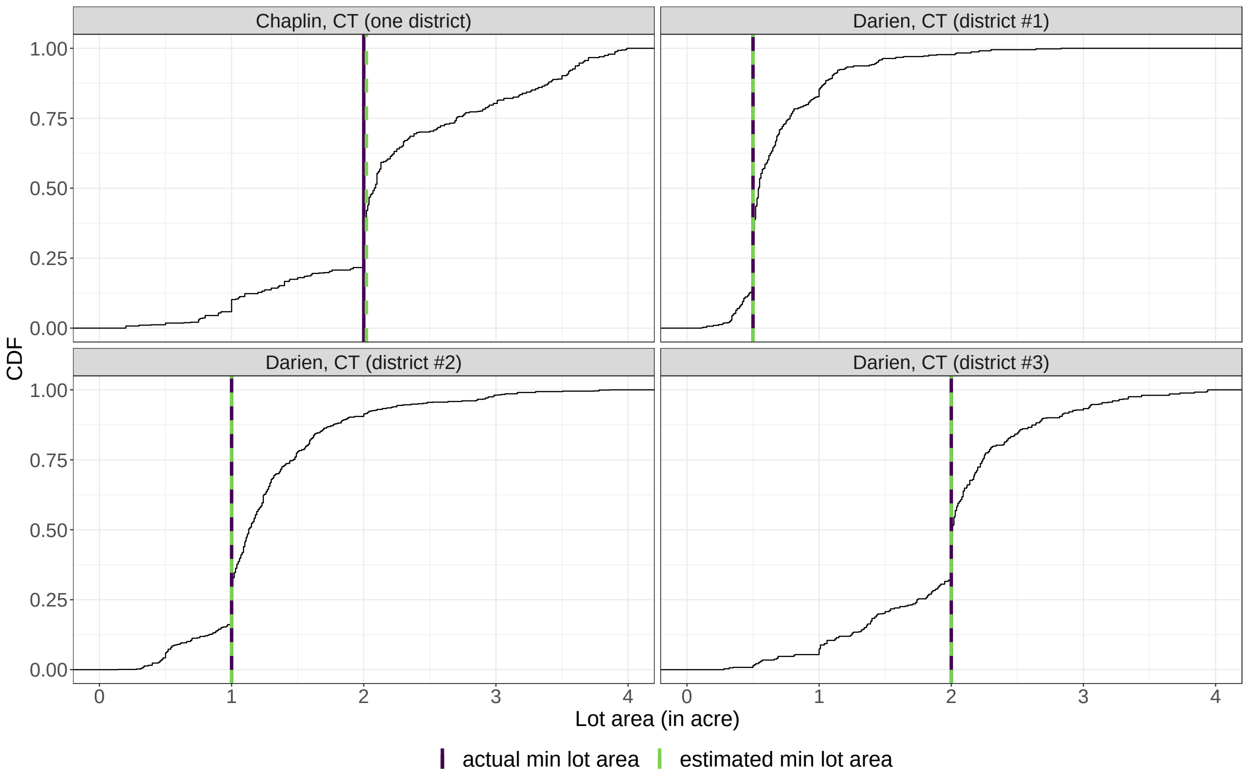
Zoning laws have come under scrutiny globally as many places face housing affordability problems. As zoning laws define the minimum requirements of housing construction, restrictive zoning removes affordable housing options and increases housing prices. This paper contributes to measuring the stringency of zoning in neighborhoods across the U.S. and quantifying its impact in the housing markets. This research provides guidance for zoning reforms, which are actively being discussed in the U.S. and other countries, and opens up the possibility of future empirical research on residential zoning.

Figure 1 — Kink at Minimum Lot Sizes (MLS) in MAPC Zoning Atlas



Note. The figure depicts the distribution of deviation from minimum lot size, defined as $\frac{\text{lot size}-MLS}{MLS}$, of single-family homes built after 1940 (top: histogram with 100 bins, bottom: empirical distribution function). The underlying data is from CoreLogic tax data merged with MAPC Zoning Atlas, including 656,695 single-family home construction in 769 zoning districts with minimum lot sizes ranging from 1,500 square feet to 4 acres (median = 20,000 sqft).

Figure 2 — Minimum Lot Size Detection from the Distribution of Constructed Lot Sizes



Note. The figure depicts the distribution of lot sizes of single-family homes built after 1940 in four example zoning districts. For each district, actual minimum lot sizes are denoted by purple solid lines, and estimated minimum lot sizes are denoted by dashed green lines. In Chaplin, the actual minimum lot size is 2 acres municipality-wide. In Darien, which is comprised of three districts, districts #1–#3 respectively have 1/2-acre, 1-acre, and 2-acre minimum lot sizes. The estimated minimum lot sizes are constructed by detecting structural breaks in the lot size distribution.

Table 1 — Validation of Estimated Minimum Lot Sizes Using MAPC Zoning Atlas Data

Decile	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
	0.45 %	0.58 %	1.00 %	6.67 %	11.7 %	28.0 %	49.8 %	50.0 %	67.3 %

Note. This table reports the error in estimated minimum lot sizes (estimated MLS) when it is compared to actual minimum lot size regulations (actual MLS) found in MAPC Zoning Atlas. It reports the deciles of error rate (in %) in estimated MLS, defined as = $\frac{\text{estimated MLS - actual MLS}}{\text{actual MLS}}$. The data is restricted to single-family homes subject to minimum lot size regulations in MAPC, excluding outliers with minimum lot sizes larger than or equal to 5 acres.

Table 2 — Minimum Lot Size Simple Proxies Compared to MAPC Zoning Atlas Data

A. Error rate in MLS, when proxied by 1st percentile constructed lot sizes

Decile	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
	50.1 %	55.9 %	61.6 %	68.0 %	73.8 %	75.5 %	77.6 %	82.1 %	86.4 %

B. Error rate for in MLS, when proxied by 10th percentile constructed lot sizes

Decile	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
	14.3 %	32.0 %	33.3 %	42.1 %	45.6 %	53.5 %	58.9 %	66.2 %	74.0 %

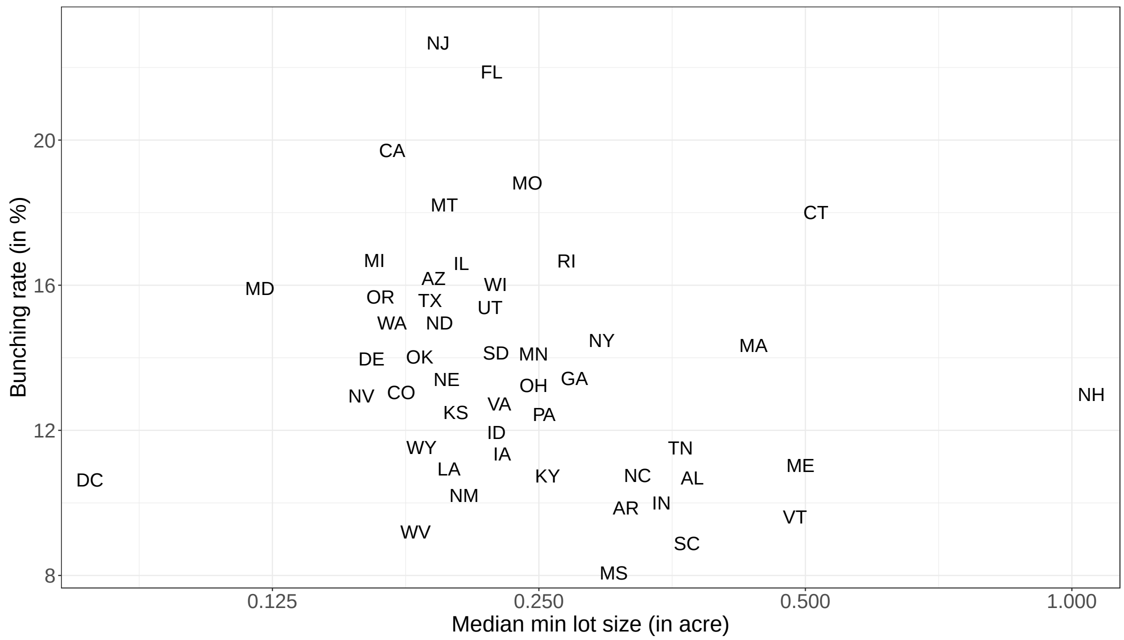
Note. This table reports the deciles of error rate (in %) in other possible minimum lot size proxies, defined as = $\frac{\text{estimated MLS - actual MLS}}{\text{actual MLS}}$, when it is compared to actual minimum lot size regulations (actual MLS) found in MAPC Zoning Atlas, excluding outliers with minimum lot sizes larger than or equal to 5 acres. The data is restricted to single-family homes subject to minimum lot size regulations (actual MLS > 0). Panel A uses the 1st percentile of single-family lot sizes in each proxy zoning district, constructed since 1940, as proxies for minimum lot sizes. Panel B uses the 10th percentile of single-family lot sizes in each proxy zoning district, constructed since 1940, as proxies for minimum lot sizes.

Table 3 — Wharton Index Compared to MAPC Zoning Atlas Data

		% municipalities by reporting error		
Wharton DRI	# municipalities	Correctly	Under	Over
No larger than 0.5 acres	6	50%	50%	.
0.5 – 1 acres	11	54.5%	9.1%	36.4%
1 – 2 acres	7	85.7%	14.3%	0%
More than 2 acres	6	33.3%	.	66.7%
All parcels	30	56.7%	20%	23.3%

Note. The table reports the reporting errors in Wharton Density Restriction Index (DRI). To calculate the reporting errors, Wharton DRI is compared to the largest minimum lot size in MAPC Zoning Atlas, restricted to residential districts while excluding planned development zones, in each municipality.

Figure 3 — Stringency and Restrictiveness of Zoning by State



Note. The figure depicts the stringency of zoning (x-axis), measured by the median min lot sizes, and the restrictiveness of zoning (y-axis), measured by bunching rate at the min lot sizes, by state.

Table 4 — Sample description (border distance = 0.5 km)

A. Deed sample statistics (within 0.5km of borders)					
# obs.			10,598,492		
# parcels			5,788,051		
# borders			10,315		
Variable	Q1	Q2 (Median)	Q3	Mean	SD
Sales price (in $1000s)	112	183	295	242	225
Min lot size (in sqft)	7,002	8,999	13,068	13,580	18,225
Lot size (in sqft)	6,400	8,834	13,995	27,955	1,954,375
Building footage (in sqft)	1,300	1,695	2,288	1,905	1,690
(Effective) Year built	1966	1990	2002	1982	26
# Bedrooms	3	3	4	3.25	0.81
# Bathrooms	2	2	3	2.31	0.93

B. Deed+HMDA sample statistics (within 0.5km of borders)					
# obs.			4,662,670		
# parcels			3,457,169		
# borders			9,594		
Variable	Q1	Q2 (Median)	Q3	Mean	SD
Sales price (in $1000s)	138	206	320	268	219
Min lot size (in sqft)	7,000	8,712	12,600	13,091	17,505
1 (Non-Hispanic white)				0.689	
1 (Asian Pacific)				0.059	
1 (Black)				0.061	
1 (Hispanic)				0.096	
Income (in $1000s)	49	74	113	100	794

C. MLS rental listing sample statistics (within 0.5km of borders)					
# obs.			300,305		
# parcels			167,597		
# borders			3,023		
Variable	Q1	Q2 (Median)	Q3	Mean	SD
Monthly Rent (in $)	1,175	1,479	2,045	5,008	22,587
Min lot size (in sqft)	6,787	8,250	10,720	11,105	12,644
Lot size (in sqft)	5,976	7,500	10,290	25,661	2,549,502
Building footage (in sqft)	1,420	1,820	2,371	1,984	816
Year built	1979	1997	2005	1990	20
# Bedrooms	3	3	4	3.32	0.78
# Bathrooms	2	2	3	2.37	0.84

Note. This table reports sample statistics. Panel A describes the sales price data from CoreLogic deed data. Panel B describes the homeowner demographics data from CoreLogic deed data merged with Home Mortgage Disclosure Act data. Panel C describes the rental listing data from CoreLogic Multiple Listings Services. All samples are restricted to single-family homes within 0.5km of municipal borders with minimum lot size estimates.

Table 5 — Baseline Price Effect Estimates in Municipal Border Analysis

Dependent Variable:	log sales price				log rent			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log MLS	0.2092***	0.2226***	0.1510***	0.1382***	0.1413***	0.1453***	0.1028***	0.0884***
	(0.0095)	(0.0095)	(0.0081)	(0.0092)	(0.0146)	(0.0135)	(0.0122)	(0.0096)
Controls								
county × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1940 IPUMS		Yes	Yes	Yes		Yes	Yes	Yes
border FE			Yes	Yes			Yes	Yes
loc gov't × SD FE				Yes				Yes
Observations	10,598,492	10,598,492	10,598,492	10,598,492	300,305	300,305	300,305	300,305
R^2	0.38858	0.39130	0.48888	0.52182	0.70169	0.70330	0.75335	0.76417

Robust standard-errors in parentheses clustered at the municipality level
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note. This table reports baseline coefficient estimates of regression 1. The outcome variable in Columns (1) to (4) is log sales price, and the outcome variable in Columns (5) to (8) is log rent. All columns include county-by-transaction year fixed effects, and Columns (2) to (4) and (6) to (8) additionally include 1940 neighborhood characteristics as controls. Columns (3), (4), (7), and (8) implement the border discontinuity design by including border fixed effects. Finally, Columns (4) and (8) also include municipality (city, town, township, or county for unincorporated places)-by-school district fixed effects.

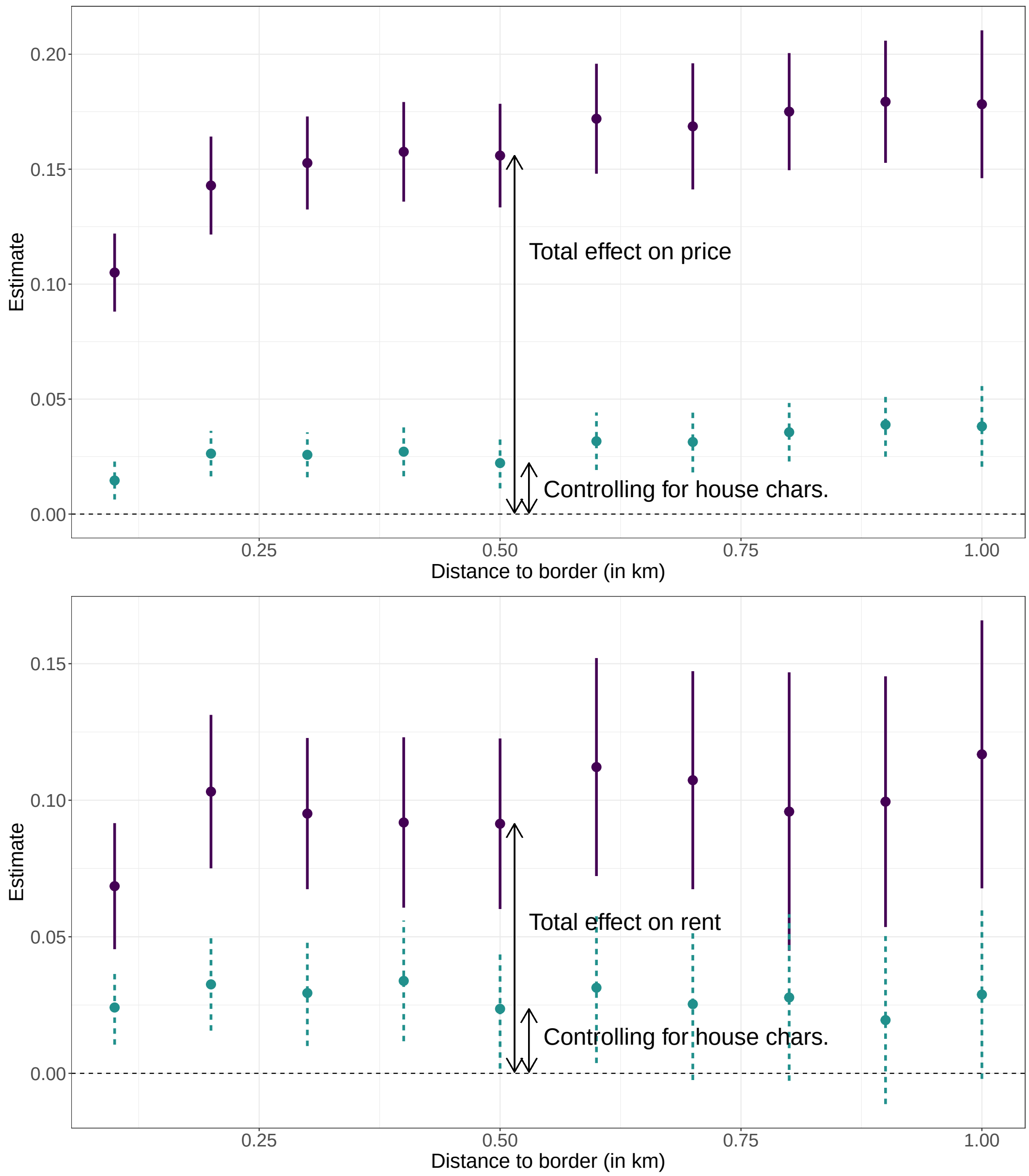
Table 6 — Price Regressions with Building Characteristics Controlled

Dependent Variable:		log sales price		
	log rent			
Model:	(1)	(2)	(3)	(4)
log MLS	0.1382***	0.0220***	0.0884***	0.0302***
	(0.0092)	(0.0042)	(0.0096)	(0.0061)
Building characteristics				
log lot size		0.0803***		0.0271***
		(0.0029)		(0.0040)
log bldg. sqft.		0.6222***		0.4750***
		(0.0081)		(0.0105)
age		-0.0051***		-0.0022***
		(0.0001)		(0.0002)
# bed		-0.0151***		-0.0028
		(0.0016)		(0.0025)
# bath		0.0687***		0.0347***
		(0.0018)		(0.0029)
Controls				
Full Controls and FEs	Yes	Yes	Yes	Yes
Observations	10,598,492	10,598,492	300,305	300,305
R^2	0.52182	0.61884	0.76417	0.81092

Robust standard-errors in parentheses clustered at the municipality level
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note. This table reports coefficient estimates of regression 1 with building-level controls. The outcome variable in Columns (1) and (2) is log sales price, and the outcome variable in Columns (3) and (4) is log rent. All columns include county-by-transaction year fixed effects, 1940 neighborhood characteristics, border fixed effects, and municipality-by-school-district fixed effects. Columns (2) and (3) additionally include transaction-level building characteristics.

Figure 4 — Price effects by border regions



Note. The figure depicts the coefficient estimates (dots) and 95% confidence intervals (lines) of β_{MLS} for sales prices (top panel) and rents (bottom panel) by border regions (x-axis). Each panel shows 10 border regions: [0,0.1km], [0.1,0.2km], [0.2,0.3km], [0.3,0.4km], [0.4,0.5km], [0.5,0.6km], [0.6,0.7km], [0.7,0.8km], [0.8,0.9km], [0.9,1km]. For each outcome variable and border region, the total price/rent effects are depicted in purple dot and solid line, and the remaining effects after controlling for building level characteristics are in green dot and dashed line. All specifications include county-by-transaction year fixed effects, 1940 neighborhood characteristics, and border fixed effects.

Table 7 — Homeowner Demographic sorting

Dependent Variables:	1(race = non-Hispanic white)			log income		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
log MLS	0.0532***	0.0334***	0.0304***	0.1671***	0.1314***	0.1217***
	(0.0027)	(0.0024)	(0.0029)	(0.0057)	(0.0047)	(0.0058)
Controls						
county × year FE	Yes	Yes	Yes	Yes	Yes	Yes
1940 IPUMS		Yes	Yes		Yes	Yes
border FE		Yes	Yes		Yes	Yes
loc gov't × SD FE			Yes			Yes
Fit statistics						
Observations	4,662,670	4,662,670	4,662,670	4,662,670	4,662,670	4,662,670
R^2	0.13266	0.18778	0.19690	0.21095	0.31306	0.33482

Robust standard-errors in parentheses clustered at the municipality level
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note. This table reports baseline coefficient estimates of regression 1 with homeowner race and income from HMDA data as outcome variables. The outcome variable in Columns (1) to (3) is the indicator variable of whether the homeowner is non-Hispanic white, and the outcome variable in Columns (4) to (6) is log homeowner income. All columns include county-by-transaction year fixed effects, and Columns (2), (3), (5), and (6) additionally include 1940 neighborhood characteristics as controls and border fixed effects. Finally, Columns (3) and (6) also include municipality (city, town, township, or county for unincorporated places)-by-school district fixed effects.

References

Albouy, David, and Gabriel Ehrlich. 2018. “Housing productivity and the social cost of land-use restrictions.” *Journal of Urban Economics* 107 101–120.

Anagol, Santosh, Fernando Ferreira, and Jonah Rexer. 2021. “Estimating the Economic Value of Zoning Reform.” *Working Paper.*

Andrews, Donald WK. 1993. “Tests for Parameter Instability and Structural Change with Unknown Change Point.” *Econometrica: Journal of the Econometric Society* 821–856.

Avenancio-Leon, Carlos, and Troup Howard. 2019. “The Assessment Gap: Racial Inequalities in Property Taxation.” *Available at SSRN 3465010.*

Bartik, Alexander, Arpit Gupta, and Daniel Milo. 2023. “The Costs of Housing Regulation: Evidence From Generative Regulatory Measurement.” *Available at SSRN.*

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. “A Unified Framework for Measuring Preferences for Schools and Neighborhoods.” *Journal of political economy* 115 (4): 588–638.

Black, Sandra E. 1999. “Do Better Schools Matter? Parental Valuation of Elementary Education.” *The quarterly journal of economics* 114 (2): 577–599.

Bronin, Sara C. 2021. “Exclusion, Control, and Consequence in 2,622 Zoning Districts.” *Control, and Consequence in* 2.

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

Cui, Tom. 2023. “The Emergence of Exclusionary Zoning Across American Cities.”

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

Freemark, Yonah. 2023. "Zoning Change: Upzonings, Downzonings, and Their Impacts on Residential Construction, Housing Costs, and Neighborhood Demographics." *Journal of Planning Literature* 08854122231166961.

Glaeser, Edward L, and Joseph Gyourko. 2002. "The Impact of Zoning on Housing Affordability." *NBER Working Paper* 8835.

Glaeser, Edward L, and Bryce A Ward. 2009. "The Causes and Consequences of Land Use Regulation: Evidence from Greater Boston." *Journal of urban Economics* 65 (3): 265–278.

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

Gyourko, Joseph, and Jacob Krimmel. 2021. "The Impact of Local Residential Land Use Restrictions on Land Values across and within Single Family Housing Markets." *Journal of Urban Economics* 126 103374.

Gyourko, Joseph, Albert Saiz, and Anita Summers. 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies* 45 (3): 693–729.

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

Jackson, Kristoffer Kip. 2018. "Regulation, Land Constraints, and California's Boom and Bust." *Regional science and urban economics* 68 130–147.

Kahn, Matthew E, Ryan Vaughn, and Jonathan Zasloff. 2010. "The Housing Market Effects of Discrete Land Use Regulations: Evidence from the California Coastal Boundary Zone." *Journal of Housing Economics* 19 (4): 269–279.

Kok, Nils, Paavo Monkkonen, and John M Quigley. 2014. "Land Use Regulations and the Value of Land and Housing: An Intra-Metropolitan Analysis." *Journal of Urban Economics* 81 136–148.

Kulka, Amrita. 2019. “Sorting into Neighborhoods: The Role of Minimum Lot Sizes.”Technical report, Mimeo., University of Wisconsin.

Kulka, Amrita, Aradhya Sood, and Nicholas Chiumenti. 2023. “How to Increase Housing Affordability? Understanding Local Deterrents to Building Multifamily Housing.” *Understanding Local Deterrents to Building Multifamily Housing (March 17, 2023)*.

Mawhorter, Sarah, and Carolina Reid. 2018. “Terner California Residential Land Use Survey.” *University of California, Berkeley*.

Menendian, Stephen, Samir Gambhir, Karina French, and Arthur Gailes. 2020. “Single-Family Zoning in the San Francisco Bay Area.” *Berkeley, CA: Haas Institute for a Fair and Inclusive Society*.

Mleczko, Matthew, and Matthew Desmond. 2023. “Using natural language processing to construct a National Zoning and Land Use Database.” *Urban Studies* 00420980231156352.

Nechamkin, Emma, and Graham MacDonald. 2019. “Predicting Zoned Density Using Property Records.” *Washington, DC: Urban Institute*.

Pendall, Rolf, Robert Puentes, and Jonathan Martin. 2006. “From Traditional to Reformed: A Review of the Land Use Regulations in the Nation’s 50 Largest Metropolitan Areas.”

Quigley, John M, Steven Raphael, and Larry A Rosenthal. 2008. “Measuring Land-Use Regulations and their Effects in the Housing Market.”

Resseger, Matthew. 2022. “The impact of land use regulation on racial segregation: Evidence from Massachusetts zoning borders.” *Mercatus Research Paper*.

Rothwell, Jonathan, and Douglas S Massey. 2009. “The effect of density zoning on racial segregation in US urban areas.” *Urban Affairs Review* 44 (6): 779–806.

Ruggles, Steven, Catherine A Fitch, Ronald Goeken, J David Hacker, Evan Roberts, and Matthew Sobek. 2017. “IPUMS Restricted Complete Count Data. Version 2.0.” *University of Minnesota*.

Sahn, Alexander. 2021. “Racial diversity and exclusionary zoning: Evidence from the great migration.”Technical report, Working Paper. Online: https://drive. google. com/file/d/10 ....

Trounstine, Jessica. 2020. "The Geography of Inequality: How Land Use Regulation Produces Segregation." *American Political Science Review* 114 (2): 443–455.

Turner, Matthew A, Andrew Haughwout, and Wilbert Van Der Klaauw. 2014. "Land Use Regulation and Welfare." *Econometrica* 82 (4): 1341–1403.

Wu, JunJie, and Seong-Hoon Cho. 2007. "The Effect of Local Land Use Regulations on Urban Development in the Western United States." *Regional Science and Urban Economics* 37 (1): 69–86.

Zabel, Jeffrey, and Maurice Dalton. 2011. "The impact of minimum lot size regulations on house prices in Eastern Massachusetts." *Regional Science and Urban Economics* 41 (6): 571–583.

Zeileis, Achim. 2005. "A Unified Approach to Structural Change Tests based on ML Scores, F Statistics, and OLS Residuals." *Econometric Reviews* 24 (4): 445–466.

Appendix A Figures and Tables

Table A1 — Municipality definition by state

	Description	2010 Census count	# identified
AL	All incorporated places (167 cities, 293 towns)	460	461
AR	All incorporated places (311 cities, 191 towns)	502	501
AZ	All incorporated places (45 cities, 45 towns)	90	91
CA	All incorporated places (459 cities, 21 towns)	480	482
CO	All incorporated places (75 cities, 196 towns)	271	271
CT	All county subdivisions (169 towns)	169	169
DE	All incorporated places (10 cities, 44 towns, 3 villages)	57	57
DC	District of Columbia (1 city)	1	1
FL	All incorporated places (268 cities, 124 towns, 19 villages)	411	412
GA	All incorporated places (425 cities, 105 towns, 2 balances)	535	536
ID	All incorporated places (201 cities). Lost River city is inactive and not included in the data	201	200
IL	1 incorporated place independent of any township (Chicago). 12 incorporated places in counties where county subdivisions are nonfunctioning election precincts. 1432 functioning Census county subdivisions (townships) excluding the 13 incorporated places	1445	1444
IN	2 incorporated places independent of any county subdivisions (Indianapolis, Terra Haute). 1009 county subdivisions that are not undefined or unorganized territories excluding Indianapolis and Terra Haute cities (1005 townships, 3 towns, 1 cities)	1011	1008
IA	County subdivisions excluding unorganized territories (1598 townships, 59 cities that are either wholly or partially independent of MCDs that create 62 CCDs)	1660	1661
KS	All county subdivisions (1403 townships while 129 of them are inactive, 120 cities creating 127 MCDs)	1530	1530
KY	All incorporated places (420 cities, 1 urban county, 1 balance)	422	417
LA	All incorporated places (69 cities, 128 towns, 107 villages)	304	304
ME	County subdivisions that are not plantations, gore, American Indian reservations, or unorganized territories (433 towns, 22 incorporated places)	455	453
MD	All incorporated places (29 cities, 123 towns, 5 villages)	157	157
MA	53 incorporated places independent of MCDs. 298 county subdivisions that are towns	351	349
MI	County subdivisions that are not undefined MCDs (1123 townships, 117 charter townships, 275 incorporated places that are cities creating 293 MCDs)	1533	1540

	Description	2010 Census count	# identified
MN	County subdivisions that are not unorganized territories or undefined MCDs excluding 23 nonfunctioning townships in Lake of the Woods County (1785 active townships, 845 incorporated places creating 893 MCDs)	2678	2672
MS	All incorporated places (110 cities, 169 towns, 19 villages)	298	298
MO	331 active townships, 959 incorporated places (637 cities, 110 towns, 212 villages while 8 of the incorporated places are inactive) that are not dependent on 331 active townships	1087	1087
MT	All incorporated places (52 cities, 75 towns, 1 city with no description, 1 balance)	129	129
NE	County subdivisions that are not election precincts or election districts (435 active townships, 77 incorporated places that are independent of MCDs creating 79 CCDs)	512	501
NV	All incorporated places (18 cities, 1 place with no legal descriptor)	19	19
NH	County subdivisions that are not townships, locations, purchases, grants, or undefined, excluding Livermore town (inactive). 13 incorporated places that are independent	234	234
NJ	County subdivisions that are not undefined MCDs (242 townships, 324 incorporated places that are independent of MCDs comprised of 254 boroughs, 52 cities, 3 towns, and 3 villages)	566	563
NM	All incorporated places (35 cities, 19 towns, 48 villages)	102	105
NY	County subdivisions that are not boroughs, American Indian reservations, or undefined MCDs (932 towns, 61 cities creating 62 MCDs)	994	994
NC	All incorporated places (76 cities, 456 towns, 21 villages)	553	552
ND	County subdivisions that are not unorganized territories (1317 townships, 357 incorporated places that are independent of MCDs creating 364 CCDs)	1681	1673
OH	County subdivisions that are not undefined MCDs (1324 townships, 258 incorporated places that are wholly or partially independent creating 274 CCDs)	1598	1601
OK	All incorporated places (164 cities, 433 towns while four are inactive)	597	590
OR	All incorporated places (233 cities, 9 towns while 1 city is inactive)	242	240
PA	County subdivisions that are not undefined MCD (1547 townships while 1 of them is inactive, 1015 incorporated places creating 1027 CCDs)	2574	2572
RI	County subdivisions that are not undefined MCDs (31 towns, 8 incorporated cities independent of MCDs)	39	39
SC	All incorporated places (69 cities, 200 towns)	269	270
SD	County subdivisions that are not unorganized territories (915 townships, 311 independent incorporated places creating 320 CCDs)	1234	1222

	Description	2010 Census count	# identified
TN	All incorporated places (182 cities, 162 towns, 1 metropolitan government, 1 with no descriptor, 1 balance)	347	345
TX	All incorporated places (956 cities, 234 towns, 24 villages while 2 places are inactive)	1214	1218
UT	All incorporated places (144 cities, 101 towns)	245	250
VT	County subdivisions that are not gores or grants (237 towns actively functioning, 9 cities independent of MCDs)	246	251
VA	All incorporated places (39 cities, 190 towns)	229	228
WA	All incorporated places (208 cities, 73 towns)	281	281
WV	All incorporated places (77 cities, 148 towns, 6 villages, 1 corporation)	232	232
WI	County subdivisions that are not undefined MCDs (1257 towns, 594 incorporated places that are independent of MCDs creating 651 CCDs)	1908	1911
WY	All incorporated places (19 cities, 80 towns)	99	99
Total		32252	32232

Note. This table presents types of municipalities by state with local government entities according to the 2010 Census Guide to State and Local Census Geography. Note that the counts of municipalities include both municipalities in Core-Based Statistical Areas and those not part to any Core-Based Statistical Areas.

Table A2 — **Summary statistics of min lot size estimates (weighted by # home)**

	Percentile					
	5th	25th	50th	75th	95th	Mean
Nationwide	4,726	7,080	9,200	14,810	43,996	15,971
AL	7,000	12,323	15,515	20,308	43,560	18,383
AZ	5,227	7,209	7,989	9,148	35,223	10,947
AR	7,000	9,448	13,000	15,246	30,056	16,158
CA	4,750	6,098	7,080	7,980	12,850	7,830
CO	5,750	7,000	8,015	10,000	22,216	10,680
CT	6,534	11,326	21,344	43,996	87,991	33,000
DE	3,485	4,356	7,405	10,019	15,682	8,018
DC	3,081	3,081	3,563	5,000	5,058	4,056
FL	4,356	7,000	9,213	10,019	16,117	10,197
GA	4,879	8,712	11,361	17,999	43,560	16,508
ID	6,970	8,028	9,365	10,585	24,394	12,197
IL	3,750	6,600	9,225	11,325	43,560	13,190
IN	6,534	9,757	14,400	25,920	87,120	27,588
IA	6,820	7,841	9,525	11,996	64,033	18,210
KS	6,600	7,920	9,200	14,000	92,347	19,903
KY	5,400	8,276	10,625	13,939	25,395	12,868
LA	6,000	6,000	9,000	15,000	43,560	13,639
ME	6,300	10,019	22,651	80,150	130,680	43,461
MD	1,120	1,938	5,000	7,840	14,000	5,921
MA	7,000	11,543	20,038	40,075	80,436	26,823
MI	5,227	5,445	7,405	13,939	44,867	15,952
MN	5,663	9,583	11,326	16,510	113,256	27,307
MS	8,031	11,175	13,939	15,000	22,512	14,495
MO	4,650	8,712	10,019	12,589	22,651	13,966
MT	6,011	7,405	8,102	10,759	14,810	10,051
NE	6,350	6,500	9,000	12,197	39,600	12,690
NV	4,792	6,098	6,534	7,405	12,197	7,373
NH	7,500	16,117	43,560	87,120	133,294	52,873
NJ	2,500	5,001	8,760	15,002	44,997	14,998
NM	7,492	8,503	8,503	9,200	12,197	10,956
NY	5,001	7,500	13,463	32,942	91,912	27,978
NC	5,875	10,019	14,810	20,038	43,560	16,770
ND	5,880	7,452	8,800	11,326	67,082	16,834
OH	5,279	7,500	11,199	20,038	75,359	21,104
OK	6,998	7,919	8,400	11,761	43,560	13,785
OR	4,792	6,534	7,600	9,575	12,197	9,311
PA	2,913	5,932	10,560	21,780	50,530	18,527
RI	5,000	7,518	11,250	20,909	87,991	24,644
SC	6,875	11,500	15,246	21,780	43,560	19,221
SD	7,766	9,240	9,274	16,000	87,120	23,857
TN	7,405	10,000	15,000	22,500	43,560	19,607
TX	6,025	7,250	8,588	9,620	20,400	10,607
UT	6,534	8,276	9,148	11,326	21,780	11,682
VT	9,386	14,810	22,216	43,560	100,188	37,232
VA	3,920	6,970	10,400	15,207	26,250	12,905
WA	5,000	6,400	7,841	10,019	16,553	9,648
WV	5,702	7,501	8,337	11,931	21,653	10,798
WI	4,800	7,105	9,300	11,761	23,958	10,987
WY	6,250	7,500	8,470	9,525	13,504	11,778

Note. This table reports the nationwide and statewide summary statistics of estimated minimum lot sizes. The 5th, 25h, 50th, 75th, and 95th percentiles and the average of minimum lot sizes are computed by weighting by the number of homes in each geographic area.

Table A3 — Summary statistics of min lot size estimates (weighted by land area)

	Percentile					
	5th	25th	50th	75th	95th	Mean
Nationwide	2,499	8,985	18,730	43,556	173,804	38,194
AL	7,722	7,909	11,871	18,644	43,348	15,787
AZ	6,499	8,177	8,253	9,142	49,397	16,334
AR	6,970	11,141	13,500	21,218	62,291	23,104
CA	4,999	6,999	7,797	8,996	18,974	9,927
CO	5,992	7,188	8,400	10,865	42,689	14,841
CT	11,325	40,055	47,406	86,970	124,146	60,913
DE	3,920	7,405	9,461	15,000	25,800	12,036
DC	3,076	3,450	4,880	4,880	7,500	4,810
FL	6,050	7,590	8,494	10,018	49,085	14,242
GA	6,965	11,308	17,982	37,026	84,942	29,371
ID	6,900	8,250	11,282	26,136	43,124	21,654
IL	7,300	11,248	12,726	39,650	209,524	37,888
IN	11,195	27,748	27,748	43,516	217,364	48,353
IA	7,564	11,587	43,521	90,169	203,861	63,255
KS	7,256	7,388	12,196	125,453	216,929	60,782
KY	7,405	10,200	15,000	41,617	51,662	23,996
LA	8,050	14,999	15,920	15,920	40,647	18,731
ME	10,001	43,485	80,150	87,991	169,448	72,497
MD	1,062	7,000	8,637	13,621	31,554	12,982
MA	10,299	22,215	43,486	65,776	119,790	49,391
MI	5,968	14,723	18,252	43,124	131,116	39,314
MN	8,880	16,995	52,272	158,994	217,606	91,317
MS	1,383	11,175	13,978	17,954	40,946	17,845
MO	7,500	10,200	14,767	42,253	217,364	49,806
MT	6,000	7,405	9,375	14,767	76,230	25,675
NE	6,992	6,992	6,992	11,625	95,396	21,661
NV	5,184	6,094	6,926	9,000	22,564	9,926
NH	16,976	45,738	86,249	111,514	202,423	87,311
NJ	2,375	2,479	2,479	10,000	43,500	10,611
NM	8,260	8,500	8,500	10,649	41,382	13,785
NY	10,888	30,272	43,520	86,267	217,364	72,528
NC	7,362	11,662	15,225	21,301	43,345	19,285
ND	7,405	10,498	27,520	67,082	129,809	52,607
OH	7,473	14,280	26,400	54,363	215,700	52,354
OK	6,999	8,399	12,276	43,556	110,400	32,121
OR	3,710	4,732	7,009	10,010	32,234	14,568
PA	7,802	19,977	43,386	48,047	87,102	42,842
RI	8,000	20,000	43,124	87,556	206,910	66,445
SC	9,017	13,050	13,996	29,575	97,574	26,842
SD	10,500	17,298	17,298	42,042	118,919	37,163
TN	7,492	13,050	18,750	43,124	48,352	25,867
TX	6,761	7,800	9,374	21,632	32,539	14,516
UT	7,100	8,668	11,500	21,344	45,302	21,417
VT	14,375	39,204	45,738	81,457	113,256	63,354
VA	5,745	10,367	15,207	18,121	130,800	25,606
WA	2,731	6,395	8,401	10,414	32,234	12,236
WV	5,998	8,276	10,000	11,927	22,198	11,970
WI	6,938	9,120	12,592	18,644	33,977	18,333
WY	6,098	7,960	8,550	11,995	108,900	20,025

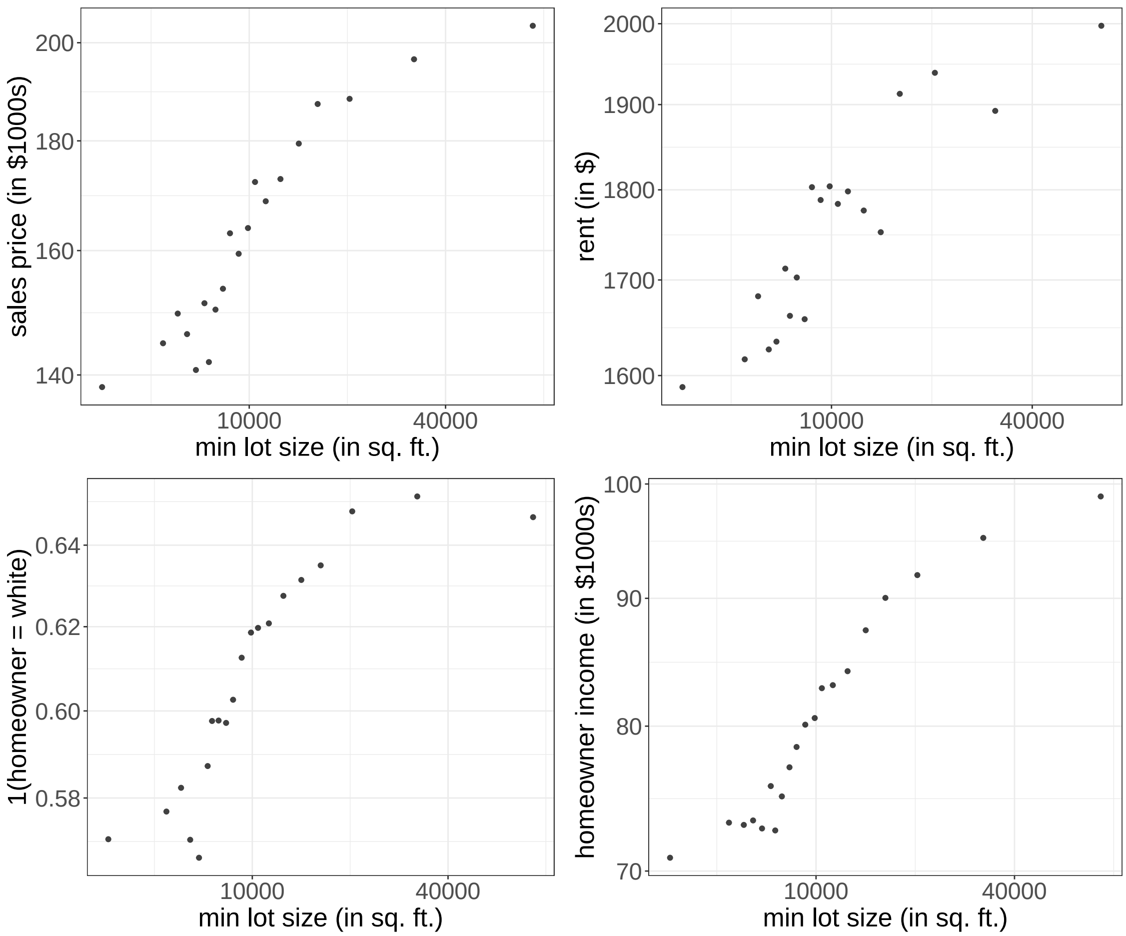
Note. This table reports the nationwide and statewide summary statistics of estimated minimum lot sizes. The 5th, 25h, 50th, 75th, and 95th percentiles and the average of minimum lot sizes are computed by weighting by land area in each geographic area.

Table A4 — Full list of control variables and their data sources

Variable Description	Source
Parcel characteristics	
Lot area	Corelogic Tax Assessor
Building square footage	Corelogic Tax Assessor
(Effective) Construction year	Corelogic Tax Assessor
# Bedrooms	Corelogic Tax Assessor
# Bathrooms	Corelogic Tax Assessor
Neighborhood characteristics (municipality-level & municipality $\times$ border region-level)	
1940 Total population	Full-Count 1940 Census
1940 Average household size	Full-Count 1940 Census
1940 Mean household wage	Full-Count 1940 Census
1940 % white	Full-Count 1940 Census
1940 % Homeownership	Full-Count 1940 Census
1940 Mean home value	Full-Count 1940 Census
1940 Mean rent	Full-Count 1940 Census

Note. This table reports the control variables used in Section 3. For the variables with missing values, I include an additional dummy variable for missing data and replace missing values with value -99 in the original variable. Lot area, building square footage, population, and mean household wage are entered log-linearly in the specifications. All other variables enter linearly in the specifications.

Figure A1 — Binscatter Plots (Full Controls With Border Distance = 0.5 km)



Note. The figure depicts the binscatter plots with 20 bins from two. The x-axis is minimum lot size (in square feet). The y-axis is sales price (in 1000 dollars), rental price (in dollars), indicator for homeowner being white, and homeowner income (in 1000 dollars), respectively. All regressions include 1940 IPUMS control variables, border region fixed effects, county-by-transaction year fixed effects, and municipality-by-school district fixed effects. Minimum lot size, prices, and homeowner income are transformed on a logarithmic scale in binscatter regressions.

Table A5 — Price Regressions without Municipality × School District Fixed Effects

Dependent Variable:	log sales price				log rent			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log MLS	0.1510***	0.0707***	0.0259***	0.0166***	0.1028***	0.0514***	0.0391***	0.0327***
	(0.0081)	(0.0054)	(0.0042)	(0.0034)	(0.0122)	(0.0060)	(0.0063)	(0.0064)
Building characteristics								
log bldg. sqft.		0.6890***	0.6383***	0.6383***		0.4993***	0.4887***	0.4862***
		(0.0085)	(0.0084)	(0.0083)		(0.0116)	(0.0116)	(0.0117)
age		-0.0048***	-0.0053***	-0.0053***		-0.0020***	-0.0022***	-0.0022***
		(0.0001)	(0.0001)	(0.0001)		(0.0002)	(0.0002)	(0.0002)
# bed		-0.0171***	-0.0183***	-0.0175***		-0.0038	-0.0038	-0.0033
		(0.0018)	(0.0018)	(0.0018)		(0.0025)	(0.0025)	(0.0025)
# bath		0.0708***	0.0742***	0.0723***		0.0366***	0.0374***	0.0361***
		(0.0022)	(0.0021)	(0.0020)		(0.0029)	(0.0029)	(0.0029)
log lot size			0.0796***	0.0793***			0.0230***	0.0228***
			(0.0030)	(0.0029)			(0.0032)	(0.0032)
Neighborhood characteristics								
HMDA % white				0.1489***				0.0128
				(0.0349)				(0.0287)
HMDA % black				-0.1864***				-0.0474
				(0.0562)				(0.0495)
HMDA Q1 income				0.0011				0.0025**
				(0.0007)				(0.0011)
HMDA Q2 income				0.0012*				-0.0020*
				(0.0007)				(0.0010)
HMDA Q3 inome				-0.0003				0.0008***
				(0.0002)				(0.0003)
Controls								
county × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1940 IPUMS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
border FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
loc gov't × SD FE	No	No	No	No	No	No	No	No
Observations	10,598,492	10,598,492	10,598,492	10,434,177	300,305	300,305	300,305	300,249
R^2	0.48888	0.59253	0.60311	0.61186	0.75335	0.80653	0.80674	0.80693
Within R^2	-0.01122	0.19384	0.21478	0.23258	0.00961	0.22315	0.22398	0.22577

Robust standard-errors in parentheses clustered at the municipality level
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note. This table reports coefficient estimates of regression 1 with building-level and neighborhood-level covariates. The outcome variable in Columns (1) to (4) is log sales price, and the outcome variable in Columns (4) to (8) is log rent. All columns include county-by-transaction year fixed effects, 1940 neighborhood characteristics, and border fixed effects but not municipality-by-school-district fixed effects. Columns (2)-(4) and (6)-(8) additionally include transaction-level building characteristics. Columns (4) and (8) also include homeowner demographics from HMDA at the neighborhood level, defined as municipality by border region.

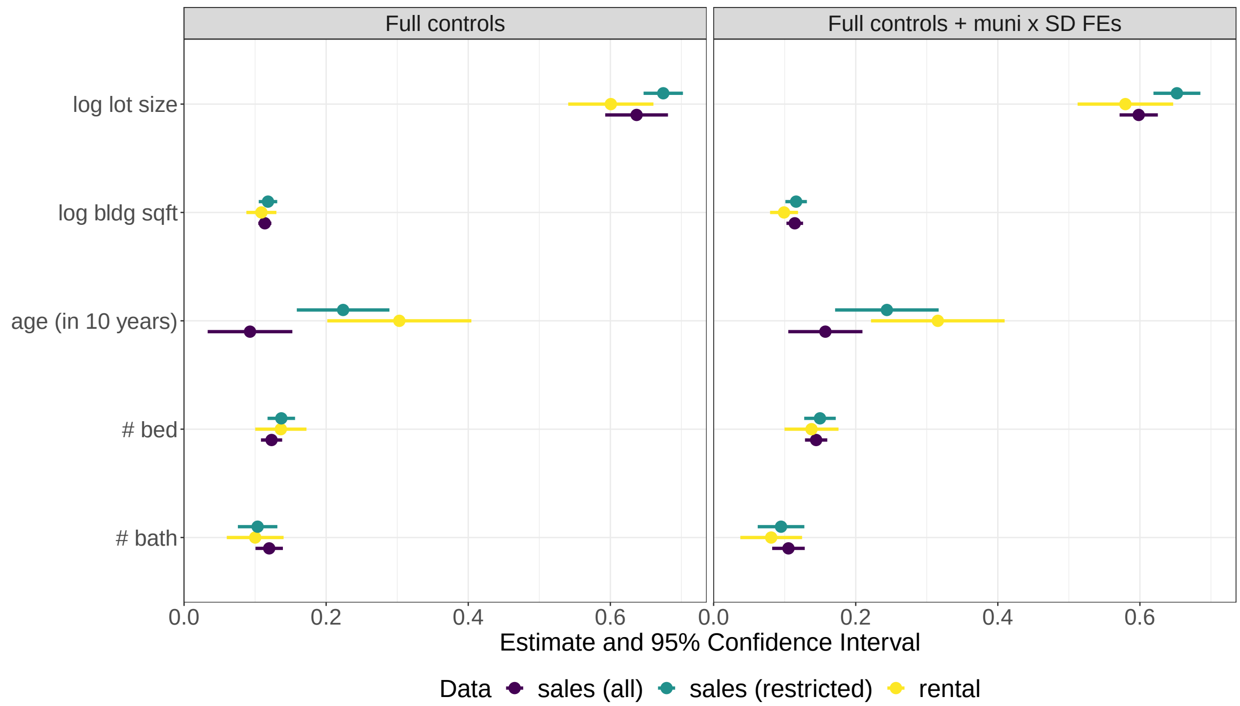
Table A6 — Price Regressions (Alternative Assumptions)

Dependent Variable:	log sales price		log rent	
Model:	(1)	(2)	(3)	(4)
log MLS	0.0638***	0.0220***	0.0884***	0.0302***
	(0.0047)	(0.0042)	(0.0057)	(0.0061)
Building characteristics				
log bldg. sqft.	0.6725***	0.6222***	0.4875***	0.4750***
	(0.0079)	(0.0081)	(0.0108)	(0.0105)
age	-0.0047***	-0.0051***	-0.0020***	-0.0022***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)
# bed	-0.0138***	-0.0151***	-0.0027	-0.0028
	(0.0017)	(0.0016)	(0.0025)	(0.0025)
# bath	0.0655***	0.0687***	0.0338***	0.0347***
	(0.0019)	(0.0018)	(0.0029)	(0.0029)
log lot size		0.0803***		0.0271***
		(0.0029)		(0.0040)
Controls				
Full Controls and FEs	Yes	Yes	Yes	Yes
Observations	10,598,492	10,598,492	300,305	300,305
R^2	0.61658	0.61884	0.81066	0.81092

Robust standard-errors in parentheses clustered at the municipality level
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note. This table reports coefficient estimates of regression 1 to understand price effect mechanisms under an alternative assumption that minimum lot size regulations do not affect other building characteristics. The outcome variable in Columns (1) and (2) is log sales price, and the outcome variable in Columns (3) and (4) is log rent. All columns include county-by-transaction year fixed effects, 1940 neighborhood characteristics, border fixed effects, and municipality-by-school-district fixed effects. The coefficient of log MLS in Column (1) measures the total price effect, and its reduction in Column (2) measures the direct effect. The coefficient of log MLS in Column (3) measures the total rent effect, and its reduction in Column (4) measures the direct effect.

Figure A2 — β_{MLS} coefficients for building characteristics (Distance = 0.5 km)



Note. The figure depicts the coefficient estimates and 95% confidence intervals of the regressions where the outcome variables are building characteristics (y-axis), and the main explanatory variable is log min lot size. Both panels include county-by-transaction year fixed effects, 1940 neighborhood characteristics, and border fixed effects. The right panel additionally includes municipality-by-school-district fixed effects. For each outcome variable in each panel, regression results using three sample data are shown: sales data (purple), sales data restricted to counties with rent data in CoreLogic Multiple Listings Service data (green), and rent data (yellow).