Are Zoning District Borders Color Lines? How Lot Size Requirements Alter Neighborhood Racial Composition *

Tianfang Cui
New York University
tc4089@nyu.edu †

Vicki Been
New York University
vicki.been@nyu.edu

October 16, 2025

Abstract

We study how minimum lot size (MLS) requirements affected the racial composition of neighborhoods since 1980. We generate a new database of the borders within local governments where we predict changes in lot size requirements. Applying dynamic border discontinuity designs over the national database, we produce causal estimates that vary over the regulations' particular characteristics. Data-driven clustering identifies which of the regulations most impact racial diversity: MLS requirements that are at least 5000 square feet larger than prevailing density patterns nearby, lowering allowed density by 3 units/acre. Our results show that those requirements persistently stymied racial integration over the following decades.

JEL Codes: H73, K42, R23, R52
Keywords: Zoning, Land use regulation, Treatment heterogeneity, Neighborhood change

*We thank Ingrid Gould Ellen, Joseph Gyourko, Jacob Krimmel, Dan Maggio, Kathy O'Regan, Martin Saavedra, Barry Scholnick and participants at Rutgers University, University of Alberta, NYU Furman Center, Toronto Metropolitan University, the Wharton School, anonymous referees and the APPAM, AREUEA and UEA National Conferences for comments. Cui gratefully acknowledges support from the Zell-Lurie Real Estate Center Research Sponsor Program at Wharton. Been gratefully acknowledges the support of the Filomen D'Agostino and Max E. Greenberg Research Account. This paper was previously circulated as "Density Zoning Interacts With Racial Diversity: New Evidence From National Data."
Although this research was conducted by researchers affiliated with the NYU Furman Center, which is affiliated with NYU's School of Law and Wagner Graduate School of Public Service, it does not purport to present the institutional views (if any) of NYU or any of its schools.

†Corresponding author.

1 Introduction

Residential racial segregation has been shown to have many costs. It increases rates of poverty in Black households, for example, and widens income disparities between Black and White households[1]. When children move out of distressed neighborhoods in racially and economically segregated areas to more diverse neighborhoods, their lives improve (Bergman et al. (2024), Chyn, Collinson and Sandler (2025)). Local governments' land use regulations, however, pose persistent barriers to achieving more diverse neighborhoods across the United States. It is well established that zoning regulations and other government land use policies segregated neighborhoods by race and class, both in intent and in practice (Rothstein (2017); Trounstine (2020)). Less well understood is which particular regulations continue to stymie integration; whether reforming those regulations is likely to result in greater racial or economic diversity; and which reforms are likely to be most effective.

Our paper quantifies the barriers to integration posed by minimum lot size (MLS) requirements, which mandate that each new housing unit use a minimum amount of land. The requirements set floors on housing quality and, in combination with other land use regulations, help to ensure that homeowners will pay the property tax required to support the community's public goods. Lot size requirements accordingly set implicit "entry fees" for access to differentiated communities (Ellickson (1977); Fischel (1987)). As the first major judicial decision to question such practices noted: "This policy of land use regulation for a fiscal end" allows each local government in effect to "build[] a wall around itself to keep out those people . . . not adding favorably to the tax base." (South Burlington County NAACP v. Mount Laurel (1975)). "Those people" the regulations keep out disproportionately include people of color, especially potential Black and Latino residents, in part because of the correlation between income and race or ethnicity, and in part because of other barriers that have limited the housing options for people of color (Rothstein (2017); Taylor (2019)). Further, because housing built according to the requirements is costly to change, those walls may persist for decades even if the regulations are changed.

MLS requirements come in different sizes, as local governments often apply multiple requirements across various zoning districts. Depending on regulatory characteristics such as the required size, the regulations' exclusionary effects may differ as well. A key barrier to studying how MLS restrictions affect neighborhood outcomes has been the lack of data on where different restrictions apply, when they were implemented, and whether they have remained in place

[1]The literature estimating the effects of persistent segregation includes, but is not limited to, Ananat (2011); Chyn, Haggag and Stuart (2025); and Chetty and Hendren (2018).

over time. We build a new dataset of where MLS requirements constrain development across the largest U.S. metropolitan areas.

Then, we study whether (and if so, which) MLS restrictions caused differences in block-level racial composition between an area subject to a particular predicted restriction and adjacent areas not subject to that restriction. We estimate effects for different MLS requirements by analyzing the new data with techniques at the intersection of causal inference and machine learning. We apply our methodology to national data since 1980, the year after which non-White households moved out of urban cores in increasing numbers. Our results indicate how MLS requirements of various sizes have, since then, impacted the racial integration of single-family neighborhoods.

We first introduce an automated method trained to detect shifts in lot sizes that signal a potential change in the applicable requirement. Our technique pieces together a national sample of *lot size discontinuities* — borders of areas in which differences in lot size are detected. We exclude those segments most likely to reflect differences in land characteristics around the border, which could drive lot size discontinuities regardless of a mandated minimum size. We also only use MLS requirements within each local government's jurisdictional boundaries.

Restricting our analysis this way allows us to filter out lot size discontinuities, in order to minimize endogeneity issues that might make some neighborhoods to be limited by a particular MLS requirement over others. The built environment, prior to MLS adoption, may have been influenced by earlier characteristics and regulations across jurisdictions. Developers may also subdivide certain land parcels into lots of standard sizes, subject to market demand expectations: consequently, we identify effects without using all blocks in a MLS regulated zoning district, especially exploiting local variation on a subsample of border segments.

The automated procedure identifies shifts in allowable density within localities across 120 of the largest U.S. metropolitan areas, where over 240 million Americans live. We estimate treatment effects using narrow bands of blocks within the same local government, but across which MLS requirements change. When policy variation across all border segments are pooled together, the average effect of minimum lot sizes on racial disparities is small. In both 1980 and 2020, the average difference across all lot size borders in block-level shares of people of color is less than one percentage point.

The average effect, however, may not reveal contexts in which the lot size requirements do matter. Suppose a MLS requirement imposes a very small difference in lot sizes between the regulated district and adjacent areas not subject to that requirement. The mandated difference may not have changed the price of housing or neighborhood characteristics enough to affect the

relative racial composition between the two. It may also be that the highest MLS requirements are across the zoning border from homes set on somewhat smaller, but still relatively large, lots. Both these kinds of homes are likely affordable only to higher income households, which means the requirements likely have only small marginal impacts on the neighborhood's future diversity.

Accordingly, we assess whether different types of MLS requirements, or different cross-border contexts, might cause different effects. We do that by estimating effects across four *impact clusters* of minimum lot sizes. Using a data-driven process based on the causal forest approach to heterogeneous treatment effects (Athey and Wager, 2019), we identify a subsample of areas we refer to as the *high-impact cluster.* MLS requirements in this cluster yield residential densities between 2 to 7 dwellings per acre (DUPAC), which means lot sizes of between 6,000 to 22,000 square feet. Relative to adjacent areas not subject to the MLS requirement, high-impact requirements decrease the density by about 2.8 DUPAC, Requirements in the high-impact cluster account for 14 percent of our national sample. MLS requirements in different cities are matched to this cluster, or one of three others, if the regulatory minimums, as well as the degree to which they changed density in their respective neighborhoods, cause similar changes in racial composition relative to surrounding residential development.

Taking into account the heterogeneity of the MLS requirements, we show three main results. First, in the high-impact cluster, disparities across the discontinuity borders in the share of non-Hispanic white residents are between 2 to 5 percentage points. The effects of "high-impact" MLS requirements are driven mainly by disparities in where Black residents live. Null effects cannot be rejected for the share of Hispanic and Asian residents. Second, the estimates show that disparities in racial diversity caused by high-impact cluster MLS requirements grew between 1980 and 2020. For regulations in the other clusters, disparities are less substantive in magnitude and do not increase over time, but are persistent.

Finally, requirements in the high-impact cluster rarely involve the very large lot requirements most subject to existing legal criticism and analysis (Ellickson (2021); Desegregate CT (2025)). More surprisingly, few MLS requirements of more than a half acre yield large drops in density from their surrounding areas. Any effects those large lot requirements have on racial composition are too small for us to detect with statistical significance. In general, a majority of detected border segments (including ones involving smaller sizes) do not involve substantial changes in the regulated density compared to areas across the border. Our findings imply that a research design using comparisons only to adjacent areas offer conservative estimates of how

the largest MLS alters racial diversity within local governments. [2]

We verify that other potential differences in housing and neighborhood quality across predicted zoning border segments do not drive our results. We break our sample down by the median year built of properties near our predicted border segments, which reveals that similar effects hold for both developments built closer to the 21st Century and for those built earlier. Our results also are robust to specification choices with the border discontinuity design. We provide evidence suggesting that the predicted segments are not driven by the location of other spatial features that can segregate neighborhoods, nor were they associated with racial disparities among neighborhoods leading up to 1980. Hence, the prevalence of predicted segments across different U.S. urban contexts also serves as a measure of MLS requirement restrictiveness at different spatial geographies.

Our paper makes several contributions. First, it provides a new source of data predicting the location and timing of MLS requirements at a neighborhood level. Recent papers that generate spatial data regarding density restrictions like minimum lot sizes (Bronin et al. (2023), Gallagher, Shertzer and Twinam (2024), Song (2025), Macek (2024)) do not try to predict which regulatory borders were drawn as an endogenous response to limit change around specific types of development. We identify those borders where MLS requirements were most likely to bind development, and less likely to merely follow spatial borders and barriers already in place, through a reproducible machine learning procedure. Our dataset of border segments is designed to isolate those border discontinuities that can identify causal effects of MLS requirements. The dataset is also unique in incorporating information on when the homes subject to particular requirements were built, which allows us to better estimate MLS effects over time and to control for housing age as a potential confounder.

Second, our paper contributes to a wide literature on the factors behind neighborhood racial composition, with implications for understanding metropolitan segregation more broadly. A growing literature tests the extent to which different U.S. ethnic and racial groups have been constrained in accessing neighborhoods, both due to coercive collection action in the past (Li, 2023) as well as the compounded consequences of past policies (Aaronson, Hartley and Mazumder (2021); Sood and Ehrman-Solberg (2024); Bagagli (2023)). [3] Policy reform efforts that cite postwar zoning regulations like MLS requirements as a constraint on residential integration (White House Council of Economic Advisors, 2021), refer mainly to results from

[2] Had localities been more willing to map one-acre MLS districts to 10,000 square feet MLS districts, controlling for all other factors, causal inference on observational data could find larger and more precise effects on racial diversity of the largest MLS requirements.

[3] *Constraints* such as regulatory barriers are distinct and can complement the importance of *preferences* over neighborhood composition, such as a preference to living next to people of the same race or ethnicity as oneself.

Rothwell and Massey (2009). A more recent contribution using national data is Song (2025), who estimates demographic sorting with Home Mortgage Disclosure Act data. Our primary methodological improvement over both papers is in recognizing that the effects of regulatory constraints may vary by the characteristics of the regulation and the context in which it is applied. We then verify the robustness of results on which specific characteristics make certain MLS requirements impactful, using two research designs and multiple specifications. We extend beyond the 50 MSAs that Rothwell and Massey (2009) surveyed. Compared to Song (2025), we provide results for over 40 years and for multiple racial and ethnic groups to further understand the dynamic mechanisms behind MLS effects.

Third, to identify regulations that we argue matter for racial diversity, we also refine recent empirical methods that both estimate heterogeneous treatment effects and improve their interpretability. [4] MLS requirements are the keystone of density zoning: as noted economist and land use law expert Robert Ellickson has written: "Drafters of zoning ordinances typically regard minimum lot-size requirements as the most salient of their zoning controls." (Ellickson, 2021) By highlighting the specific regulations that cause the greatest changes in *outcomes*, we connect the land use regulation literature defining the overall restrictiveness of a jurisdiction's land use practices (Gyourko, Saiz and Summers, 2008) with the literature focused on measuring how land use regulations distort outcomes from market equilibrium (Albouy and Ehrlich (2018); Gyourko and Krimmel (2021)).

Fourth, as state and local governments across the United States grapple with how to reform land use regulations to address the housing crisis, our evidence helps bolster the case that allowing denser housing could both increase housing supply and increase racial diversity. Moreover, our evidence suggests that the first priority should be reducing or eliminating those MLS requirements that have the characteristics of what we identify as the high-impact cluster. That is, increasing density may be more effective in neighborhoods with less restrictive lot size requirements than the one or two acre minimums that often serve as the poster child of exclusionary zoning (Gardner, 2023). What may seem like only moderately restrictive MLS requirements do in fact affect the racial composition of the neighborhood. In neighborhoods affordable for a wider range of potential residents, allowing what has been called "gentle density" (Baca, McAnaney and Schuetz, 2019) may be a feasible reform to increase housing supply that has the added benefit of increasing racial and ethnic diversity in the affected neighborhoods.

Finally, in its most recent decision regarding the Fair Housing Act (Texas Department of

[4] Rehill (2025) provides an overview grounded in interpretations of the causal forest model.

Housing and Community Affairs v. Inclusive Communities Project, Inc., (2015), the Supreme Court insisted that plaintiffs seeking to prove that a policy has a disparate impact on different racial groups must point to a specific policy (as opposed to, for example, a "one-time decision"), and show "robust causality" between that policy and the disparate impact. On the one hand, our findings should help litigants satisfy that burden of proof for challenges to the specific lot size requirements that share the characteristics of the cluster we identify as having significant and meaningful effects on racial composition. On the other hand, the difficulties we illustrate in isolating the causal impact of specific MLS requirements also reveals that the Court's requirements for causal proof render the concept of disparate impact almost impossible to use.

The rest of the paper proceeds as follows. Section 2 explains the automated procedure we use to detect border segments where we predict that MLS requirements change. Section 3 discusses how MLS requirements might affect the racial composition of the regulated neighborhoods. Section 4 explains how we clustered MLS requirements according to the extent of their impact, and describes our identification strategies. Section 5 presents our main results, while Section 6 discusses the results' policy and legal implications. Section 7 concludes.

2 Identifying Lot Size Discontinuities

In this section, we first describe how minimum lot size requirements function. Then, we describe the detection procedure we use to create our dataset of border segments where we predict that minimum lot size requirements change.

2.1 Regulatory context

Recent legal research shows that MLS requirements were first used to regulate density between the 1920s and the 1940s (Gardner, 2023). After New York City adopted the first comprehensive citywide zoning ordinance in 1916, jurisdictions across the country began to adopt zoning ordinances that established different districts within the locality (Appelbaum, 2025). The ordinances specified individual regulations for each district, including, but not limited to, whether residential use is limited to single-family; the size of the lots; size and height of buildings; share of a lot that a building may occupy; size of yards, courts and other open spaces; and other rules that determined the general density at which land within the district could be developed. The ordinances were accompanied by zoning maps that assigned land to those districts and recorded the boundaries of each district.

While primary sources on when and what kinds of density restrictions were adopted in jurisdictions across the United States over the ensuing decades are rarely available, Cui (2024) uses an algorithm to predict which MLS requirements local governments adopted, and when they did so. Appendix Figure B.1 plots results from that paper, which infers the rate of MLS adoption from 1930 to 2010. Cui (2024) finds that MLS requirements were imposed widely after World War II. Four-fifths of the towns and cities that had at least 5,000 residents by 2010 had adopted at least one MLS requirement by 1970. Jurisdictions commonly adopted different MLS requirements for different neighborhoods; by 1970, two-thirds of that same subpopulation of cities had adopted some MLS requirement above 7,500 square feet.

2.2 Overview of border detection

The MLS requirements predicted by Cui (2024), like other *non-spatial* zoning data used in the literature, cannot produce the borders between districts regulated by different MLS requirements.[5] Information about those borders would allow us to evaluate the effects different MLS requirements have on outcomes of interest by comparing the outcomes between otherwise similar districts. Because some parts of the border between two districts may have been drawn to separate areas already developed before MLS requirements were adopted, the comparison must also isolate those developers' decisions from the effects of MLS requirements. We build a dataset that predicts the borders that likely demarcate changes in MLS requirements for districts in which it is probable that the MLS requirement, rather than market forces, is the binding constraint.

We begin by identifying the presence of lot size discontinuities — areas in which we see lot sizes differ in what appear to be systematic patterns. The intuition is as follows: if we knew a certain MLS requirement applied in one zoning district within a jurisdiction, we could start to identify the district's boundaries by locating clusters of residential lots with sizes around the minimum requirement. If we cannot detect any lots smaller than the MLS around a cluster, we would assume that the cluster is in the interior of a district with the requirement. If we instead find that surrounding lots are smaller or larger than the MLS requirement, that would suggest the cluster is close to a border between two districts with different binding MLS requirements.

Our procedure operationalizes this intuition by splitting up jurisdictions we analyze into *hexagonal tiles*. Each tile has a radius of 300 meters, which generally covers several blocks of

[5] The existing literature that retrieves regulatory outputs by surveying local officials ((Gyourko, Saiz and Summers, 2008); Pendall, Puentes and Martin (2006); Rothwell and Massey (2009)), or natural language processing of ordinance text listing different requirements (Mleczko and Desmond (2023), Bartik, Gupta and Milo (2024)) also cannot produce this information, as their source material is not the jurisdictions' district maps.

development. Given an external dataset of predicted MLS requirements in a jurisdiction, but one that lacks spatial context, our procedure iterates over the requirements. Our procedure examines geocoded lots within each tile and determines whether a meaningful number of lots within the tile fall above and below that potential MLS requirement.

Using machine learning techniques further explained in the next subsection, we then estimate the zoning district borders by drawing a line that separates all lots smaller than the MLS requirement from lots equal to or larger than the requirement. The existing literature points out, however, that a district's borders may curve and wind to follow the natural landscape, separating more desirable locations from less desirable ones (Turner, Haughwout and van der Klaauw (2014); further referenced in Kulka, Sood and Chiumenti (2023)).[6] To be conservative about whether lot size discontinuities were likely to have been driven by MLS requirements, rather than by market judgments about the desirability of the areas on either side of a natural feature, we restrict our predictions about zoning district borders to only those straight lines that are as long as a hexagonal tile's diameter, and which still adequately separate lots above and below the MLS requirement.

Our procedure accordingly is not intended to recover the entire border of each zoning district within a jurisdiction. Even if we had access to the zoning map for each jurisdiction and could see the exact border of every district in a jurisdiction, the map would tell us the borders only for a particular point in time. We would be unable to discern whether development within the district had been bound by other regulations prior to the date the map was adopted. We would also be unable to tell whether the district boundaries were drawn around existing development that had been shaped by market forces, or other factors, rather than earlier zoning regulations. To use the maps for causal inference, we would ideally filter out those segments by linking each segment to administrative data unavailable on maps, such as statistics about surrounding homes' age.

Recent literature that generates national spatial data in density restrictions like minimum lot sizes (Bronin et al. (2023), Song (2025), Macek (2024)) do not gauge how likely borders reflect the districts' endogenous features. Gallagher, Shertzer and Twinam (2024) advances the literature by using changes in zoning boundaries identified from a panel of amended zoning maps to account for that endogeneity problem. However, their data are only available for one metropolitan area. We improve upon the literature by seeking to isolate those border segments most appropriate for causal inference for jurisdictions all across the United States. We identify

[6] Conversely, Turner, Haughwout and van der Klaauw (2014) observes that straight sections of boundaries used by U.S. local governments are more likely to be drawn based on land surveying formulas, which separate lots "without regard for local physical geography."

a subsample of border segments that satisfy three properties. First, the subsample is national in scope. Second, it identifies only the borders of zoning districts within the same jurisdiction, not of districts that are in different jurisdictions, in order to control for interjurisdictional differences that might affect the racial composition of neighborhoods. Third, the border segments are selected so they are less likely to be drawn around natural features or other differences in land quality that might affect the racial composition of neighborhoods on either side of those features.

2.3 Implementing the automated procedure

Identifying tiles that contain predicted MLS changes. We first load boundaries of all local governments holding zoning power, whose boundaries are standardized by the U.S. Census Bureau as the TIGER shapefiles. In order to process multiple governments in parallel, we group the jurisdictions by the Metropolitan Statistical Area (MSA) under which they fall. Over all land in a MSA, we construct a regular hexagonal tiling; each hexagon, in every MSA, has uniform side length R. no matter the MSA, each hexagon has side length and hence radius R. We focus on *interior tiles*, in contrast to *boundary tiles* that overlap with boundaries of separate jurisdictions.

We merge those tiles with geocoded records in the CoreLogic tax assessor database that contain lot-level characteristics for individual single-family and duplex homes. We also merge in the estimates of what MLS requirements were imposed in jurisdictions across the nation from Cui (2024). That paper predicts MLS requirements based upon where lots bunch at particular sizes. When significantly more lots are developed at one size than at marginally smaller sizes, for both old and newer homes, the bunching behavior signals market distortions caused by MLS requirements. The final bunching-based predictions are fine-tuned to match MLS requirements found in historical zoning ordinance, so they identify the requirements applied on districts where developers would have built on lots smaller than the requirement allows.[7]

We keep the tiles that, for at least one of a jurisdiction's MLS requirements predicted through bunching, contain sufficient numbers of lots sized around a particular requirement $\underline{\ell}$. By "around", we mean lots either at size $\underline{\ell}$ or somewhat larger than that size (formally defined as at least $\underline{\ell}$ square feet and no more than an upper bound $M \cdot \underline{\ell}$. Additionally, we keep only those tiles that have some positive number of lots at sizes less than $\underline{\ell}$ square feet. Panel (a)

[7] The MLS requirements predicted in Cui (2024) will exclude MLS requirements that do not produce bunching. At the same time, limiting analysis over just the predicted lot sizes in that dataset is a guard against straight line divisions between two subdivisions. The "false positive," in that case, stem from a developer's decision to create lots somewhat different in size than neighboring subdivisions for market-based reasons.

of Figure 1 shows how the tile filtering occurs for one requirement of Lower Merion Township, a suburb of Philadelphia, PA: a 30,000 square foot MLS applied to one zoning district.[8] Out of all the interior tiles shown in the left inset of Panel (a), we filter out those tiles failing to meet one of the two conditions listed above.

Returning to our illustrative jurisdiction, The right inset of 1 (a) highlights all the lots between $\underline{\ell}$ and $M \cdot \underline{\ell}$ square feet. Consistent with our goal of finding the border of a particular contiguous MLS district, the tiles we find are not scattered randomly and disconnected from each other. They instead are clustered, with multiple tiles neighboring each other. Some tiles we filtered out contain isolated collections of lots with sizes close to the MLS. However, if there are less than N lots sized at or somewhat above the predicted MLS requirement, we exclude the tile to be conservative about whether the larger lots in the area were based on developers' choices or on MLS requirements. Alternatively, those tiles could lack any lots that are smaller in size than the MLS requirement, in which case we infer that a zoning district border does not run through the tile.

As implemented, three hyperparameters can be adjusted — the hexagon length R, the upper bound parameter M and a lot count threshold $\underline{N}$ — to meaningfully affect which tiles would be processed. Table 1 lists values for all hyperparameters in our preferred calibration, which are discussed further below and in Section 2.4.

Identifying border segments where MLS requirements change. For each tile that remains after our filters, we then must estimate which areas inside the tile are likely in the MLS constrained zoning district, so only lots equal to or somewhat larger than $\underline{\ell}$ square feet are allowed? Conversely, which areas likely are not bound by that requirement and lots accordingly could be smaller? We can formulate the problem as the estimation of a *decision boundary,* such that one side of the boundary is classified as being in the MLS constrained district and the remaining area is classified as not being subject to that constraint. The decision boundary then functions as an estimate of the zoning border segment running through the tile.

A common classifier to construct such a decision boundary is the support vector machine (SVM). In the left inset of Figure 1 Panel (b), we fit a SVM on lot-level data within the tile, restricting the boundary to be linear over the lots' longitude and latitude.[9] The output border segment, in blue, minimizes misclassification between the category of lots at least as large as

[8] Lower Merion has five MLS requirements as estimated by the Cui (2024) data, so the procedure runs five times.

[9] Using function expansions such as the "kernel trick" would allow us to fit nonlinear boundaries using SVMs. However, that goes against our goal of discovering straight segments of zoning boundaries to minimize boundaries more likely to follow natural features that might affect market demand for the land.

the requirement demands and the category of lots smaller than the MLS requirement. The training problem is complicated because MLS requirements typically specify a minimum, but no maximum: development larger than the MLS requirement may occur on either side of the border. Accordingly, a predicted border segment that perfectly overlaps with a straight section of an actual zoning border may not produce a misclassification rate of 0%.[10]

To visualize an acceptable amount of misclassification error m_{err}, the right inset of Panel (b) highlights both the lots equal to or larger than $\underline{\ell}$ square feet where the MLS requirement is not predicted to apply, and the lots smaller than $\underline{\ell}$ where the requirement is predicted to apply. The border segment is dropped at this stage if $m_{err} \leq \overline{M}_{err}$, a threshold hyperparameter. In our example, $m_{err} = 16\%$. Table 1 shows the calibrated level of the error threshold, which is set at $\overline{M}_{err} = 35\%$. In Appendix C, we explain the exact objective function in more detail.

In addition, Appendix C explains the technical details about how we resolve border segment conflicts resulting from running the procedure for the different MLS requirements predicted by Cui (2024), independently from one another. The last row of Table 1 references a parameter that governs a K-Nearest Neighbor (KNN) algorithm we use in resolving conflicts. Appendix Table C.1 also offers a full list of every calibrated hyperparameter we use across our multiple machine learning steps.

2.4 Training the automated procedure

The border segments identified by the automated procedure reflect where there are observed changes in the lot sizes of what was built. Up to now, the main way to ensure that the border segments we predict are based on real zoning boundaries is running the procedure for only lot sizes that Cui (2024) identify as MLS requirements. It remains possible, however, that those MLS requirements are measured with error. Additionally, the actual MLS requirement in some neighborhoods may be a lot size commonly supplied by developers following market forces, rather than because regulations required that size.

To address those concerns, we further refine the final predicted border segments we use in our analysis by validating the segments generated in each jurisdiction in a major metropolitan area against a map of actual zoning boundaries for that jurisdiction, then applying what we learn from that validation exercise to the entire sample. The idea is that if the border segments we generate include "false positives" — straight segments not close to any actual zoning boundary where MLS requirements change — that would be reflected in an abnormally large

[10] Another reason is local governments have some discretion to issue *zoning variances* for certain developments, which can be invoked, in some circumstances, to build lots smaller than $\underline{\ell}$ in a district with the requirement.

distance between the segment and any real boundary.

Formally, given a zoning map with actual zoning boundary segments, our calibrated hyperparameter vector θ minimizes the average nearest point distances between the sets of predicted border segments $\hat{\mathbf{z}} \in \hat{\mathbf{Z}}(\theta)$ and of actual segments $\mathbf{Z}^{MLS}$:

$$\theta^* = \arg\min_\theta \frac{1}{|\hat{\mathbf{Z}}(\theta)|} \sum_{\hat{\mathbf{z}} \in \hat{\mathbf{Z}}(\theta)} \min_{\mathbf{c_1} \in \hat{\mathbf{z}}, \mathbf{c_2} \in \mathbf{Z}^{MLS}} ||\mathbf{c_1} - \mathbf{c_2}||. \tag{1}$$

The zoning map we use is the Metropolitan Area Planning Council (MAPC) Zoning Atlas, a spatial dataset of zoning in the Boston metropolitan area also used for validation in Song (2025) and Kulka, Sood and Chiumenti (2023). We use all actual zoning boundaries from the Zoning Atlas that separate districts with different MLS requirements for residential development and have a segment that overlaps with a straight line for 200 meters or more. For each jurisdiction within the metropolitan area, we compare all our predicted border segments with all the straight zoning segments extracted from the Zoning Atlas. We calculate the nearest point distance to zoning boundaries for each predicted border segment, with the objective over a training or validation set being the grand mean of the distances.

We also fit an additional random forest model on the MAPC map to prune outlier predicted segments far away from zoning changes. We simultaneously calibrate hyperparameters governing the border detection procedure and the random forest pruning model, using a 10-fold cross-validation scheme that iterates predicting on a tenth of Boston metro towns whose zoning is left out of the training set.

After calibrating the automated procedure parameters on the Boston metropolitan area zoning data, we verify goodness of fit with the filtered subset of actual straight zoning district border segments. We also offer some evidence that there is no serious overlap between our predicted border segments and other spatial features within cities that, as the literature reveals, form constraints that might affect the demographics of the area even absent zoning. Those results suggest causal inference using these predicted border segments is less susceptible to confounding long-run dynamics than would be the case with the full population of possibly endogenous zoning boundaries (Maheshri and Whaley, 2025).

Table 2 summarizes the results, starting with the degree of overlap between predicted border segments and actual straight zoning district boundary segments in the Boston metropolitan area. We define overlap as whether the distance between a detected discontinuity and an actual straight zoning boundary segment is 350 meters or less.[11] Panel (a) shows that without

[11] 350 meters is the median value of the tile side lengths R we considered when calibrating the procedure.

pruning of outlier border segments, about half of predicted border segments do not overlap with zoning boundaries. This rate is substantially reduced following the "pruning" of border segments using our random forest model, leaving only 27% of predicted border segments that do not overlap with actual zoning boundaries. We also note that a sizable share of actual zoning boundaries, 53%, are far away from any detected lot size discontinuities. By construction, our detected discontinuities use lots developed for single family homes or duplexes. We interpret the zoning districts we do not get close to as the ones that vary primarily on use restrictions, like restrictions on building multifamily apartments.

With Panel (b), we analyze more overlap rates involving three definitions of zoning borders: border segments predicted by our procedure, the subsample of straight zoning district boundaries used as training data, and all actual zoning district boundaries reported in the Boston Zoning Atlas. We measure overlap between spatial features within a jurisdiction that are likely to affect the demographic makeup of what surrounds the features. By "overlap," we mean whether a border segment we've identified is within a 175 meter buffer centered on those spatial features [12]

First, we map primary roads and thoroughfares in the Boston metro area, which we extract from the Census Bureau's roads database. Evidence shows that the construction of interstate highways led to sorting along racial lines (Mahajan (2024); Weiwu (2024)), and indeed that the path of interstate highways was sometimes chosen to further segregation (Archer (2021); Karas (2015)). However, whether major roads that are not interstates have the same effect on segregation appears to depend on local context (Noonan (2005), Kramer (2018)). Our definition of a "primary road" is based on roads mapped out to be 3 kilometers or longer in Boston, which includes more variety in road structure than interstates.

Then, we map all boundaries in the 1930s HOLC Maps that divide neighborhoods based on perceived mortgage risk. We also map railroad tracks from Census Bureau data. Evidence that those two spatial features accelerate racial change on one side versus another are in (Aaronson, Hartley and Mazumder, 2021) and (Ananat, 2011), respectively. Finally, we map boundaries of school catchment areas from the 2015 national SABS database (Geverdt, 2018), reflecting changes in school assignment and possibly changes in school quality that may further neighborhood racial segregation (Richards, 2014).[13]

Three main facts emerge from calculating overlap rates for our three zoning border types.

[12] Analogously, the overlap definition in Panel A can be viewed as both the discontinuity and the zoning segment falling in a buffer of 175 meters, centred around the midpoint of the two.

[13] Conversely, Saporito and Van Riper (2016) and Monarrez (2023) offer evidence the area boundaries may be drawn to integrate students by race.

First, all three types of segments overlap substantially with the road network, to a greater extent than with the other spatial features. Second, segments predicted by our automated procedure are less likely to overlap with the other spatial features than the two other border types, especially when it comes to overlap with rail tracks and the lowest graded HOLC neighborhoods. Finally, our predicted segments have significant overlap with elementary school catchment areas and with the dividing line between class B and C HOLC neighborhoods. Similar degrees of overlap with catchment areas, however, are also observed for the full universe of the actual zoning district borders and the subsample of the straight segments used to train our procedure.

A potential concern is whether our procedure performs differently on residential areas that were subdivided before usage of MLS requirements than on areas developed following regulatory adoption. To illustrate this concern's importance, we calculate overlap rates like those in Panel (a) for two subsamples of Boston municipalities. We output a sample of 13 jurisdictions that, as of 1940, had less than a fifth of their peak population.[14] As such, they represent "post-war development" municipalities where much of the housing was built after World War II, by which point MLS requirements had been widely adopted. In those jurisdictions, the overlap rate of detected segments on actual zoning boundaries remains at 68%; 57% of actual zoning boundaries are far from any detected segments. These rates correspond to the rates after the random forest pruning step, which was 72% over all municipalities in Table 2. The role of the random forest is to filter out lot size discontinuities with lots of half an acre or higher; in Boston, we see the patterns as private development in more distant suburbs, even if not mandated by MLS requirements.

Additionally, we can output a sample of 11 jurisdictions that, prior to 1940, already were at 80% or more of their peak populations. They represent "pre-zoning development" cities where, even if residential areas subdivided before MLS requirement adoption match with today's zoning lines, our procedure limits their inclusion in our sample. In those jurisdictions, the overlap rate of detected segments on actual zoning boundaries increases to 77%;[15] 72% of actual zoning boundaries are far from any detected segments.[16] For the segments we do match to actual zoning boundaries, the median year built of homes on the MLS restricted sides is 1950. This statistic suggests that in the pre-zoning development cities, our procedure detects peripheral

[14]Throughout this section, we use the Schmidt (2018) panel of Census population for incorporated places, which go back to the 19th Century for New England towns.

[15] his statistic excludes pre-zoning development jurisdictions with no MLS predictions in the Cui (2024) algorithm, such as Cambridge, MA.

[16] Inspecting what makes up the larger share of actual zoning segments lacking overlap with detected segments, we found those segments involve districts separating residential uses unrelated to single-family homes (e.g. apartments and townhomes), or "overlay" districts that stipulate requirements on redeveloping built-up areas for new residential development.

parcels of land that remained undeveloped in the 1940s, but were zoned and developed soon after.

The array of overlay comparisons we make shows that our empirical strategy is isolating those particular neighborhoods in which we can measure the causal effects of MLS requirements. In those neighborhoods, any divergence in racial composition is more likely to be caused by the governing MLS requirements than by other factors that may have determined how the zoning district's boundaries were selected. Because the spatial features we test for Boston are all available nationally, we control for the presence of overlap in our main specifications.[17] As explained in more detail in Section 2.5, our processes focus on the top tercile of U.S. MSAs by population. Appendix Table A.2 reruns overlap rates in Table 2 for the national sample. We observe that overlap rates do not differ markedly from what we found for Boston, offering assurance that variation in the selection of zoning district boundaries in the Boston metro area is representative of national trends.

2.5 Final analysis dataset

We iterate the detection procedure across 120 of the largest Metropolitan Statistical Areas (MSAs) in the United States, where 241 million Americans live as of 2020. To link our predicted border segments to demographic data, we use the smallest geographic unit where those statistics are available: Census block shapes and statistics provided by NHGIS (Manson et al., 2024). At such a fine level of granularity, public tables based on the 100% short-form Census report a limited number of demographic responses. Across all Census waves, we observe total population and households; breakdowns by race, age and gender; as well as whether the resident owns or rents their housing unit.

For any MLS requirement $\underline{\ell}$ reflected in a local lot size discontinuity, we split blocks overlapping the interior tiles around the predicted border segments into two. The blocks on the side predicted to be bound by the MLS requirement are "treated" blocks. Likewise, blocks overlapping with the areas within tiles classified as not bound by that MLS requirement form "comparison areas" with higher densities than the treated area.

For each MSA where predicted border segment data are available, we keep Census blocks in the 1980-2020 decennial censuses that have at least 70% of their area in the treated and comparison areas we identify[18]. We filter out blocks that straddle both sides — those that have

[17] Apart from Census TIGER shapefiles that are produced nationally, the SABS database contains a national sample of school districts, while the Mapping Inequality project has produced a shapefile of the HOLC maps.

[18] We note that every MSA we process is in the NHGIS from 1990 onwards, but some MSAs lack 1980 data. NHGIS is still in the process of digitizing block shapes for every MSA in the 1980 Census, though they are not

at least 7% of their area in both the treated and the comparison areas — because it is unclear how we should assign them. To avoid dropping larger Census blocks not fully contained in a single interior tile, we also employ an iterative algorithm imputing a binding MLS for all tiles neighboring a tile with a discontinuity, if lots in that tile are almost all larger than the size at which lots were bunched in the initial tile.

Our final output is a repeated cross-section of demographics at the block level, for blocks on the treated and comparison sides of different border segments exhibiting lot size discontinuities. [19] In some specifications, we also group together all block-level variables matched to the treated or comparison area of a specific border segment. Appendix Table A.1 shows that the data cover 42 states, then compares within those states by taking means of block counts across Censuses. The states with the most segments are also the most populated, like California, Florida and Texas. Our procedure is more likely to skip over sections of Southern metros, where annexation complicates the assignment of nearby lots into different jurisdictions, while processing more of the lots inside jurisdictions in the Northeast and Midwest.

Using CoreLogic assessor records, Table 3 presents summary statistics on residential built form surrounding border segments for predicted MLS requirements. Column (1) is tabulated at the block level for areas we identified as subject to a particular MLS requirement. The average block in our data belongs to a district with a predicted MLS of around 10,000 square feet. The median home in those blocks was built in 1960. Using central business district locations collected by Holian and Kahn (2015) and Cui (2024), we calculate that the average block is 26 kilometers away from the CBD for the metropolitan area it is in.

The spatial diffusion of our detected MLS requirements is important. Baum-Snow (2023) and Bartik, Gupta and Milo (2024) observe that the further a block is from the city center, the less dense housing supply is in equilibrium, consistent with the monocentric city model. If the MLS requirements we predict, based upon our border segments with lot size discontinuities, were concentrated only on the urban periphery, then the regulations would not produce meaningful wedges between observed densities and what we would see in unregulated spatial equilibrium.

Contrary to that hypothesis, blocks in treated areas have a standard deviation of 19 kilometers away from the CBD. By splitting the sample of regulated blocks based on MLS requirement sizes in columns (2) to (5) of the Table, we also see standard deviations of 19-20 kilometers from the CBD both within the treated blocks that we predict are subject to the smallest MLS requirements, up to those predicted to be subject to the largest requirements. While larger MLS

prone to releasing metro data where many constituent counties are incomplete.

[19] The analysis dataset is not a panel dataset, because block borders do not stay constant between decades.

requirements are further away from the city centers on average, the differences in means are small compared to spatial diffusion of various MLS requirements throughout the urban area.

3 How might MLS requirements shape racial diversity?

Between the start and end years of our analysis — 1980 up to 2020 — metropolitan residential segregation has declined at a steady pace (Logan and Stults, 2021). Recent work (Massey and Tannen (2018), Bartik and Mast (2023)) documents how, over these years, the racial and ethnic diversity of those who move to the suburbs has also grown. we follow a calculation in Bartik and Mast (2023) and quantify the share of People of Color (POC) residents in our sample — defined as those who are not Non-Hispanic White individuals — living beyond historical urban cores. For each MSA we analyze, we define the historical core as their central city's 1960 borders. Using small area Census tabulations, Panel (a) of Figure 2 plots the share of POC residents outside the historical borders. Over our 120 MSAs, POC residents went from a "suburbanization rate" of 39% in 1980 to 74% by 2020. Drilling down to individual racial or ethnic groups, we similarly find that Black individuals' suburbanization in major MSAs has increased at a steady pace, from a 32% suburbanization rate in 1970 to 68% by 2020.

The year built statistics in Table 3 indicate that by 1980, the start of our analysis time frame, substantial shares of development had already been built following existing MLS requirements. About 70% of treated areas surrounding predicted border segments have homes whose median year built was before 1980. The effects we investigate thus come out of the tension between demand for suburban residential development and the patchwork design of MLS requirements that hinder supply.

Increases in suburbanization rates can be explained in part by the rising incomes of all racial and ethnic groups in the United States. Panel (b) of Figure 2 shows, in the latest data available, a persistent gap between mean incomes of non-Hispanic white individuals and people of color of between \$24,000–\$40,000 in 2023 dollars. Thus, even if everyone in the population has the same preferences for lot sizes, the "income channel" can generate differences in the racial composition of neighborhoods where higher requirements take effect.

Formally, consider an environment where a metropolitan area with N_W white residents and N_{POC} POC residents are divided into neighborhoods. Each neighborhood has a perfectly binding MLS requirement, so all homes in the neighborhood are restricted to be at a specific lot size. Different racial and ethnic groups have identical preferences for neighborhood characteristics. Assuming standard preferences where housing is a normal good, each neighborhood subject to

a particular MLS requirement attracts households earning an income y.[20] Then, the observed share of POC in the neighborhood can be written as:

$$\frac{N_{POC} f_{Y|POC}(y)}{N_{POC} f_{Y|POC}(y) + N_W f_{Y|W}(y)} = \left(1 + \frac{N_W}{N_{POC}} \cdot \frac{f_{Y|W}(y)}{f_{Y|POC}(y)}\right)^{-1},$$

where $f_{Y|g}$ is the income distribution conditional on a resident belonging to demographic group g. In equilibrium, a neighborhood with a higher binding minimum level of housing quality than another sees residents at a higher income y' move in. Because observed income disparities between POC and Non-Hispanic White residents appear at the right tail of income distributions, the likelihood ratio $LR(y) = \frac{f_{Y|W}(y)}{f_{Y|POC}(y)}$ increases and the equilibrium POC share decreases.

The income channel by which MLS restrictions might affect neighborhood racial and ethnic diversity, on its own, provides two testable hypotheses regarding the heterogeneous treatment effects of MLS requirements:

Hypothesis 1: *The effect of a MLS requirement on racial composition is increasing in how much less dense neighborhoods constrained by the requirement are, relative to surrounding developments' densities.*

Hypothesis 2: *If both the MLS restricted development and surrounding development are developed at sufficiently low densities, neither area will have a high POC share: preexisting income disparities cause equilibrium sorting away from both low-density areas. Thus, the effect of a MLS requirement on racial composition is moderated by the base density level of surrounding development.*

The income channel is unlikely to be the sole mechanism: potential buyers of different races or ethnicities may have different average preferences for the size of the lots they buy or rent, or for amenities that might be correlated with lots of particular sizes or restrictions. Different racial or ethnic groups might, on average, have different tastes for the lawns and gardens the lot size may permit or discourage, such as in Kulka (2019) and Song (2025). To the extent that previous lot sizes are correlated with current racial compositions of neighborhoods, larger lot sizes could also be more desirable to particular residents, due to homophily preferences for neighborhoods (Kuminoff, Smith and Timmins (2013), Aliprantis, Carroll and Young (2024)). Homophily preferences may both have a "buyer-side" effect as described up to now, as well as a "seller-side" effect: current neighborhood residents of a particular race or ethnicity might have low risk tolerance for the implications of neighborhood change, and might consider MLS

[20] A simple equilibrium concept that rationalizes this result in a metro area with many neighborhoods is an assignment model of housing with an increasing housing price gradient (Landvoigt, Piazzesi and Schneider, 2015)

requirements to limit that risk by making it more costly to redevelop the area.

Those two types of mechanisms also may occur in tandem: if Non-Hispanic White households both have greater incomes and a stronger preference for neighborhoods bound by MLS requirements than households of other races or ethnicities, that would exacerbate the effects of MLS requirements on the racial composition of neighborhoods. The overlapping nature of the mechanisms complicates finding the requirements that matter most for racial composition absent careful empirical work.

The mechanisms we discussed apply even without having neighborhoods organized under local governments, which differ in public goods provision[21] While we exclude differences in MLS requirements across municipal borders, due to confounder concerns specified in Section 2.2, we acknowledge that better identification through within-city comparisons trades off potentially undercounting MLS border segments involving the largest size differences, wherever those segments overlap with jurisdiction borders.[22]

4 Empirical Strategy and Identification

4.1 Identifying Policy-Relevant Effects

Suppose that a residential area i contains homes of lot sizes $\{\ell\}_i$ and we are interested in how a MLS $\underline{\ell}$ in the area affects residential composition Y_i. The estimand

$$\tau(\underline{\ell}, N) = \mathbb{E}\Big[Y_i \Big| \min\{\ell\}_i \geq \underline{\ell}\ \Big] - \mathbb{E}\Big[Y_i \Big| \operatorname{median}\{\ell\}_i = N\ \Big], \quad \underline{\ell} > N, \tag{2}$$

is then an average treatment effect on the treated corresponding to a counterfactual: how would Y_i differ if the lot size restricted area i is in was instead developped at smaller lot sizes of around N square feet? The discussion in Section 3 suggests that while this function should be nonnegative and monotonic in $\underline{\ell}$, multiple overlapping mechanisms introduce complexity and heterogeneity to the function. Characterizing the function, however, is crucial for the

[21] The income channel of minimum lot sizes is thus separate from the "benefit tax" view of zoning in theory dating back to Hamilton (1975), though the empirical literature following that paper assumes higher income individuals are also being willing to pay more for high-quality local public goods.

[22] Our approach contrasts with Rothwell and Massey (2009), who instrument density zoning restrictiveness with year of statehood, claiming that the between-state instrument satisfies the exclusion restriction. More broadly, our results are not directly comparable with between-MSA regressions that use metropolitan segregation as an outcome, like what was used in Rothwell and Massey (2009). However, metropolitan segregation measures are weighted averages of both segregation between jurisdictions in a metro, as well as within-jurisdiction segregation (Monarrez and Schönholzer, 2023). We plan to extend our estimates to simulate changes in both components of metropolitan segregation.

enforcement of fair housing: the effects inform outcomes had a particular zoning jurisdiction not adopted some, or all, of their MLS requirements.

In our empirical context, our definition of residential area is as granular as the Census block. Even with national data, the sample size needed to identify the entire treatment effect surface over $(\underline{\ell}, N)$ may be large. To make estimation and inference more tractable with our block-level aggregates, we report a set of point-identified effects across *impact clusters*. We set up a partition of the space $\{(\underline{\ell}, N)\} \subset \mathbb{R}^{2++}$ as $P_1, \ldots, P_K$. The *effect of regulations belonging to impact cluster* k is thus $\tau^k = \mathbb{E}_{P_k}\left[\tau(\underline{\ell}, N)\right]$.

Effects that vary at the impact cluster level allow for more heterogeneity than the existing economics literature estimating an average effect over the pooled sample (Severen and Plantinga (2018), Schönholzer (2024)). We allow less heterogeneity, however, than an approach like Keele and Titiunik (2015), which estimates a separate treatment effect at different boundary segments without parametric restrictions. However, just as individual estimates over the boundary can be equally weighted to produce an averaged impact cluster effect, our preferred estimates are weighing underlying heterogeneous effects by how frequently they show up among all within-cluster built contexts.

Our estimates are most informative about "first-order" magnitudes of policy hypotheticals involving marginal changes in zoning, applied to one jurisdiction at a time. Counterfactuals involving rezoning of large amounts of land, or reforms that apply to many jurisdictions, also alter so many neighborhood characteristics in general equilibrium that they could independently affect neighborhood choice. [23]

4.2 Identifying Zoning Impact Clusters

Identification of the treatment effects in Equation 2 requires two terms. The first term is observed for blocks on the side of a predicted border segment that is restricted by a particular MLS requirement. The second counterfactual term is how those MLS restricted blocks would look had they been developed at a higher density. Our empirical strategies are all based on estimating the counterfactual using observational data: we employ blocks in the comparison area on the other side of the border segment as a control group, and we use our block-level

[23] If a hypothetical zoning reform introduces many new households with different races and ethnicities than incumbent residents, sufficient strong homophily preferences would make the neighborhood attractive to more POC residents. Additionally, those preferences would make incumbent residents more likely to move elsewhere. The combination of two channels could mean our estimates are lower bounds for the effect a metro-wide reform limiting MLS requirements would have on neighborhood racial composition.

demographic data and within-city spatial features to implement regression adjustment.[24] We classify MLS requirements over different neighborhoods or jurisdictions to the same cluster if the regulatory minimums, as well as the degree to which they change density in their respective neighborhoods, cause similar changes in racial composition relative to surrounding residential development.

In an ideal experiment, we would estimate effects using only clusters of zoning district borders that were randomly assigned in space, prior to any development on land zoned one way or the other. We could then derive the *unconfounded* average treatment effect of a MLS T_b on development constrained by that requirement, relative to nearby developments at an unconstrained density: $(Y_b(\underline{\ell}), Y_b(N)) \perp\!\!\!\perp T_b$. In practice, some parts of zoning districts' borders are set up endogenously. The borders may follow city boundaries diverging in public good provision prior to zoning, or they may have been drawn around preexisting development or land parcels that would have attracted different buyers regardless of zoning.

In the first stage, where we discover impact clusters, we apply strategies to control for endogenous border definitions. While we elaborate on our research designs in the next section (4.3), one of our designs relies on unconfoundedness conditional on observables. Our analysis dataset is already limited to a sample of predicted borders that follow straight lines and do not overlap with any local government boundaries, as was explained in Section 2.2. Straight borders are less likely to reflect endogenous drawing of borders around preexisting development. To the extent that we know when the straight predicted borders overlay within city spatial features described in Section 2.4, we explicitly control for those conditions with flexible regression adjustment.

To identify impact clusters while also controlling for endogeneity, we employ the generalized random forest (GRF) algorithm developed by Athey, Tibshirani and Wager (2019). The algorithm, as implemented in the `grf` package in R, simultanenously estimates treatment effects, an additive function of historical confounders on Y_i and the propensity score function for the binary treatment — in this case, whether a block is in the restricted or comparison area. Through the GRF algorithm, we derive individualized treatment effect estimates for each pairing of a restricted area and a comparison area, as a function of regulation characteristics and covariates.

We deploy the GRF algorithm using variation between areas, achieved by first aggregating block-level data to the restricted or comparison area it overlaps. These blocks must also have centroid lengths of no more than 400 meters from the predicted border segment they're

[24]The comparison areas may themselves be subject to MLS requirements. But, by construction (as explained in Section 2.3), they are less restrictive than the MLS requirement imposed on the restricted area.

matched to. Our outcome of interest is the share of all people of color across the treated and comparison areas, while we estimate over a pooled sample of blocks observed after 2000. We include granular fixed effects and all the historical boundary overlap covariates we tested in Section 2.4, which are used to fit the additive function of confounders.[25] We also include 1980 shares of rental units, calculated separately for comparison and restricted blocks, in case pre-existing differences in rental rates influenced both racial composition and the design of zoning boundaries.

From these individualized estimates $\hat{\tau}(\underline{\ell}, N, X_i)$, we follow the logic of Semenova and Chernozhukov (2021) and we output model-generated treatment effects from the second-order polynomial model:[26]

$$\tilde{\tau} = \tilde{b}_0 + \tilde{b}_1 \log(\underline{\ell}) + \tilde{b}_2 \log(\underline{\ell} - N) + [\log(\underline{\ell} - N) + \log(\underline{\ell})]^2 \tilde{\mathbf{b}}. \tag{3}$$

Panel (a) of Figure 3 plots the surface of model-generated treatment effects on block-level POC shares, over the two axes over which we allow treatment effect heterogeneity: the MLS applied to the lot size restricted area in question, $\underline{\ell}$, as well as the density differences with the comparison area by subtracting from $\underline{\ell}$ the median lot size in the comparison area, N.[27] The log transform minimizes the influence of the largest minimum lot sizes on predicted effects.[28] The output curve shows, unsurprisingly, how the most negative effects on the POC resident share are found in areas where $\underline{\ell} - N$ is the largest, or where N is closest to zero. Surprisingly, the effects are not simply monotonic in the density change variable: when the MLS requirement is both at its smallest and largest values, model-generated effects are close to zero.

Panel (b) of Figure 3 shows how we use the model-generated effects information to define our impact clusters. We first identify the range of model-generated treatment effects that are negative, $(0, \tilde{T}]$, from our model of $\tilde{\tau}$. Then, our definition of a *high-impact cluster* includes all

[25] These overlap variables are calculated at the predicted border segment level, so are the same for the treatment area and its paired comparison area.

[26] If we restricted this model to have only the first order term $\tilde{b}_1 \log(\underline{\ell})$, the model reduces to a specification like in Song (2025). To replicate that paper's specification, we would control for confounders through a linear specification and not the GRF algorithm. Then, we would impute for comparison areas the MLS requirement applied to them instead of realized density, after which we would run the regression over all treated and comparison areas.

[27] Appendix Figure B.2 shows the density of observed border segments over these two characteristics. Minimum lot sizes between 6,000 square feet to 22,000 square feet (1/2 acre) are common, and it is also common for the change in density between the treatment and comparison areas to be small, and at the other end of that continuum, for the comparison group to have lot sizes at least 50% smaller than the treatment area.

[28] The log transformation is one of several ways to reduce the influence of outliers when the underlying distribution of treatment intensity is skewed or fat-tailed. Similarities exist between our approach and the estimator in Pons (2024), which outputs a two-dimensional treatment curve defined over quantiles of treatment intensity instead of levels.

lot size discontinuities whose model-generated treatment effect is more negative than $\tilde{T}/2$.

To further characterize MLS requirements in the high-impact cluster, we calculate moments for areas treated on one side of border segments in the high-impact cluster, aggregated over blocks as explained in Section 2.5. The median and average MLS requirement in effect in the treated areas is 10,500 and 11,700 square feet, respectively. Neither moment is very far from what was estimated using bunching data in Cui (2024), where the median and mean requirements were at 7,500 and 14,000 square feet, respectively.[29] In practice, nearly all requirements in this cluster are between 6,000 square feet and half an acre (22,000 square feet): requirements less than 6,000 square feet and those larger than 22,000 square feet account for just 1% and 2% of treated blocks, respectively.

In Panel (b), we also see three other clusters into which border segments could fall. Where model-generated effects are smaller than a 1 percentage point magnitude change in block-level POC shares, we classify the predicted border segment to belong to a *low-impact cluster*. The low-impact cluster includes all MLS requirements of less than half an acre that also are no more than 1,500 square feet larger than median lot sizes of surrounding development. Lot size discontinuity border segments in the *moderate-impact cluster* have model-generated effects between the high-impact and low-impact thresholds. A sparse set of segments involves very large predicted MLS requirements, where crossing over to the less dense treated area is predicted to cause increases in POC shares. We assign them to a fourth *large lot outlier cluster*. Further summary statistics, produced for all segments, can be found in Appendix Table A.3.

4.3 Identification Under Two Research Designs

Our national data allow us to pool outcome changes for blocks near border segments, with a well-powered sample to infer heterogeneous effects. While we limit the sample to those blocks with centroids within 400 meters from border segments, it is possible for that threshold to still include developments built at different points in time, or disconnected from other neighborhoods for the border segments that overlap with other within-city boundaries.[30] The blocks further away from the discontinuity, for example, might have had different rates of unit availability and tenure by 1980, which on its own can attract different types of incoming residents to jurisdictions (Resseger (2022), Furth and Webster (2022)) To further address concerns that

[29]The median MLS requirement in the high-impact cluster is even closer to other estimates of the median national MLS requirement in the literature: Song (2025) reports it to be 10,000 square feet using her methodology, while Bartik, Gupta and Milo (2024) reports it to be 11,000 square feet from analyzing municipal ordinances.

[30]The checks we have done on border overlap with spatial features does not rule out a spatial feature running in parallel with the border segment, in which case there are barriers between blocks belonging to only one side of the border.

our estimates could be driven by factors other than the restrictiveness of the MLS requirements, we flexibly adjust for sample composition issues and reverse causality issues from endogenous borders through two separate research designs.

Causal Forest design. The first design reuses the causal forest methods we ran in Section 4.2. Where our predicted straight border segments overlap with the railroad tracks and other within-city spatial boundaries discussed in Section 2.4, we control for those spatial boundaries. We further control for built environment features, like the median year built of the housing stock surrounding the predicted straight border segments. More formally, the causal forest identifies effects in an impact cluster due to conditional independence:

Assumption CI (conditional independence) For outcomes taken over analyzed border segments with characteristics $(\underline{\ell}, N) \in P_k$, belonging to an impact cluster P_k, $(Y_b(\underline{\ell}), Y_b(N)) \perp\!\!\!\perp T_b | X_b$.

Border Discontinuity design. The second research design uses outcome variation between lots that are closer and those further from the border segments, to model unobservable trends correlated with distance to the border. By fitting separate regression functions in the treated and the comparison areas, the design estimates treatment effects of MLS requirements not influenced by neighborhood confounders.

Formally, we apply a *border discontinuity design* that identifies treatment effects if potential outcomes are *continuous at the border* determining where the requirement applies:

Assumption CC (conditional expectation function continuity) For all points **s** on border segments classified within an impact cluster P_k, and for **s'** of blocks b in space:

$$\lim_{\mathbf{s}' \to \mathbf{s}} \mathbb{E}\Big[Y_b(\mathbf{s}') \Big| \text{median}\{\ell | T_b = 0\}_{\mathbf{s}'} < \underline{\ell} \Big] = \mathbb{E}\Big[Y_b(\mathbf{s}) \Big]$$
$$\lim_{\mathbf{s}' \to \mathbf{s}} \mathbb{E}\Big[Y_b(\mathbf{s}') \Big| \min\{\ell | T_b = 1\}_{\mathbf{s}'} \geq \underline{\ell} \Big] = \mathbb{E}\Big[Y_b(\mathbf{s}) \Big].$$

We note that under this assumption, the border discontinuity design estimates a different effect from the average treatment on the treated (ATT) presented in Section 4.1: the design identifies the effect of MLS requirement changes on the marginal lot at the border segment, instead of for an entire MLS restricted neighborhood. While the border discontinuity strategy offers a more flexible way to correct for neighborhood confounders on two sides of the MLS border, it is by design a more conservative estimate of the policy-relevant effect: some lots in MLS restricted neighborhoods in our data are not developed at the size mandated by the

requirement.[31]

Estimation and inference under both designs can be modeled as a linear regression, with differences in how the control function is set up in the regression equation and in how sample restrictions are applied.

$$Y_{bt} = \alpha_{j(b)t} + \beta^t \mathbf{1}[Dist_b \geq 0] + g(\mathbf{X}_b) + \varepsilon_{bt} \quad (CF), \tag{4}$$

$$Y_{bt} = \alpha_{j(b)t} + \beta^t \mathbf{1}[Dist_b \geq 0] + \eta^t_{-} Dist_b + \eta^t_{+} Dist_b \cdot \mathbf{1}[Dist_b \geq 0] + \varepsilon_{bt} \quad (BD), \tag{5}$$

where $g(\mathbf{X}_b)$ is a flexible additive function estimated using the GRF algorithm; $Dist_b$ is the nearest distance from the block centroid to the border segment, in meters; and in the baseline we apply fixed effects $\alpha_{j(b)t}$ at either the segment or jurisdiction level. Importantly, the causal forest model is estimated over blocks aggregated into treatment and comparison areas s. The comparable block-level regression in the border discontinuity model is also weighed by block population in any given year. Adding fixed effects further facilitates pooling regulatory variation across cities nationwide, because fixed effects net out additional characteristics that could explain across-city divergence. The fixed effects limit bias, and also increase the precision of estimates.

In our empirical context, balance in predetermined covariates between the treated and comparison areas around the boundary informs the validity of both research designs.[32] In Appendix Table A.4, we take simple differences in means over our main subsample of interest: predicted MLS requirements in the high-impact cluster. By filtering to blocks up to 400 meters away from the discontinuity, these samples also reflect differences in neighborhood composition for areas far away from the border. Some characteristics are better balanced than others. For example, single-family homes in treated blocks were on average only built 5 years later than those in comparison blocks. Imbalance issues are more apparent with variables like unit tenure, where we see a 10 percentage point difference between the treated and comparison areas.[33] Due in part to this finding, we added 1980 rental unit shares to the covariates included in the causal forest model.

[31] The lots may be bigger, or even smaller if they are non-conforming uses or were granted a variance or other discretionary approval.

[32] We conduct two other tests separate from covariate balance and present them in Appendix Figure B.3. We verify that overlap of propensity scores appears satisfied in Panel (a). For the Cattaneo, Jansson and Ma (2018) test of manipulation of block density around the discontinuity, we cannot reject the null that a break appears in the distance distribution around border segments. We discuss in Section 5.6 particularities with the distance distribution when our distance measure is the block centroid distance.

[33] Analogous difference in means tables for the other two principal clusters are shown in Appendix Table A.5.

Table 4 estimates covariate discontinuities at the border segment, using the local linear model in Equation 5 with the same bandwidth of 400 meters. Estimates of discontinuity magnitudes at the border are listed in the first four rows: we cannot detect breaks in 1980 demographics. We find significant point estimates for age distribution outcomes that are small in magnitude, close to 1 percentage point. We find more sizable discontinuity estimates for the block rental share. Because the effect remains significant at a 90% level, we take care to account for the rental channel in presenting our results.

5 Effects on Racial Composition

5.1 Lot Size Effects Vary By Regulatory Context

We run the border discontinuity design, as specified in Section 4.3 for blocks near every border segment in our national sample, estimating the effect of the predicted MLS requirements on block-level shares of a specific racial group m, denoted $Share^m_{bt}$, or of the share of all residents whose race and ethnicity is other than non-Hispanic and White — the people of color share — denoted as $Share_{bt}$.

Table 5 presents separate effects over the pooled segment sample for three years: two recent Census years (2020 and 2010) and the earliest year in our sample (1980). Even with the large sample size, the results across some specifications do not reject the null hypothesis at 90% confidence. The results of other specifications produce the counter-intuitive finding that larger lot size requirements, which result in lower density, increase the block-level share of people of color in the areas subject to those requirements.

The latter result is hard to explain as a matter of theory, because of income disparities between POC residents and non-Hispanic White residents, as discussed in Section 3. The null results over multiple decades also appear inconsistent with other evidence that large MLS requirements and other density restrictions were designed to exclude lower-income residents (Inman and Rubinfeld (1979); Trounstine (2018)), and that such restrictions are associated with fewer Black and Hispanic residents in the jurisdiction (Pendall (2000); Quigley, Raphael and Rosenthal (2004)).[34]

But average results for the pooled sample obfuscate the wide variety of contexts in which MLS requirements are applied. The pooling might hide effect differences between, for example,

[34]We have run the pooled specification over all Census years, as well as breaking down the POC resident share separately by Black, Hispanic and Asian residents. We do not find any negative effect of our predicted MLS requirements with greater magnitude than 0.5 percentage points. Appendix Table A.6 also presents pooled regressions using the causal forest model, which more consistently finds negative effects.

a treated area with a 7,000 square foot predicted MLS requirement and a comparison area with lot sizes around 5,000 square feet versus a predicted requirement of 22,000 square feet near untreated areas where lot sizes average to 16,000 square feet, even though the ratio of the prevailing lot sizes is about the same in those two scenarios. Similarly, pooling might hide effect differences where the MLS requirements are similar but there is substantial variation in how they differ with their comparison areas. We have accounted for both those dimensions of heterogeneity when identifying the impact clusters in Section 4.2.

Figure 4 reports descriptive statistics, for blocks belonging to each of the four impact clusters, on whether the share of people of color residents in the treated areas lagged behind the increase in nearby comparison blocks. Over all the MLS requirements classified within each impact cluster, we sum counts of all POC residents, divided over total population, separately for blocks on the treated side of the border and those in the comparison areas not bound by the restriction.

In all four clusters, blocks around the border segment see steady increases in the POC resident share, similar to the aggregate trends in POC suburbanization shown in Figure 2 Panel (a). In 1980, the high-impact and moderate-impact clusters had shares of POC in both the treated and comparison areas that were around the national average of 20%. In the high-impact cluster, however, that we see visible divergence in overall POC shares in each decade after 1990. By 2020, the comparison blocks next to the predicted MLS requirements that are high-impact have shares of POC residents that are 4 percentage points higher than the blocks subject to those requirements.[35]

5.2 High-impact MLS Requirements Cause Shifts In Racial Composition

In each panel of Figure 5, we show the same binned scatterplot for the POC share outcome as of 2020, residualized once using fixed effects at the border segment level, with bins aggregated from block data in the data-driven way recommended by Cattaneo et al. (2024). Adjusting for county-level endogeneity gives us a sample with a baseline POC share of 48%, close to the 46% POC share found for MSA residents in the Census's 2021 American Community Survey Ruggles et al. (2025). Once granular fixed effects are applied, there is no obvious divergence in POC share for treated areas close to the boundary compared to treated areas further away, denoted as having a high positive value in the running variable. The same holds true for outcome trends

[35] In the low-impact cluster, the POC resident share is slightly higher for treated area blocks than the blocks to which we compare them. This result may be due to differences in built form or housing age across the border segment that we control formally in our main specifications.

on the comparison side (denoted with negative sign).[36]

The solid lines highlighted in Panel (a) represent the equal-weighted outcome levels for each side, whose difference reflects the effect estimate in a regression with fixed effects but without any covariate adjustment term $g(X)$. The dashed lines reflect reweighted outcome levels following covariate adjustment based on the causal forest algorithm, showing a smaller effect estimate due to a larger downward adjustment on the comparison side than the treated side.

The leftmost two columns of Table 6 report how predicted MLS requirements in the high-impact cluster affect the racial composition of the treated areas, following a causal forest design. Column (2) shows our preferred specification using border segment-level fixed effects, which produces an estimated decline of -2.8 percentage points in the POC share for the average treated area, compared to the average nearby area without the requirements. Employing all the treated and comparison areas around the border segments, our effects are statistically significant at the 99% level. The effect is around 6 percent of the baseline POC share of 46%. Even disregarding statistical insignificance concerns, effect magnitudes based on the pooled data would be less than 1% of the baseline share.

We further assess the effect for the high-impact cluster by estimating separate effects for different categories of race and ethnicity. The second row of Table 6 estimates, for the high-impact cluster and based on the causal forest design, a decline of 3.1% in the share of Black residents in the average block governed by a predicted MLS restriction. That effect is a substantial 22% of the fixed effects adjusted baseline Black resident share of 14%.

In Panel (b) of Figure 5, we illustrate the border discontinuity design: we plot the piecewise linear regression functions based on a population-weighted regression, corresponding to the design's control functions, against the same set of racial composition data. Because granular fixed effects worked well to remove divergence in outcomes based on distance, the treatment effect from the BD design in this case is close to that from our causal forest approach.[37]

In the rightmost two columns of Table 6, we show results from our preferred specification analogous to what is shown in Panel (b), fixing the border discontinuity sample to include blocks whose centroid distances are no more than 400 meters from the border segment.[38]

[36] We process these binned plots over other outcomes, at different time periods and with less granular fixed effects. Less granular fixed effects, and outcomes like the Black resident share, exhibit more divergence based on the running variable.

[37] Panel (a) of Appendix Figure B.4 conducts the same visualization for the POC share in 1980, or visualizing the pre-period balance test we conducted in Section 4.3. Panel (b) does the same for a statistically significant confounder discussed in that Section: share of rental units at the Census block level.

[38] This is the same criterion used for the causal forest design, and is 100 meters shorter than the diameter of the interior tile used in our border detection procedure.

While the specification based on jurisdiction fixed effects produce a noisy estimate, results using our preferred specification in Column (2) are not far off from the causal forest estimate.

The point estimate magnitudes based on the border discontinuity design are smaller than the ones for the causal forest (which, in turn, were smaller than the 4% difference in raw means between restricted and comparison areas, suggested by Figure 4). That said, the estimate from our preferred specification on the overall POC residential share (-0.019), as well as on the Black residents' share of the neighborhood (-0.026), remain significant at a 95 and 99% level of confidence, respectively. Furthermore, for any conventional threshold of statistical significance, we cannot reject the null hypothesis that the effects identified by the two designs are equal. When discussing the remaining results, we prioritize using the border discontinuity estimates as a conservative estimate of MLS requirements' effects.

When we estimate results for Hispanic residents and Asian residents, we cannot reject the null hypothesis of zero effects for either group. The causal effects estimated for the overall POC share accordingly appear to be mainly driven by changes in Black residents' location decisions. Furthermore, they are suggestive about the mechanisms that underlie the racial diversity results.

In the most recent analysis of its kind, Cilluffo and Kochhar (2018) shows the income distributions of Asian Americans closely resemble that of non-Hispanic White Americans in 1970 and in 2016. Without strong income disparities between the racial groups at any quantile of the income distribution, the income mechanisms detailed in Section 3 have little effect on how MLS requirements cause Asian resident composition. We lack the evidence needed to determine whether mechanisms of the second kind apply to Asian residents: the evidence is unclear about whether Asian households have a stronger preference for larger lot sizes than Non-Hispanic White households, or whether they are constrained in their search process in housing markets [39]

Cilluffo and Kochhar (2018) also show few strong income disparities over time between Black and Hispanic adults; both groups earn less than Non-Hispanic White adults. *A priori*, that fact would suggest that MLS requirements would have a negative effect on the share of Hispanic residents in neighborhoods governed by the requirements, contrary to our findings. Because point estimates are negative for Hispanic resident shares in the causal forest design, it is possible that we lack power over our sample to detect smaller effect sizes for Hispanic

[39] In recent correspondence studies testing discrimination in rental markets, there is conflicting evidence on whether Asian Americans are steered into certain neighborhoods (Turner and Ross (2003), Christensen and Timmins (2022)).

residents. [40] Differences in effect magnitudes between the groups could possibly be explained by different preferences for neighborhood characteristics that are associated with high-impact cluster MLS requirements — the non-income disparity mechanisms discussed in Section 3.

5.3 Requirements in Other Clusters Have Modest to Small Effects

In the top half of Table 7, we estimate the border discontinuity model over the two other clusters: the moderate-impact and low-impact clusters, which each account for about 40% of our sample of border segments. Reproducing the same model used to estimate effects over the high-impact cluster, we confirm that in 2020, neighborhood composition effects are small and insignificant for the other clusters. No point estimate in the clusters is larger than 1 percentage point in magnitude.[41] Apart from the effects on the share of Black residents in the moderate-impact cluster, other estimates are not statistically significant at the 95% level.

The less meaningful results could be explained in part by how, in the clusters other than the high-impact cluster, the average border segment involves a smaller shift in density between the two sides of the border. In other words, the MLS requirement under which the relevant district is regulated does not substantially deviate from the density in comparison blocks without the same requirement. We visualize the density changes in the data by segments in Figure 6, where we present binned scatterplots overlaid with a border discontinuity model as in Panel (b) of Figure 5. Our measure of block-level density in this case is based on the Census estimates, where we divide all housing units, occupied or not, in the block with the block area in the TIGER shapefiles.

Panel (a) of Figure 6 visualizes the average density change for MLS requirements classified in the high-impact cluster. On the sides not bound by the MLS requirements, we estimate an average comparison density of 1482 units per square kilometer. On the side bound by the MLS requirements, density drops by an average of 558 units/km2 to 924 units/km2. In Panel (b), the same specification over the moderate-impact cluster results in a drop from 1135 to 1030 units/km2, a drop of 105 units/km2. Converting to dwelling/acre (DUPAC) terms, the differences in density between treated and comparison areas equate to an average fall of 0.6 DUPAC in the moderate-impact cluster (contrasted to 2.8 DUPAC in the high-impact cluster). For the low-impact cluster, the average density decline is 0.3 DUPAC.[42]

[40] To corroborate this, estimates we ran on the share of residents who are neither Non-Hispanic White nor Asian are different from zero at a 95% level of statistical significance; in line with effect estimates on the POC share; and smaller than the effect estimates on the Black share.

[41] Appendix Table A.7 replicates Table 6 in full for requirements in the moderate-impact cluster. We confirm no effects for Hispanic and Asian residents are statistically significant.

[42] Alternatively, the average high-impact cluster requirement induces a proportional decline of 38% in residential

Particularly for the latter cluster, true effect sizes may be so small that we lack the power to detect them with our national sample of border segments. Compared to a DUPAC change that is 4.7 times greater for the high-impact cluster than for the moderate-impact cluster, the ratio of effects on the Black resident share between the two clusters is 3.3 (2.6/0.8). The difference in the ratios could indicate an interesting nonlinearity in effects moving from small to large DUPAC changes.

5.4 Dynamics of Effects Across Segments

Analyzing one year of the most recent data, as the estimates reported in Table 6 do for 2020, shows whether MLS requirements affect the current racial composition of neighborhoods. Table 6 is uninformative, however, about whether those requirements had different effects on racial composition in the past relative to today. MLS requirements may have been more impactful previously if, for example, people were more reluctant to be among the first of their race or ethnicity to move into a housing market's most expensive neighborhoods (which neighborhoods with larger lots were more likely to be). However, POC households may have been averse to move into *any* neighborhood made up primarily of white residents. Then, suburbs with high shares of white residents have a MLS requirement or not matters less on the margin, so MLS requirement effects may be larger into the present. The same implications hold if any POC household moving out of dense areas faced more discrimination in applying or securing credit for any kind of suburban home in the past than today.[43]

Figure 7 provides estimates for high-impact cluster effects over time, with block-level POC share as the outcome. Requirements in this cluster caused larger neighborhood composition effects between 1980 and 2010 than in the most recent decade. In 2010, the MLS requirement effect was as much as 5.5 percentage points for the POC share, or about 15% of a 36% baseline share. In 2020, however, the estimates diminished in magnitude.

In Appendix Figures B.5 and B.6, we break down the analysis by the shares of Black and Hispanic residents. The time trends indicate that the between-group effect differences seen in Table 6 have appeared repeatedly throughout time. Our conservative estimates from the border discontinuity design cannot find a composition effect for the Hispanic American share throughout 1980 to 2020, while the dynamics of the Black resident share effects show magni-

lot sizes, relative to surrounding development. For the average moderate-impact cluster, the decline is 9% in residential lot sizes.

[43] Such a conclusion is supported by both the redlining literature over the course of the 20th Century (Aaronson, Hartley and Mazumder, 2021) as well as prior audit studies conducted by the US Department of Housing and Urban Development (Oh and Yinger, 2015)

tude increases without clear reversion as of 2020.[44]

Finally, we check whether there were greater neighborhood composition effects in previous decades for the two lower impact clusters. Appendix Figures B.7 and B.8 run the analysis for the POC share over, respectively, the low-impact and minimal impact segments. Effects over the low-impact cluster have persistently been small or statistically insignificant across specifications. In the moderate-impact cluster, we detect negative effects of about 1 percentage point as early as in 1980. Negative effects persisted up to 2010, but reverted to 0 by 2020. As Table 7 shows, we can rule out the null hypothesis that in the moderate-impact segment, minimum lot sizes do not affect the 2020 Black American share.

Overall, the results suggest that MLS requirements cause racial composition changes, especially after 2000 and excluding low-impact requirements that generate minimal declines compared to adjacent density. In Appendix Table A.8, we compare the magnitude of the MLS effects in the high-impact and moderate-impact clusters to four other papers in the literature on the persistence of legal constraints on neighborhood diversity. We consider specifications where variation around border discontinuities are used and where the outcome is the share of Black residents at a sub-Census tract geography.

For Black residents, effect estimates for MLS requirements in the high-impact cluster are in line with the magnitude of Black share declines across the average local government border, estimated in Monarrez and Schönholzer (2023). While that paper emphasizes the role of local public good provision in producing neighborhood racial sorting between jurisdictions, our estimate magnitudes suggest the presence of high-impact lot size zoning within jurisdictions could affect neighborhood racial composition in comparably important ways.

In contrast, MLS effect magnitudes over the moderate-impact cluster, in both 2010 and in prior decades, more resemble the results of historical interventions in suburban housing markets, whether it is unequal credit access (Aaronson, Hartley and Mazumder, 2021) or legal covenants banning sales to certain races (Sood and Ehrman-Solberg, 2024).[45] The effects of the historical interventions, as of 2010, are not noticeably more impactful than effects were in 1980, which is also the case with moderate-impact cluster effects in Appendix Figure B.7.

[44]Between 2010 and 2020, the Census Bureau also revised the way in which respondents report combinations of race and ethnicity and introduced a new disclosure avoidance algorithm that adjusts block-level counts for privacy reasons Asquith et al. (2022) The 2020 Census vintage more explicitly adds "noise injection" to each block, shifting both the denominator of the outcome variable (population) and the numerator (race/ethnicity counts). The magnitude of reversion for the POC share specifically could be due to attenuation bias, coming from a form of intentional measurement error.

[45]Sood and Ehrman-Solberg (2024) remarks as much that on average, development in Hennepin County, MN with a racial covenant did not have much larger lot sizes compared to surrounding development.

5.5 From causal effects to policy guidance

To make our estimates more helpful for policymakers and the courts, we offer a back-of-the-envelope calculation to answer the question: "Will the number of POC residents increase, and by how much, if the neighborhoods to which they might move lacked the lot size requirements those areas currently have?" Our methods and results let us consider counterfactuals where, without the requirements not in place, developers would have built up to the density of surrounding blocks not subject to the same requirement.[46] We then calculate the level of growth in the share and number of POC residents we would expect to see on average in the counterfactual MLS restricted areas. We perform this calculation separately for requirements across the three principal impact clusters, multiplying our estimates on changes in POC resident shares s^{POC} with those on changes in resident densities $n = N/sqmi$.

We offer these results with the same caveats noted in Section 4.1: other general equilibrium effects could alter market equilibrium densities throughout the metropolitan area if certain MLS requirements were prohibited everywhere in the MSA. Our guidance better corresponds to the result of a marginal relaxation of a single jurisdiction's MLS requirement in the districts where it is applied, such as over a few blocks or around one street, holding other jurisdictions' strategies constant. We also prioritize using our conservative estimates of effects from the border discontinuity designs: if the absence of MLS requirements caused development built on lots even larger than the MLS required to densify, our estimate would be lower than the true effects.

The bottom half of Table 7 reports the policy effects in units of residents per square mile across the three clusters. Furthermore, it offers a decomposition of the policy effect, based on the following observation:

$$\begin{aligned}\Delta\frac{POC}{sqmi} =& n_2 \times s_2^{POC} - n_1 \times s_1^{POC} \\ =& n_2 \times s_2^{POC} + n_2 \times s_1^{POC} - n_2 \times s_1^{POC} - n_1 \times s_1^{POC} \\ =& s_1^{POC}(n_2 - n_1) + n_2(s_2^{POC} - s_1^{POC})\end{aligned}$$

The first decomposition component is a *densification* term: if observed shares of POC residents on the MLS restricted blocks were unchanged, how many more residents of color would reside there in the counterfactual just because the same amount of land is more densely developed?

[46] In other words, we assume the realized density of surrounding blocks reflect a counterfactual equilibrium outcome without regulation. It is possible that with large-scale reforms, all the surrounding neighborhoods could have been developed for bigger lots, to cater to a different market segment. It is also possible the neighborhoods could have been developed at higher densities.

The second term converts the causal effects on *neighborhood composition* we estimated into level terms, based on the counterfactual neighborhood density. An alternative interpretation is that the densification term alone is a plausible lower bound for the counterfactual growth in POC residents: it is the predicted growth assuming that densification does not increase the area's attractiveness to POC residents considering whether to move to the neighborhood, due to either the income disparity or preference heterogeneity channels highlighted in Section 3. The combination of the densification and neighborhood composition terms reflect our preferred estimate of the full counterfactual growth effect.[47]

For both all POC residents and Black residents, Table 7 shows that eliminating the MLS requirements in the high-impact cluster would have meaningfully altered the diversity and racial integration of whichever jurisdictions that did in fact adopt those MLS requirements.[48] Generally, these requirements reflect large differences between the lot sizes in the restricted area and those in the comparison areas. If a jurisdiction had not applied the MLS requirement in the high impact cluster that is in place today on a square mile of land, development at the density of the surrounding comparison areas would have led to 1,480 more POC residents on that square mile, of which over 600 would have been Black residents.

The densification term by itself explains a majority of total predicted policy effects. Because even the areas subject to relatively large MLS requirements already have meaningful numbers of POC and Black residents, even our most conservative estimates of how MLS impacts racial diversity imply notable policy effects — a rise of 610 Black residents per square mile. Effects vary by racial or ethnic group, and less of the effect for Black household growth than for POC households is attributable to the densification term alone.

For the 41% of moderate-impact MLS requirements, we find modest effects on growth in POC resident levels on the order of 100-200 additional residents per square mile. We do not evaluate requirements classified under the low-impact cluster, as neither effects on neighborhood composition or on density in that cluster are statistically significant. Further work to identify specific MLS requirements imposed in the contexts that characterize the high-impact cluster, in addition to some in the moderate-impact cluster, could allow targeted reforms to

[47] The magnitude of our causal effects are also lower than the 4% raw difference in shares, noted in Figure 4. Our preferred estimates are then bounded by the densification effect as well as by the estimate arising from predicting counterfactual shares using the average composition shares in comparison blocks.

[48] One data issue involves how the density change effects are estimated over a definition of land area that includes undeveloppable land, like street frontage, that comes with residential development. The *POC/sqmi* statistics we report are in units of all land, developable or not. However, when we convert lot sizes in square feet to DUPAC, we implicitly *exclude* undevelopable land (i.e. using net, not gross, DUPAC). To convert gross DUPAC to net DUPAC, we assume for simplicity that 19% of the land is used for street space, based off of recent estimates for U.S. suburbs in Guerra, Duranton and Ma (2024).

those areas to encourage greater diversity in the future (and perhaps to redress the persistent effects of the past policies).

5.6 Further Robustness Results

Robustness over different housing vintages. One alternative explanation for the effects we estimate is that those effects are biased because the homes being compared around the border segment were built at different times. If homes in either the treated or comparison areas were built earlier than the other, the older homes might be more likely to house more diverse residents due to filtering effects (Rosenthal, 2014). Contrary to that explanation, however, our results across the three impact clusters remain robust to differences in housing age.

Table A.9 reruns effects in Table 6 but excludes discontinuities where one of the two surrounding areas has preexisting development: the areas in which median year built of the area's development is at least 5 years older than the inferred MLS adoption year for the jurisdiction, based on Cui (2024).

We also stratify each of the three principal impact clusters in Section 4.2 based on median year built. We define a *contemporary housing sample* as areas where both sides of the lot size border have homes with a median year built after 1980. The remaining tiles are referred to as the *postwar sample*. As Cui (2024) finds that most jurisdictions began MLS adoption by the 1970s, blocks built up after 1980 are unlikely to be areas already built up when the jurisdictions first considered these regulations.

We reestimate the border discontinuity specification from Equation 5 over six (2×3) samples stratified by impact cluster and housing stock age. Appendix Figure B.9 plots these estimates and 95% confidence intervals for our primary outcomes of POC share and Black resident share. We present time-varying estimates like we did in Figure 7 starting from 1990, the first meaningful post-period for the contemporary housing sample.

Across all the panels, the effects estimated on just the contemporary sample remain negative across many segments. The effects for Black resident shares are around a decline of 2 percentage points, remaining statistically significant at the 95% or 90% levels in recent times, for both the high-impact and moderate-impact clusters. We also cannot conclude, in any panel, that effects on the contemporary sample are statistically different at the 95% level from the postwar sample estimates.[49]

[49] To the extent that point estimates in the contemporary sample can be more negative than those on the postwar segment, it is possible that new builds in lot size restricted areas are also more likely to be bundled with more onerous homeowner association conditions that affect buyer composition. (Clarke and Freedman, 2019)

Robustness of border discontinuity design to bandwidth choices. In both our causal forest and border discontinuity design, we impose a bandwidth that drops blocks with centroid distance of more than 400 meters from their closest border segment. Increasing the bandwidth, while growing the sample, may include divergent blocks changing for non-zoning reasons. Conversely, decreasing the bandwidth may hinder the statistical power of the design. As displayed in Appendix Figure B.3, our definition of centroid distance for block observations, while also dropping blocks that overlap across the detected discontinuities, lead to underrepresentation of blocks with small distances relative to blocks that are 200 to 500 meters away.

Figure B.10 highlights estimates using both our baseline bandwidths in red, along with estimates from a variety of different bandwidths. Moreover, we present results from both a predetermined set of bandwidths, as well as estimates in blue that use data-driven bandwidths following the procedure in Calonico, Cattaneo and Titiunik (2014) (CCT). For each of the three specifications in the Figure, we fit the CCT bias-minimizing bandwidth from a border discontinuity design based on 1980 outcomes, passing them on to the design based on 2020 outcomes. We fit different bandwidths derived from different "donut hole" specifications, where we drop blocks with less than 25 or 50 meter centroid distances from the discontinuity.

Across all these combinations of possible bandwidths, we find most yield point estimates close to what we derive in the baseline specification. Estimate robustness holds for both outcomes under the border discontinuity design, and for a computationally tractable subset of bandwidths for the causal forest design. One exception is that estimates attenuate somewhat when bandwidths are smaller than 300 meters, resulting in a heavily truncated sample of blocks. The statistical power concerns mentioned earlier could be behind greatly attenuated effects when choosing a bandwidth of 250 meters.

Alternative outcome definitions. We have also run specifications that use the log transformed count of all people of color, Black residents, or Hispanic residents per block, rather than block level shares. Limiting the analysis to discontinuities with positive counts of residents in each group, we estimate an intensive margin effect of lot size restrictions on the rate of growth of those households across border segments (Chen and Roth, 2024). Our results are similar to the per-acre estimates in Table 7, which imply an average drop of 40% to 60% in POC and Black household counts when moving from the comparison area to the restricted area.

6 Implications for Housing Reforms and for Fair Housing

Many state and local governments are considering land use reforms to remove inefficient or unfair barriers to housing production that much of the research shows are, at least in part, responsible for the housing affordability crisis across the nation. As of 2025, at least twenty-two states have now adopted significant reforms that seek to curb local governments' ability to limit density. Many cities have revised their zoning ordinances as well to allow denser development.

Among the reforms are specific restraints on MLS requirements to allow accessory dwelling units such as backyard cottages or over-the-garage apartments to be built on a lot that already has one residential unit (e.g., Cal. A.B. 68, 2019-20 Leg., §§1-2. 2019). Similarly, some reforms allow existing lots to be split into two or more lots, thereby reducing the lot size by half or more (Cal. S.B. 9, 2021-22 Leg., §1. 2021; Been, Zhang and Kazis (2023)). Others require local governments to allow "gentle density" — such as duplexes and triplexes — to be built on areas zoned for (and applying minimum lot sizes to) single family homes (Baca, McAnaney and Schuetz (2019), Been, Zhang and Kazis (2023)). Such reforms will result in modest increases in density. Allowing a lot of about 5000 square feet to be subdivided into lots as small as 1400 square feet, as Houston did (Wegmann, Baqai and Conrad, 2023), would increase density at least to the levels of our moderate-impact cluster, for example. Allowing attached two to four family homes on lots of typical sizes for single-family homes, as several states have mandated (e.g., Ore. H.B. 2001, 80th Leg. Assemb., 2019 Reg. Sess. §8 (2019)) would increase density at the levels of our high-impact cluster. Our results show that those increases in density could result in important increases in the share of people of color in the neighborhoods.

As state and local governments consider reforms, they will be faced with many political tradeoffs, because zoning change motivates competing interest groups to limit how widely applied the reforms are in order to avoid additional density near them (Hankinson and Magazinnik, 2023). Reforms that override existing requirements in relatively large-lot, low-impact MLS districts near housing already developed at densities similar to those requirements will need to be carefully considered if they jeopardize the passage of broader legislative packages. On the other hand, some of the reforms have exempted smaller jurisdictions from their requirements. Oregon, for example, requires all local governments with 2020 populations of 25,000 or more to allow duplexes, triplexes, quadplexes, townhouses, and cottage clusters on land zoned for single-family dwellings (H.B. 2001, 80th Leg. Assemb., 2019 Reg. Sess. §2 (Or. 2019), Act of Aug. 8, 2019, ch. 639, 2021 Or. Laws). Our conservative prediction is that 27%

of the land area where MLS requirements in the high and moderate impact clusters are located are in jurisdictions below that population threshold, so states should carefully weigh the costs and benefits of exempting those jurisdictions.

Further, as noted in Section 3, a likely mechanism by which MLS requirements may affect the racial composition of neighborhoods is by raising the price of housing, and thereby pricing out lower-income households. Because of the relationship between income or wealth and race and ethnicity, MLS requirements also may price out people of color. MLS requirements are only one of the ways in which local governments regulate density, and thereby affect the price of housing. Rather than seeking to fine-tune the MLS requirements in every district, only to see recalcitrant local governments use other regulations to achieve the same purpose, the results suggest that requiring local governments to allow a specified minimum density across the jurisdiction, regardless of existing MLS or other requirements, may be a more effective route to change.

Some states are mandating equity assessments of their existing land use regulations (e.g. California Assembly Bill 686 (2018)) (and some local governments are undertaking such assessments on their own accord, e.g., New York City Administrative Code §25-117 and §25-118). Our findings should help those jurisdictions analyze how reforming different MLS requirements may reduce current patterns of residential segregation.

Finally, our findings may be helpful to litigants seeking to establish which lot size minimums qualify for *disparate impact liability*, thereby violating federal or state constitutions, the Federal Fair Housing Act and similar state legislation (Schoenbrod, 1968-1969). Our study uses causal inference methods and granular variation to estimate the effect of specific MLS requirements, relative to the surrounding built form context. We have taken care to control for other factors that could be contributing to differences in the effects MLS requirements have on racial composition across neighborhoods. The evidence can also help litigants and judges shape remedies for the harms exclusionary requirements have already caused.

This study also reveals, however, just how onerous — if not impossible — it is to meet the evidentiary standard for disparate impact based on the United States Supreme Court's theories (Williams and Seicshnaydre, 2017). The Court has demanded that those challenging land use regulations point to a specific policy and prove "robust causality" between that policy and disparate effects by race (Texas Department of Housing, (2015)). The Court also noted that plaintiffs must identify a "policy" that is causing the alleged disparate impact, and cast doubt on whether a "one-time decision" could even be a "policy." As we've noted, even when pooling data across 120 metropolitan areas, we had to trade off refinements to the analysis against

concerns with sample size. Courts could consider our findings and determine that the "policy" at issue is the decision to impose MLS requirements, not the choice of whether to apply a particular minimum (such as one that would fall into the high-impact cluster we identify) to a particular neighborhood.

Similarly, the Supreme Court noted that when many factors go into a policy decision, it may be difficult to prove causation. To the extent that we also find zoning district boundaries overlap with other issues (neighborhood features, congruence with other boundaries, and much more), a court could refuse to find that a particular zoning district border, drawn for a myriad of reasons, "causes" the disparate impact that we estimate for the MLS requirements in particular contexts that are in the high-impact cluster. While the Supreme Court purported to uphold the use of disparate impact theories in *Inclusive Communities*, our work shows that the requirements it imposes to prove disparities are likely to make that approach often unavailable in actuality.

7 Conclusion

We provide evidence on which specific minimum lot size requirements matter for the policy objective of increasing the access all races and ethnicities have to more advantaged communities. We show that a not insignificant share of the restrictions do affect the demographics of neighborhoods.

Our paper develops a procedure to create a novel dataset of MLS requirements. The procedure that generates the resulting data predicts where local governments were most likely to apply a requirement that was a constraint on levels of density that developers could supply, judging from the density of less regulated adjacent development. Applying modern causal inference techniques that uncover treatment heterogeneity in big data, our analysis sheds light on how a myriad of MLS requirements from half an acre to as low as 6000 square feet, planned decades in the past to impose densities substantially more restrictive than adjacent neighborhoods, have statistically significant and practically important effects on the racial composition of affected neighborhoods.

More broadly, we employ machine learning and econometric methods to produce land use regulation data that are designed for causal inference over descriptive analysis. With our curated dataset of MLS requirement borders and longitudinal data, we plan to further uncover which jurisdictions persist in restricting the use of land and the externalities they impose on regional development.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder. 2021. "The Effects of the 1930s HOLC "Redlining" Maps." *American Economic Journal: Economic Policy*, 13(4): 355–392.

Albouy, David, and Gabriel Ehrlich. 2018. "Housing Productivity and the Social Cost of Land-Use Restrictions." *Journal of Urban Economics*, 107: 101–120.

Aliprantis, Dionissi, Daniel R. Carroll, and Eric R. Young. 2024. "What Explains Neighborhood Sorting by Income and Race?" *Journal of Urban Economics*, 141: 103508.

Ananat, Elizabeth Oltmans. 2011. "The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality." *American Economic Journal: Applied Economics*, 3(2): 34–66.

Appelbaum, Yoni. 2025. *Stuck: How the Privileged and the Propertied Broke the Engine of American Opportunity.* Random House Publishing Group.

Archer, Deborah N. 2021. "Transportation Policy and the Underdevelopment of Black Communities." *Iowa Law Review*, 106(2125).

Asquith, Brian, Brad Hershbein, Tracy Kugler, Shane Reed, Steven Ruggles, Jonathan Schroeder, Steve Yesiltepe, and David Van Riper. 2022. "Assessing the Impact of Differential Privacy on Measures of Population and Racial Residential Segregation." The MIT Press Special Issue 2.

Athey, Susan, and Stefan Wager. 2019. "Estimating Treatment Effects with Causal Forests: An Application." *Observational studies*, 5(2): 37–51.

Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. "Generalized Random Forests." *The Annals of Statistics*, 47(2): 1148–1178.

Baca, Alex, Patrick McAnaney, and Jenny Schuetz. 2019. ""Gentle" Density Can Save Our Neighborhoods."

Bagagli, Sara. 2023. "The (Express) Way to Segregation: Evidence from Chicago." *Working Paper*.

Bartik, Alexander, Arpit Gupta, and Daniel Milo. 2024. "The Costs of Housing Regulation: Evidence From Generative Regulatory Measurement."

Bartik, Alexander W., and Evan Mast. 2023. "Black Suburbanization: Causes and Consequences of a Transformation of American Cities." W.E. Upjohn Institute.

Baum-Snow, Nathaniel. 2023. "Constraints on City and Neighborhood Growth: The Central Role of Housing Supply." *Journal of Economic Perspectives*, 37(2): 53–74.

Been, Vicki, Helen Zhang, and Noah Kazis. 2023. "Allowing More and Different Types of Housing." The NYU Furman Center for Real Estate and Urban Policy.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer. 2024. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice." *American Economic Review*, 114(5): 1281–1337.

Bronin, Sara C., Scott Markley, Aline Fader, and Evan Derickson. 2023. "How to Make a Zoning Atlas 2.0: The Official Methodology of the National Zoning Atlas."

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica : journal of the Econometric Society*, 82(6): 2295–2326.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2018. "Manipulation Testing Based on Density Discontinuity." *Stata Journal*, 18(1): 234–261.

Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Yingjie Feng. 2024. "On Binscatter." *American Economic Review*, 114(5): 1488–1514.

Chen, Jiafeng, and Jonathan Roth. 2024. "Logs with Zeros? Some Problems and Solutions*." *The Quarterly Journal of Economics*, 139(2): 891–936.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates*." *The Quarterly Journal of Economics*, 133(3): 1163–1228.

Christensen, Peter, and Christopher Timmins. 2022. "Sorting or Steering: The Effects of Housing Discrimination on Neighborhood Choice." *Journal of Political Economy*, 130(8): 2110–2163.

Chyn, Eric, Kareem Haggag, and Bryan A. Stuart. 2025. "The Effects of Racial Segregation on Intergenerational Mobility: Evidence from Historical Railroad Placement." *American Economic Journal: Applied Economics*.

Chyn, Eric, Robert Collinson, and Danielle H Sandler. 2025. "The Long-Run Effects of America's Largest Residential Racial Desegregation Program: Gautreaux*." *The Quarterly Journal of Economics*, 140(3): 2213–2267.

Cilluffo, Anthony, and Rakesh Kochhar. 2018. "Income Inequality in the U.S. Is Rising Most Rapidly Among Asians."

Clarke, Wyatt, and Matthew Freedman. 2019. "The Rise and Effects of Homeowners Associations." *Journal of Urban Economics*, 112: 1–15.

CT, Desegregate. 2025. "Small Lots in Smart Places." *https://www.desegregatect.org/lots*.

Cui, Tianfang. 2024. "Did Race Fence Off The American City? The Great Migration and the Evolution of Exclusionary Zoning."

Ellickson, Robert. 2021. "The Zoning Straitjacket: The Freezing of American Neighborhoods of Single-Family Houses." *96 Indiana Law Journal 395 (2021)*, 96(2).

Ellickson, Robert C. 1977. "Suburban Growth Controls: An Economic and Legal Analysis." *The Yale Law Journal*, 86(3): 385–511.

Fischel, William A. 1987. *The Economics of Zoning Laws: A Property Rights Approach to American Land Use Controls.* JHU Press.

Furth, Salim, and MaryJo Webster. 2022. "Single-Family Zoning and Race: Evidence from the Twin Cities."

Gallagher, Ryan, Allison Shertzer, and Tate Twinam. 2024. "Zoning and the American Suburb."

Gardner, Charles. 2023. "Urban Minimum Lot Sizes: Their Background, Effects, and Avenues to Reform."

Geverdt, Doug. 2018. "School Attendance Boundary Survey (SABS) File Documentation: 2015-2016."

Guerra, Erick, Gilles Duranton, and Xinyu Ma. 2024. "Urban Roadway in America: The Amount, Extent, and Value."

Gyourko, Joe, 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.

Hamilton, Bruce W. 1975. "Zoning and Property Taxation in a System of Local Governments." *Urban Studies*, 12(2): 205–211.

Hankinson, Michael, and Asya Magazinnik. 2023. "The Supply-Equity Trade-Off: The Effect of Spatial Representation on the Local Housing Supply." *The Journal of Politics*, 85(3): 1033–1047.

Holian, Matthew J., and Matthew E. Kahn. 2015. "Household Carbon Emissions from Driving and Center City Quality of Life." *Ecological Economics*, 116(C): 362–368.

Inman, Robert P., and Daniel L. Rubinfeld. 1979. "The Judicial Pursuit of Local Fiscal Equity." *Harvard Law Review*, 92(8): 1662–1750.

Karas, D. 2015. "Highway to Inequity: The Disparate Impact of the Interstate Highway System on Poor and Minority Communities in American Cities."

Keele, Luke J., and Rocío Titiunik. 2015. "Geographic Boundaries as Regression Discontinuities." *Political Analysis*, 23(1): 127–155.

Kramer, Rory. 2018. "Testing the Role of Barriers in Shaping Segregation Profiles: The Importance of Visualizing the Local Neighborhood - Rory Kramer, 2018." *Environment and Planning B: Urban Analytics and City Science*.

Kulka, Amrita. 2019. "Sorting into Neighborhoods: The Role of Minimum Lot Sizes."

Kulka, Amrita, Aradhya Sood, and Nicholas Chiumenti. 2023. "Under the (Neighbor)Hood: Understanding Interactions Among Zoning Regulations."

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

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider. 2015. "The Housing Market(s) of San Diego." *American Economic Review*, 105(4): 1371–1407.

Li, Nicholas Y. 2023. "Racial Sorting, Restricted Choices, and the Origins of Residential Segregation in US Cities."

Logan, John R, and Brian J. Stults. 2021. "Metropolitan Segregation: No Breakthrough in Sight."

Macek, James. 2024. "Housing Regulation and Neighborhood Sorting across the United States."

Mahajan, Avichal. 2024. "Highways and Segregation." *Journal of Urban Economics*, 141: 103574.

Maheshri, Vikram, and Kenneth Whaley. 2025. "The Ubiquity of Tiebout Sorting Across Neighborhood Boundaries."

Manson, Steven, Jonathan Schroeder, David Van Riper, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles. 2024. *IPUMS National Historical Geographic Information System: Version 19.0 [Dataset].* Minneapolis, MN:IPUMS.

Massey, Douglas S., and Jonathan Tannen. 2018. "Suburbanization and Segregation in the United States: 1970–2010." *Ethnic and Racial Studies*, 41(9): 1594–1611.

Mleczko, Matthew, and Matthew Desmond. 2023. "Using Natural Language Processing to Construct a National Zoning and Land Use Database." *Urban Studies*, 00420980231156352.

Monarrez, Tomás, and David Schönholzer. 2023. "Dividing Lines: Racial Segregation across Local Government Boundaries." *Journal of Economic Literature*, 61(3): 863–887.

Monarrez, Tomás E. 2023. "School Attendance Boundaries and the Segregation of Public Schools in the United States." *American Economic Journal: Applied Economics*, 15(3): 210–237.

Noonan, Douglas S. 2005. "Neighbours, Barriers and Urban Environments: Are Things 'Different on the Other Side of the Tracks'?" *Urban Studies*, 42(10): 1817–1835.

Oh, Sun Jung, and John Yinger. 2015. "What Have We Learned From Paired Testing in Housing Markets?" *Cityscape*, 17(3): 15–60.

Pendall, Rolf. 2000. "Local Land Use Regulation and the Chain of Exclusion." *Journal of the American Planning Association*, 66(2): 125–142.

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." The Brookings Institution.

Pons, Martina. 2024. "Quantile on Quantiles."

Quigley, John M., Steven Raphael, and Larry A. Rosenthal. 2004. "Local Land Use Controls and Demographic Outcomes in a Booming Economy." *Urban Studies*, 41: 389–421.

Rehill, Patrick. 2025. "How Do Applied Researchers Use the Causal Forest? A Methodological Review." *International Statistical Review*, 93(1).

Resseger, Matthew. 2022. "The Impact of Land Use Regulation on Racial Segregation: Evidence from Massachusetts Zoning Borders."

Richards, Meredith P. 2014. "The Gerrymandering of School Attendance Zones and the Segregation of Public Schools: A Geospatial Analysis." *American Educational Research Journal*, 51(6): 1119–1157.

Rosenthal, Stuart S. 2014. "Are Private Markets and Filtering a Viable Source of Low-Income Housing? Estimates from a "Repeat Income" Model." *American Economic Review*, 104(2): 687–706.

Rothstein, Richard. 2017. *The Color of Law: A Forgotten History of How Our Government Segregated America.* . First edition. ed., W.W. Norton.

Rothwell, Jonathan, and Douglas S. Massey. 2009. "The Effect of Density Zoning on Racial Segregation in U.S. Urban Areas." *Urban Affairs Review*, 44(6): 779–806.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari C.W. Williams. 2025. *IPUMS USA: Version 16.0 [Dataset].* Minneapolis, MN:IPUMS.

Saporito, Salvatore, and David Van Riper. 2016. "Do Irregularly Shaped School Attendance Zones Contribute to Racial Segregation or Integration?" *Social Currents*, 3(1): 64–83.

Schmidt, Benjamin M. 2018. "Historical US City Populations."

Schoenbrod, David S. 1968-1969. "Large Lot Zoning." *Yale Law Journal*, 78: 1418.

Schönholzer, David. 2024. "Measuring Preferences for Local Governments."

Semenova, Vira, and Victor Chernozhukov. 2021. "Debiased Machine Learning of Conditional Average Treatment Effects and Other Causal Functions." *The Econometrics Journal*, 24(2): 264–289.

Severen, Christopher, and Andrew J. Plantinga. 2018. "Land-Use Regulations, Property Values, and Rents: Decomposing the Effects of the California Coastal Act." *Journal of Urban Economics*, 107(C): 65–78.

Song, Jaehee. 2025. "The Effects of Residential Zoning in U.S. Housing Markets." *Journal of Urban Economics*, 149: 103784.

Sood, Aradhya, and Kevin Ehrman-Solberg. 2024. "The Long Shadow of Housing Discrimination: Evidence from Racial Covenants."

Southern Burlington County NAACP v. Mt. Laurel. 1975.

Taylor, Keeanga-Yamahtta. 2019. *Race for Profit: How Banks and the Real Estate Industry Undermined Black Homeownership. Justice, Power, and Politics*, Chapel Hill:University of North Carolina Press.

Texas Dept. of Housing and Community Affairs v. Inclusive Communities Project, Inc. 2015.

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

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

Turner, Margery, and Stephen Ross. 2003. "Discrimination in Metropolitan Housing Markets Phase II: Asians and Pacific Islanders." *Economics Working Papers*, , (No. 200318).

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

Wegmann, Jake, Aabiya Noman Baqai, and Josh Conrad. 2023. "Here Come the Tall Skinny Houses: Assessing Single-Family to Townhouse Redevelopment in Houston, 2007?2020." *Cityscape*, 25(2): 171–202.

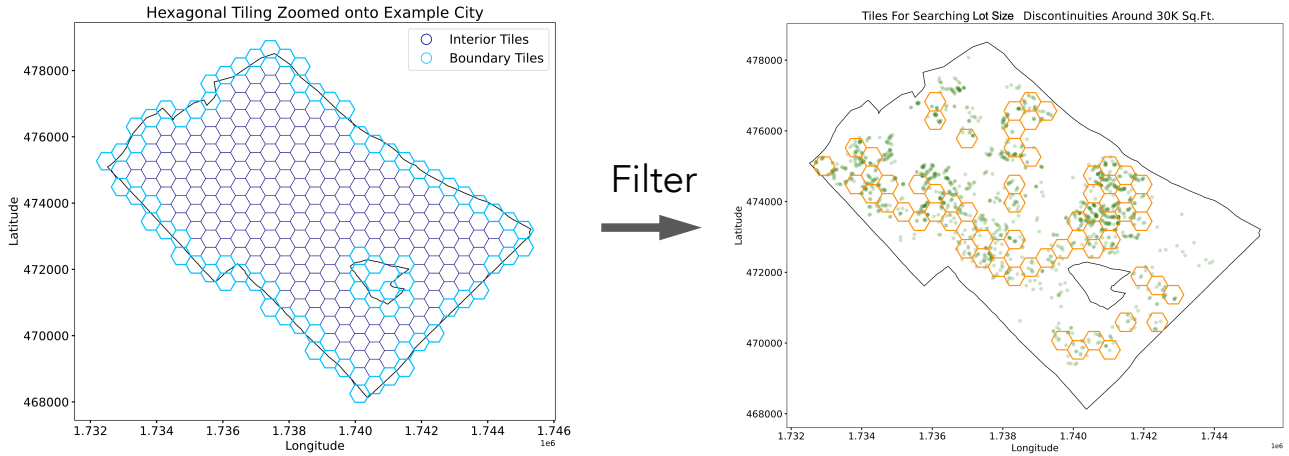
Weiwu, Laura. 2024. "Unequal Access: Racial Segregation and the Distributional Impacts of Interstate Highways in Cities."

White House Council of Economic Advisors. 2021. "Exclusionary Zoning: Its Effect on Racial Discrimination in the Housing Market | CEA."

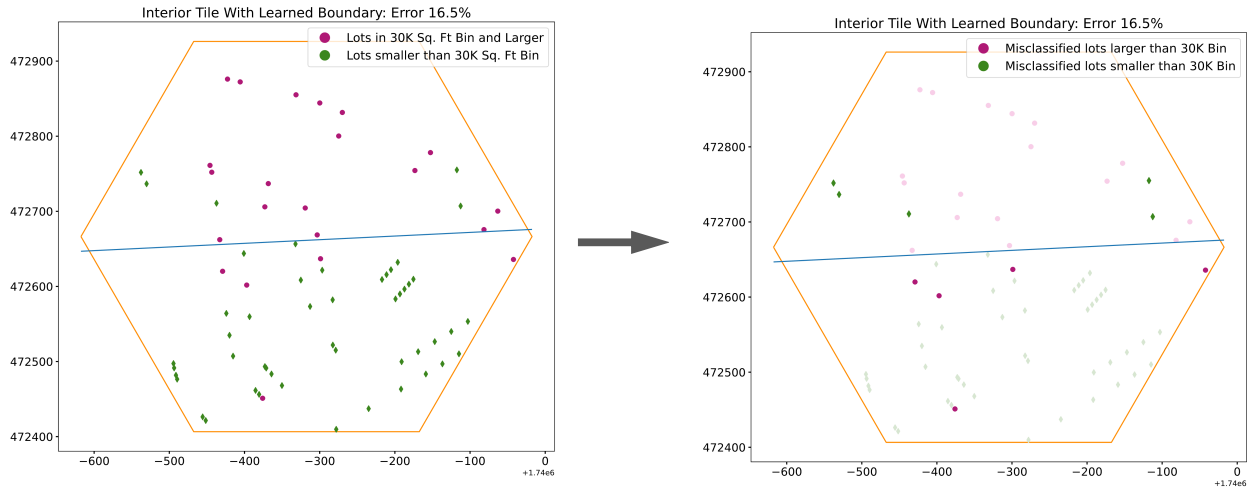
Williams, Morgan, and Stacy Seicshnaydre. 2017. "The Legacy and the Promise of Disparate Impact." In *The Fight for Fair Housing: Causes, Consequences, and Future Implications of the 1968 Federal Fair Housing Act*. Routledge.

Figure 1: Illustrating detection of border segments with lot size discontinuities

(a) Filtering down to tiles that include lot size discontinuities



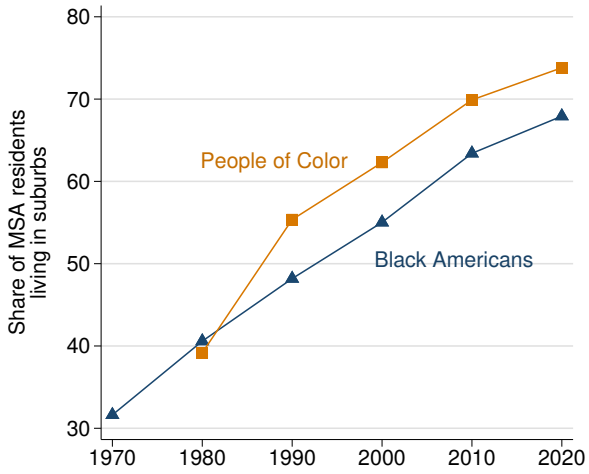
(b) Local detection of MLS border segments



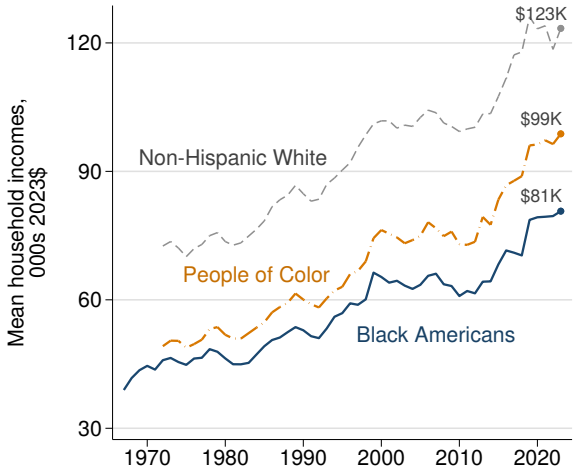
Notes: Using multiple panels, this figure illustrates the workflow to detect lot size discontinuities that are evidence of borders between different MLS requirements. The example case is the 30,000 square foot minimum lot size in Lower Merion Township, PA. Section 2 describes in further detail the two steps of the procedure: Checking for presence of lot size discontinuities across tiles in Panel (a) and finding linear borders consistent with regulatory district borders in Panel (b), following application of support vector machines. Black boundaries reflect the boundaries of a specific zoning jurisdiction, while orange boundaries reflect the limits of each interior tile. Details of the algorithm are written in Section 2.
Sources: Calculations from CoreLogic Tax Records.

Figure 2: National trends behind the diversification of U.S. suburbs

(a) Suburbanization of People of Color



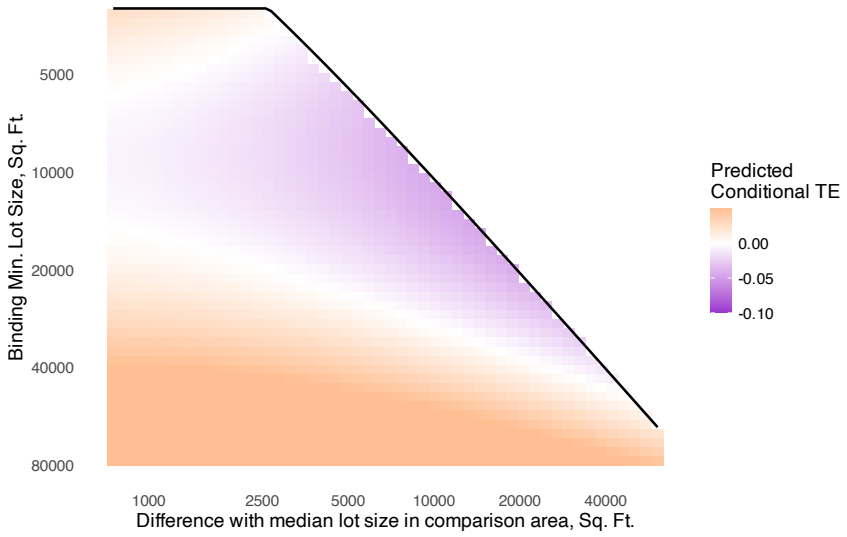
(b) Real Wage Growth for People of Color



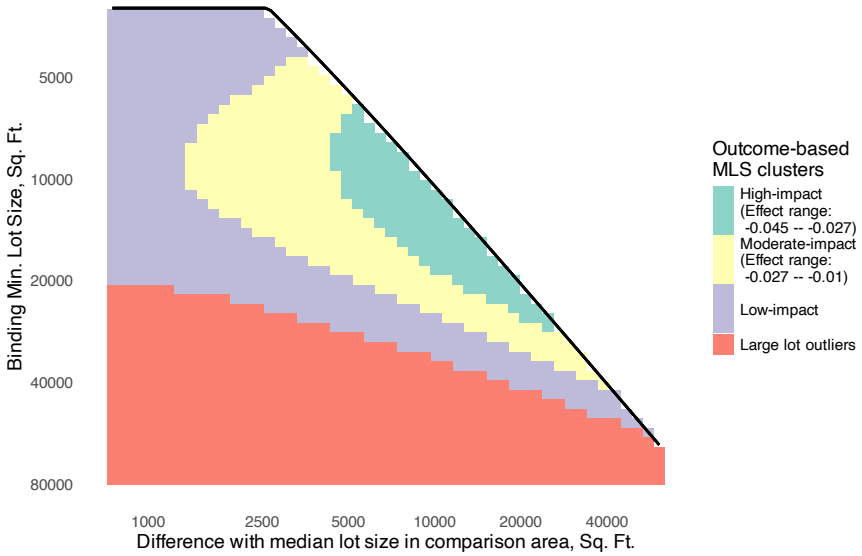
Notes: Panel (a) of this Figure separates the tract geography within Metropolitan Statistical Areas (MSAs) into a time-consistent "central city core" based on the central city's 1960 borders, and the remaining "suburban area" of the MSA. We visualize dynamics from decadal Censuses on the share of people in racial or ethnic categories in MSAs who live in the suburban area instead of the city core. "People of Color" includes all Census respondents not identifying as Non-Hispanic White. Panel (b) plots average household income time series from Census historical tables. The series for "People of Colour" is the overall mean subtracted by the mean for the Non-Hispanic White category.
Sources: Calculations from 1970–2020 NHGIS Tables (Manson et al., 2024)), Census Bureau, Census Historical Income Tables, Table H-5.

Figure 3: Constructing impact clusters for MLS requirements

(a) Model-generated treatment effects on neighborhood composition

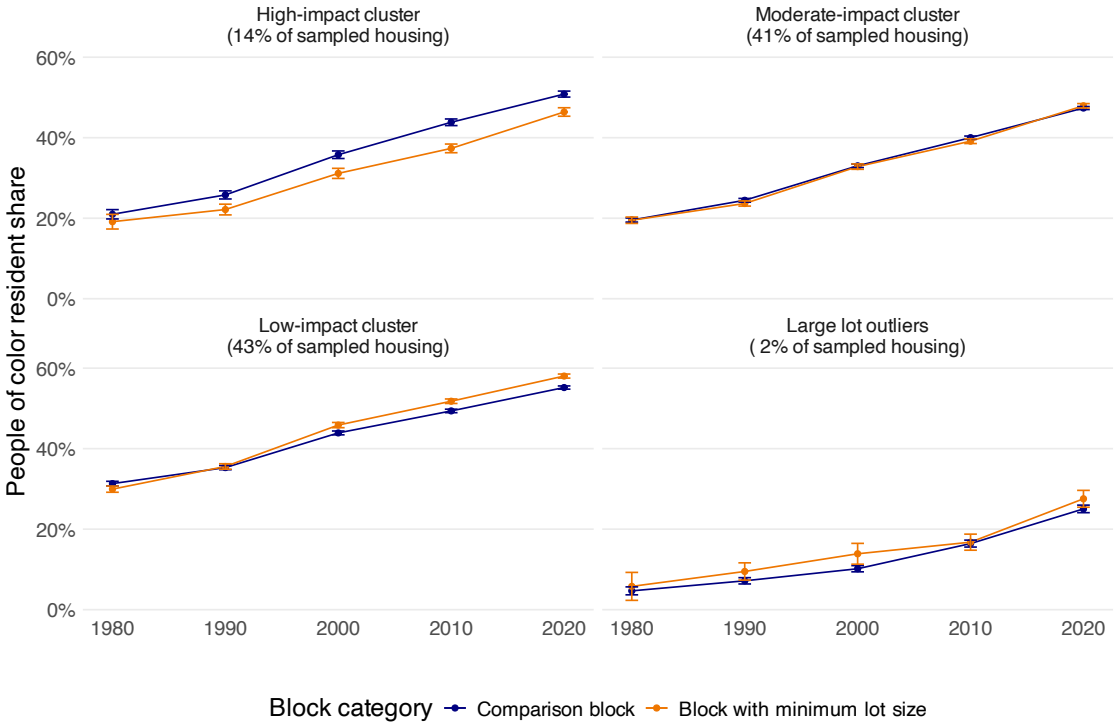


(b) Defining impact clusters based on treatment effects



Notes: This figure illustrates the division of MLS border segment, following the methodology of Section 2, into four impact clusters. Border segments are in the same impact cluster if they have similar model-generated treatment effects on block-level people of color resident shares in the treated blocks. Panel (a) shows model-generated treatment effects, which smooths out the discontinuity-level treatment effect predictions from the Athey and Wager (2019) causal forest model, as it varies over two features: the MLS applied on treated blocks and the gap in density compared to the surrounding comparison area. Panel (b) splits the model-generated treatment effects based on the criteria listed in the legend, as well as those in Section 4.2.
Sources: Calculations from 2000–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

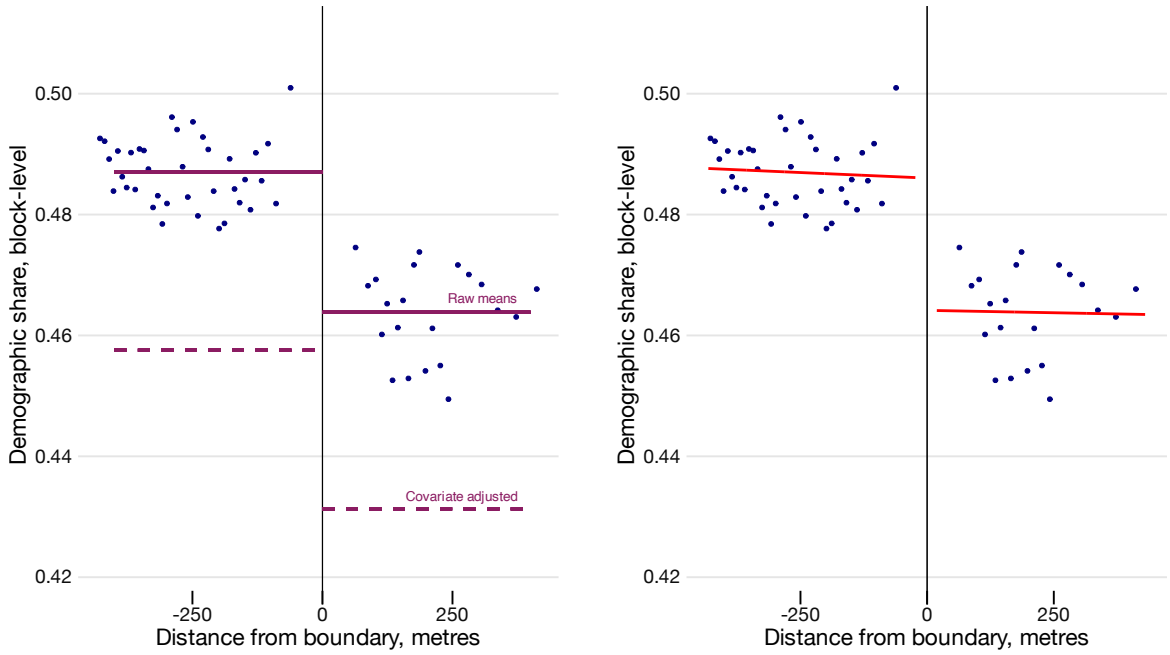
Figure 4: Dynamics of neighborhood diversity across impact clusters



Notes: To construct this figure, we classify all border segments featuring lot size discontinuities, defined following Section 2, into the impact clusters defined following Section 4.2. Blocks around the borders are separated into ones on the sides of lot size restricted areas, and the ones within the comparison areas with higher densities than the requirement. In each Census wave, we plot the population-weighed share, aggregating by impact cluster all the block-level shares on one of the two sides.
Sources: Calculations from 1980–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Figure 5: Nonparametric effects of MLS requirements producing high-impact lot size changes

(a) Output from causal forest design (b) Output from border discontinuity design

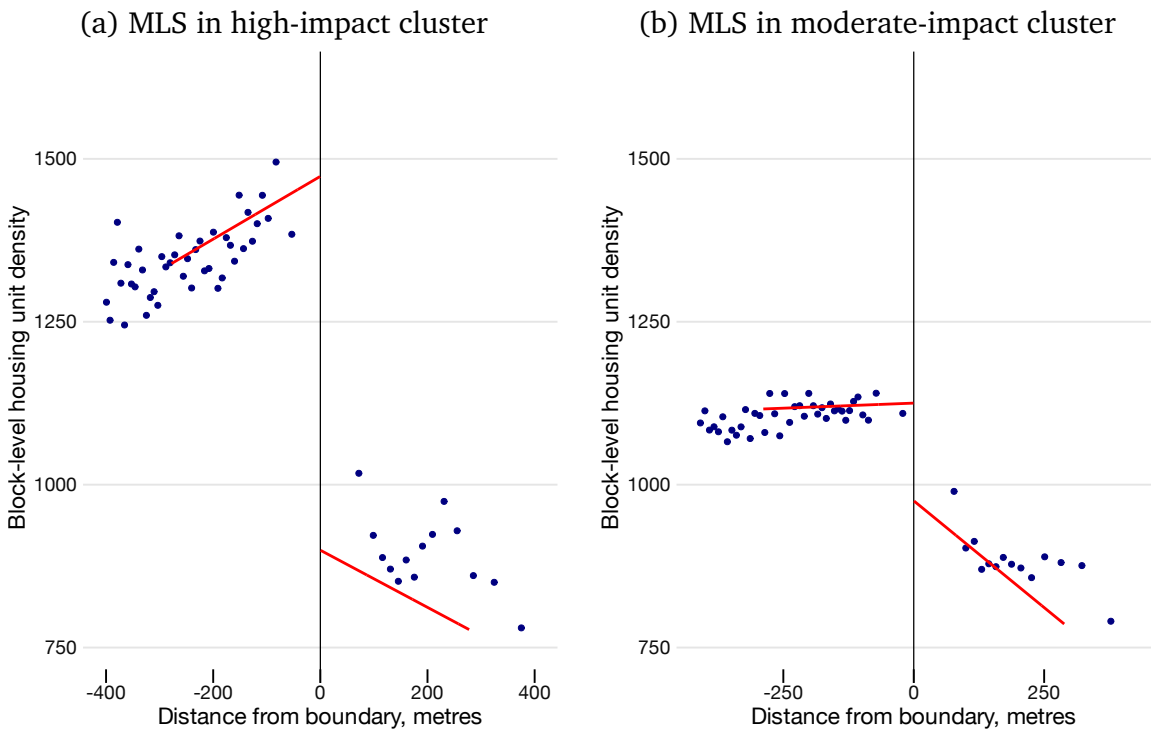


Notes: This figure presents multiple approaches to infer MLS requirements effects on racial composition as of 2020, estimated over the high-impact cluster of minimum lot size border segments constructed in Section 4.2. People of color share is defined as all residents in a block who are not Non-Hispanic White. In each panel, the points are binned means outputted using the procedure from Cattaneo et al (2024), and after being demeaned by county-level fixed effects. Overlaid on the points are regression functions based on the fixed effects model

$$Y_b = \alpha_{\text{cty}(b)} + \beta^t \mathbf{1}[Dist_b \geq 0] + g(\mathbf{X}_b) + \varepsilon_b,$$

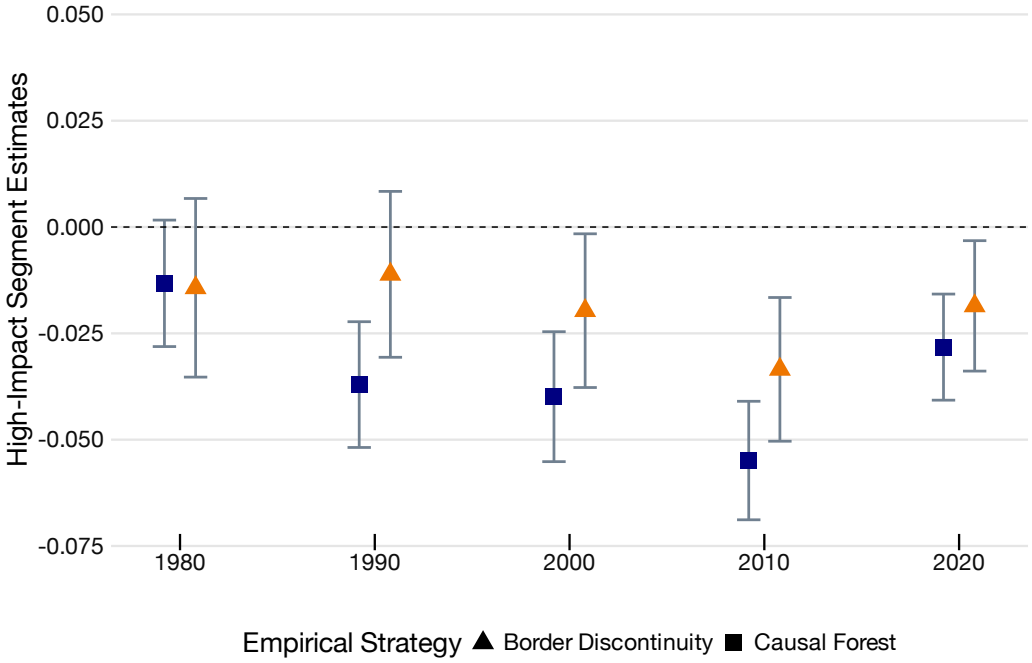
where Panel (a) shows means following regression adjustment based on the control function g trained on a causal forest (Athey and Wager, 2019) and Panel (b) shows trends fit on local linear functions in distance. Panel (a) shows, for comparison, the equally-weighted mean over all the binned scatterplot points.
Sources: Calculations from 2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Figure 6: Nonparametric effects of MLS requirements on observed density



Notes: This figure presents border discontinuity effects to infer MLS requirements effects on built density as of 2020, estimated over the two impact clusters of minimum lot size border segments constructed in Section 4.2. Of the two, the high-impact cluster is predicted to have stronger average effects on limiting racial composition change based on the preliminary regression in Section 4.2. Built density is calculated using Census counts of all housing units recorded in a block, divided by the area of the underlying block in square kilometers. In each panel, the points are binned means outputted using the procedure from Cattaneo et al (2024), and after being demeaned by county-level fixed effects. Overlaid on the points are regression functions based on trends fit on local linear functions in distance.
Sources: Calculations from 2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Figure 7: Dynamic effects of MLS requirements producing high-impact lot size changes



Notes: This figure presents border discontinuity effects for the block-level shares of people of color $Share^m$, using data from 1980 to 2020. Effects are estimated over the high-impact cluster defined in Section 4.2. The sample is limited to blocks whose centroid distances to the nearest MLS border segment, detected through the procedure in Section 2, is 400 meters or less. We present effects estimated from two models, with different identifying assumptions listed in 4.3:

$$Y_b = \alpha_{\text{seg}(b)} + \beta^t \mathbf{1}[Dist_b \geq 0] + g(\mathbf{X}_b) + \varepsilon_b,$$

where Panel (a) shows estimates following regression adjustment based on the control function g trained on a causal forest (Athey and Wager, 2019) and Panel (b) shows estimates controlling for local linear functions in distance. Fixed effects are applied at the border segment level. 95% confidence intervals are presented with standard errors clustered at the border segment level.
Sources: Calculations from 1980–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Table 1: Value of parameters in automated border detection method

Parameter name	Symbol	Value
Hexagonal tile radius, meters	R	300
Count of lots at bunched sizes	$\underline{N}$	8
Bunching range factor	$\overline{M}$	1.25
Misclassification rate threshold	$\overline{M}_{err}$	0.35
Neighboring lots for KNN extension	k_{mult}	5

Notes: This Table lists the key parameters used in different stages of the automated detection procedure for lot size discontinuities, as detailed in Section 2. Except for dimensionless parameters, parameter units are given in the leftmost column.

Table 2: Detected lot size discontinuities' proximity to other borders (350m buffer)

Panel (a): Share of detected lot size discontinuities near real zoning borders			
	Near zoning segment	Far from zoning	Zoning far from discontinuity
Percentage, raw	50.6%	49.4%	53.2%
Percentage, pruned	72.9%	27.1%	53.2%
Panel (b): How zoning border segments overlap with other urban features			
	Cui & Been segments	Straight zoning segments	All zoning segments
Primary roads & thoroughfares	39.8%	49.4%	50.4%
HOLC Class A	8.3%	5.3%	5.6%
HOLC B–C	14.4%	16.4%	18.1%
HOLC Class D	3.5%	5.8%	8.1%
Rail tracks	8.3%	14.6%	16.3%
Elem. school catchment area	18.9%	20.4%	21.8%
Middle school catchment area	3.6%	2.8%	3.5%
High school catchment area	1.1%	0.6%	0.9%

Notes: Over Boston metropolitan jurisdictions identified in the MAPC Zoning Atlas, Panel (a) calculates the share of procedurally detected border segments with lot size discontinuities (The "Cui & Been segments") that are 350 meters or less away from a straight line zoning border recorded in the MAPC Atlas. We report shares at two different stages of sample post-processing. The other conditional probability — the share of MAPC straight line zoning borders close to a detected border segment — is recorded in the rightmost column. Panel (b) considers separately the two types border segments evaluated in Panel (a), along with the universe of zoning segments in the MAPC Atlas where MLS requirements change. For each of these three types of segments, we compare overlap rates of segments with other historical boundaries in the Boston area that can cause neighborhood stratification. Overlap in Panel B is defined as the border segment to be within any 175-meter radius buffer centered at a historical boundary. Detailed definitions of border segment types and of the post-processing stages are provided in Section 2.4.
Sources: Calculations from CoreLogic records; the MAPC Zoning Atlas; digitized HOLC "redlining" maps from Mapping Inequality; national roads and railroads data from the Census Bureau; and the SABS sample of school catchment areas.

Table 3: Summary statistics for areas treated with MLS requirements

Statistic		Range of MLS in subsample				
		Total	1–6000	6000–11000	11000–22000	22000+
MLS level (sq. ft.)	Mean	9,667.9	5,157.6	8,176.9	14,959.8	37,063.7
	St. Dev.	7,845.1	803.9	1,469.3	3,602.9	19,436.4
	N	78,402	24,847	35,473	14,948	3,134
Median year built	Mean	1960	1953	1961	1967	1974
	St. Dev.	22	25	21	19	17
	N	65,726	19,749	30,000	13,177	2,800
Distance from CBD (km)	Mean	25.9	23.0	26.2	28.4	32.1
	St. Dev.	19.3	18.6	19.6	18.8	20.4
	N	78,402	24,847	35,473	14,948	3,134
Population of treated area, 2010	Mean	203.0	371.7	209.2	103.7	60.9
	St. Dev.	297.3	463.5	257.2	132.7	81.7
	N	24,791	4,932	10,720	7,058	2,089

Notes: This summary table plots statistics on the predetermined character of residential development, for blocks determined to be in district imposing an MLS requirement based on the detection procedure in Section 2. The level of observation is Census block based on 2010 boundaries. Both the full sample is summarized, along with four subsamples based on the level of lot size regulation applied in treated areas. MLS level and year built variables are defined at the level of interior cell, then matched to blocks contained in those cells. The final variable is observed at the level of treatment area, which is the union of all blocks identified to be surrounding a border segment and subject to a MLS requirement.
Sources: Calculations from 1980–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Table 4: Covariate balance around border segments at start of analysis

1980 Outcomes	(1)	(2)
People of Color share	0.013	-0.014
	(0.027)	(0.011)
Black resident share	0.008	-0.012
	(0.020)	(0.009)
Hispanic resident share	0.001	-0.006
	(0.011)	(0.006)
Asian resident share	0.003	0.002
	(0.004)	(0.004)
Rental units Share	-0.046	-0.051*
	(0.033)	(0.030)
Share of pop. $>$ 62 years	-0.013	-0.017*
	(0.010)	(0.010)
Share of pop. $\leq$ 18 years	0.017**	0.013*
	(0.008)	(0.008)
Impact Cluster	High-Impact (14% of sample)	
Jurisdiction FE	X	
Border FE		X
Total N	11545	

Significance levels: * = 10%; ** = 5%; *** = 1%.

Notes: This table presents outputs of border discontinuity designs over Census blocks b in year t for residents of race or ethnicity m,

$$Z_{bt}^{m} = \alpha_{seg(b)t} + \beta^{t}\mathbf{1}[Dist_b \geq 0] + \eta_{-}^{t} Dist_b + \eta_{+}^{t} Dist_b \cdot \mathbf{1}[Dist_b \geq 0] + \varepsilon_{bt},$$

where Z describes a confounding variable at the block level. Each column plots a separate confounder available in the Census tabulations. All outcomes are calculated based on available 1980 block-level data, and using blocks on both sides of high-impact MLS border segments identified in Section 4.2. Point estimates and standard errors are based on the robust nonparametric procedure in Calonico, Cattaneo and Titiunik (2014). Two specifications are provided with two levels of fixed effects, at the jurisdiction or border segment level. Standard errors are calculated clustering at the county-year level.

Sources: Calculations from 1980–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Table 5: Pooled effects of MLS requirements on racial composition

	Block Nonwhite Shares		Block Black Shares	
BD Estimates	(1)	(2)	(1)	(2)
2020 Data	0.004	0.011***	-0.003	-0.002
	(0.003)	(0.004)	(0.003)	(0.003)
2010 Data	-0.000	0.004	-0.003	-0.000
	(0.004)	(0.004)	(0.003)	(0.003)
1980 Data	0.007	-0.001	0.006	0.001
	(0.010)	(0.004)	(0.006)	(0.003)
Jurisdiction FE	X		X	
Border FE		X		X
Total N	666919		666919	

Significance levels: * = 10%; ** = 5%; *** = 1%.

Notes: This table presents outputs using the border discontinuity research design to estimate causal effects of MLS requirements on neighborhood resident shares by race and ethnicity. Over Census blocks b in year t, we run the regression:

$$\text{Share}_{bt} = \alpha_{j(b)t} + \beta^t \mathbf{1}[Dist_b \geq 0] + \eta^t_- Dist_b + \eta^t_+ Dist_b \cdot \mathbf{1}[Dist_b \geq 0] + \varepsilon_{bt},$$

over the people of color (POC) share at the block level Share_{bt} as well just for Black residents, $\text{Share}^{Black}_{bt}$. Results are presented for two sets of fixed effects, and for a sample including blocks with centroid distances of 400 meters or less from the nearest border segment detected using the procedure in Section 2. The total number of blocks used in reported regressions is given, and all regressions are weighed by block-level population. Standard errors are calculated clustering at the border segment-year level.

Sources: Calculations from 1980–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Table 6: How MLS requirements producing high-impact lot size changes change diversity

2020 Outcomes	Causal Forest		Border Discontinuity	
	(1)	(2)	(1)	(2)
People of Color Share	-0.031***	-0.028***	-0.003	-0.019**
	(0.006)	(0.006)	(0.010)	(0.008)
Black American Share	-0.031***	-0.030***	-0.019**	-0.026***
	(0.005)	(0.005)	(0.008)	(0.006)
Hispanic American Share	-0.011***	-0.007*	0.012	0.004
	(0.004)	(0.004)	(0.009)	(0.006)
Asian American Share	0.002	0.001	-0.002	-0.001
	(0.003)	(0.002)	(0.005)	(0.004)
Impact Segment	High-Impact (14% of sample)			
Jurisdiction FE	X		X	
Border FE		X		X
Total N	7628		28116	

Significance levels: * = 10%; ** = 5%; *** = 1%.

Notes: This table presents outputs from two separate research designs to estimate causal effects of MLS requirements on neighborhood resident shares by race and ethnicity. The results here are specifically for MLS requirements classified to be in the high-impact cluster, identified in Section 4.2. Over Census blocks b in 2020 we run regressions of the form:

$$Share_b^m = \alpha_{\text{seg}(b)} + \beta^t \mathbf{1}[Dist_b \geq 0] + g(\mathbf{X}_b) + \varepsilon_b,$$

where the two leftmost columns show estimates following regression adjustment based on the control function g trained on a causal forest (Athey and Wager, 2019) and the two rightmost columns show estimates controlling for local linear functions in distance. Average effects over the high-impact cluster of MLS requirements are calculated for four different racial and ethnic categories in the Census, following 2020 definitions. The analysis sample includes blocks with centroid distances of 400 meters or less from the nearest border segment detected using the procedure in Section 2. The causal forest design treats aggregations of blocks as an observation, while the border discontinuity design is at the block level and weighed by population. Standard errors are calculated clustering at the border segment-year level.

Sources: Calculations from 1980–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.

Table 7: Full Suburban Diversification Effects of Hypothetical MLS Reforms

2020 Outcomes	High-impact Clust.	Moderate-impact Clust.	Low-impact Clust.
People of Color Share	-0.019**	-0.003	0.004
	(0.008)	(0.004)	(0.004)
Black American Share	-0.026***	-0.008***	-0.001
	(0.006)	(0.002)	(0.002)
Residential Density, Persons/sqmi	−1128***	−150.8***	−49
	(111)	(51.0)	(72.2)
Additional POC residents/sqmi under reforms	1480	200	—
Component due to densification	1310	120	—
Additional Black residents/sqmi under reforms	610	130	—
Component due to densification	350	50	—
Total N	28116	95513	100905

Significance levels: * = 10%; ** = 5%; *** = 1%.

Notes: This table presents outputs using the border discontinuity design to estimate causal effects of MLS requirements on neighborhood outcomes. The results here are a subset of 2020 outcomes studied in Table 6, plus a measure of density defined as residents reported in a Census block divided by block area in km^2. We reproduce effects estimated for the high-impact cluster, but also for 2 other notable clusters identified in Section 4.2. Over Census blocks b in 2020 we run regressions of the form:

$$Y_b^m = \alpha_{\text{seg}(b)} + \beta^t \mathbf{1}[Dist_b \geq 0] + \eta_-^t Dist_b + \eta_+^t Dist_b \cdot \mathbf{1}[Dist_b \geq 0] + \varepsilon_{bt},$$

We also produce unit converted rules of thumb for what these estimates imply about how hypothetical MLS reforms in the past could have changed racial composition of different localities, which we explain more in Section 5.4. The analysis sample includes blocks with centroid distances of 400 meters or less from the nearest border segment detected using the procedure in Section 2. Standard errors are calculated clustering at the border segment-year level.

Sources: Calculations from 1980–2020 NHGIS Tables (Manson et al. (2024)) and CoreLogic Tax Records.