Did Race Fence Off The American City? The Great Migration and the Evolution of Exclusionary Zoning *

Tianfang Cui*

February 21, 2024

Latest version available here

Abstract

I study how postwar Black migration caused American cities' suburbs to widely restrict development with minimum lot sizes. When local governments use lot size regulations to restrict the supply of dense housing, bunching is detectable across lot size distributions. I develop an algorithm to measure bunching and produce a national panel for the regulation spanning the 20th Century. Most suburbs adopted lot size controls from 1945–1970, all the while Black Americans left the South for economic opportunity. I find the Great Migration of Black Americans into non-Southern cities accelerated minimum lot size adoption and caused up to half of baseline lot size restrictiveness. Migration of poorer white Americans cause null or small negative effects. A sizable driver of the Great Migration effect is local implementation of school desegregation before 1970. The results support a theory where competition in policy planning still reinforced a minority group's exclusion from local public goods. As homes remain in place longer than their first owners, past lot size restrictiveness can explain today's spatial disparities in racial sorting and economic mobility.

JEL Codes: H73, N92, R31, R52
Keywords: Exclusionary zoning, Land use regulation, Local public goods, Great Migration

*New York University Furman Center. E-mail: tc4089@nyu.edu. I thank Fernando Ferreira, Joseph Gyourko, Benjamin Keys, David Berger, Kirill Borusyak, Ellora Derenoncourt, Edward Glaeser, Nina Harari, Benjamin Lockwood, Jacob Krimmel, Nathan Seegert, Allison Shertzer, Matthew Turner, Harrison Wheeler, Eric Zwick and seminar participants at the Wharton School, the Bank of Canada, Colgate University, the Federal Reserve Board of Governors, University of Pittsburgh, Penn State, Rotman School of Management, Western University, the NBER Summer Institute and the Urban Economics Association Meetings for their comments. I also thank Anne Leavitt-Gruberger for retrieving historical zoning records. Adam Sciara provided excellent research assistance. I gratefully acknowledge support from the Zell-Lurie Real Estate Center Research Sponsor Program at Wharton.
This paper previously circulated under the title: "The Emergence of Exclusionary Zoning Across American Cities."

People are still making zoning decisions without the slightest concern for the outside world...
They operate in the belief that their balkanized municipal universe is the happiest of worlds.

— Babcock (1966)

1 Introduction

Local governments in the United States have the legal right to shape urban form. Since the 1920s, localities possess the *zoning power* to regulate land use within their borders. To take one example, they can forbid dense housing development from prescribed neighborhoods. American courts impose few limits on how local actors can meet residents' demand for such *exclusionary zoning*. These norms held up even as segregation rose in urban areas, followed by a social mobility gap between suburbs and the urban core (Chetty and Hendren, 2018).

This paper identifies how a pivotal demographic shift in U.S. history, the Second Great Migration of Black Americans out of the South, caused local actors to adopt exclusionary zoning. Plausibly exogenous shocks to cities' Black population cause more regulations that limit market provision of dense housing. Migration of poor non-Black groups cannot produce effects of the same magnitude. In contrast, mandated desegregation of local schools substantially influence the size of the Great Migration effect. My results support a *racial exclusionary zoning* mechanism less recognized in economics. Local actors limited access to more neighborhoods, by proxy restricting local public goods, when urban growth involved racial change.

Data availability complicates verifying narratives with evidence. Statistics are scarce on which governments adopted land use regulations earlier or later. Unlike regulations in effect, historical records of local land use decisions are lost or too costly to retrieve.[1] This paper first fills the data gap by measuring the emergence and scope of a widely used regulation — the minimum lot size.

I construct the first national panel on the adoption and restrictiveness of minimum lot sizes, a key regulation for U.S. suburbs. Most Americans today live in suburbs,[2] where I document most governments first adopted minimum lot sizes from 1940 to 1970. This period of wide adoption overlaps with the accepted time frame for the Second Great Migration (Collins, 2021). Using lot size controls as a component of exclusionary zoning practices, I evaluate how much of this correlation in time is causal.

[1] National surveys (Gyourko, Hartley and Krimmel, 2021) and text parsing of active ordinances (Bronin and Ilyankou (2021), Mleczko and Desmond (2023)) substantially improve our understanding of which local land use regulations apply in the last two decades.

[2] 54% of respondents in the 2019 wave of the American Housing Survey self-report living in suburban neighborhoods (Bucholtz, Molfino and Kolko, 2020).

A neighborhood regulated by a minimum lot size requires each housing unit to use up a certain amount of land. A minimum lot size that binds on development has observable implications on homes built after lot size adoption: development should bunch at certain lot sizes relative to the distribution. I introduce an algorithm that recovers when bunching on lot sizes first appears and persists over vintages of the housing stock. I digitize a national sample of historical zoning records, training the algorithm to separate bunching caused by persistent regulations from statistical noise.

I apply the algorithm one local government at a time, using administrative records covering 86 million homes. Together, the data cover 407 U.S. metropolitan areas where 83% of the 2020 U.S. population live. Figure 1 shows one algorithm output: measures of minimum lot size adoption for 4,800 large and mid-sized local governments. I calculate facts novel to the literature: three-fifths of these localities adopted binding minimum lot sizes between 1940 and 1970. Of the adopted regulations, mandates for "restrictive" lot sizes exceeding 7,500 square feet show up in over 70% of localities today.

I also adapt methods from the bunching literature (Kleven (2016), Bertanha et al. (2023)) to introduce an excess mass measure of restrictiveness. Taking the share of homes bunching on lot sizes, I partial out counterfactual demand around those sizes in the absence of regulation. Aggregating excess mass measures across sampled cities, I calculate minimum lot sizes constrained 14% of properties since the 1940s, or at least 7.5 million properties nationwide.

If lot size regulations limit market provision of dense housing, they cause lost land rents in the short run. I interpret high excess mass places as places whose interest groups take short run rent losses to chase ulterior motives. In the long run, past homes built under minimum lot sizes are not easily demolished. I find suggestive evidence that suburbs with high lot size restrictiveness in the past continue to diverge in land use and racial diversity. When the suburbs surrounding central cities are lot size restricted, the disparity in economic mobility between the suburbs and the urban core widens further.

With my lot size outcomes, I exploit panel variation in Black racial change across metropolitan areas' central cities. The cities Black migrants chose, however, may have already seen flight to the suburbs. If those suburbs had planned to exclude further development, their non-racial motives would confound migration's effects on outcomes. Black migrants also had lower incomes than Northern incumbents. Lot size responses to Black migration could be an expression of exclusion only based on income.

I address the first concern with a shift-share instrumental variable strategy, following Boustan (2010) and Derenoncourt (2022). The shift-share instrument predicts actual migration using outmigration growth in Southern Black counties, weighed by cities' pre-1940 exposure to county migration flows. Before 1940, cities outside of the South did not vary much in the Black population share. Instead, cities differed in where in the South their Black population originated. Outmigration shocks based on differential economic change across the South provide plausibly exogenous variation for identifying effects. I carefully follow estimation guidelines proposed in Borusyak, Hull and Jaravel (2022) to justify shift-share instrument validity and derive confidence intervals with correct coverage.

For a panel of non-Southern metropolitan areas, I find a rise in the central city's share of Black residents caused early adoption and more restrictive design of minimum lot sizes. Over 1940–70, a 6 percentage point rise in a central city's share of Black residents explains 28% of minimum lot size adoption across decades. Effects on lot size restrictiveness are greater, at about 57%.

To verify that exclusion on race as opposed to income drive my main effects, I construct a separate migration instrument for shares of Southern white migrants in central cities. As late as 1960, Southern white migrants earned less on average than non-Southern white incumbents. I find cities with higher shares of Southern whites instead adopted less restrictive regulations. This is the opposite of what theories of exclusion only on income predict, where zoning would exclude lower-income residents who free ride on public good consumption.

I further analyze a mechanism that explains the Great Migration effect on lot size outcomes: the supply of exclusionary neighborhoods grows when cities plan to desegregate public good consumption. Civil Rights leaders throughout my analysis period lobbied against public goods segregation, culminating in plans to desegregate neighborhood schools. White incumbents who view this as a disamenity might then find large lot neighborhoods relatively more attractive, because schools serving those neighborhoods are unaffordable for Black families. Using both national cross-sectional variation and a panel event study, I find large effects on the interaction of greater Black migration with public school desegregation in non-Southern metropolitan areas.

1.1 Existing literature and historical context

My results frame the spread of lot size restrictions as a form of postwar *collective action* among white suburbanites, who used regulation to limit development and increase entry costs as the Black population rose. Existing literature like Cutler, Glaeser and Vigdor (1999) attributes the rise and fall of Black-White residential segregation in the 20th Century to white residents' use of legal restrictions or targeted violence.[3] Providing new estimates of how racial exclusionary zoning distorted postwar housing markets, I argue zoning regulations still in place today could have been first motivated as a reaction against racial demographic change.

The first studies of exclusionary zoning as a national phenomenon documented how coalitions of residents and developers worked together to develop the zoning power, though admitted a case arguing regulations excluded a particular race would be difficult (Babcock, 1966). Following the logic of the Supreme Court's 1926 *Euclid v. Ambler* ruling, land use regulations were justified by stating they protected "health, safety and general welfare." During my analysis period, state courts interpreted "general welfare" to include desires to preserve property values or neighborhood character [4]

American suburbanization after 1940 is also coupled with large internal migration like the Second Great Migration. While migration up to 1940 formed Black communities in the urban core, Gregory (2005) estimates Southern Black migration to the North from 1940–70 doubled flows during the preceding 1910–40 period. On one hand, contact between Black Americans and liberal-minded Americans in the North would greatly increase and contribute to the advancement of Civil Rights (Calderon, Fouka and Tabellini, 2023). However, the degree of white flight out of Northern cities suggests backlash to social change was not limited to the South (Boustan (2010), Sahn (2022)).

Interest groups who could influence the adoption and modification of zoning ordinances exercised collective action to limit denser, more affordable housing. [5] If the letter of zoning regulations were race-neutral, occasionally the interest groups who supported them would express ulterior motives. A case in point followed the 1948 *Shelley v. Kraemer* decision, where the Supreme Court rendered unconstitutional racial covenants that only permit sales to homebuyers of "the Caucasian race." A statement by the president of the Los Angeles realtors' association was uncompromising:

> The unfortunate feature of the situation is that those who suffer most are the owners of comparatively modest homes. Such home owners as a general rule have practically all of their wealth invested in their homes. Moreover, the prices of

[3] Work along these lines include Sood and Ehrman-Solberg (2023), who find persistent racial sorting around homes under a racial covenant, banning sales to non-white homeowners. Albright et al. (2021) study the decline in Black economic outcomes after the 1921 Tulsa Race Massacre.

[4] Infranca (2023) notes that such acceptance started developing in the 1930s, and that the Supreme Court affirmed "broad and inclusive conceptions of public welfare" in 1954.

[5] Among existing case studies, Freund (2007) and Reynolds (2019) study organized resistance by suburban residents to strike down development at city council meetings. Rothstein (2017) details how residents and local officials used regulation to stall racially integrated housing in the Bay Area. Gordon (2008) studies exclusionary zoning by St. Louis suburban planners and their attempts to defend their motives in judicial hearings.

> homes in such areas are well within the purchasing power of Negroes . . . The magnitude of the economic and social loss with which we are confronted is appalling. (McMichael, 1949)

Unlike other racially motivated collective action, lot size regulations survived legal challenges and remain a tool for local governance.[6] Neighborhoods constrained by minimum lot sizes are locked into a built environment that is costly to change, even if current residents lack racial animus[7]. I show these lock-in effects have persistent and consequences for residential sorting into the suburbs and disparities in social mobility within cities.[8]

Existing studies that use cross-sectional evidence from recent times to study the causes of local regulations face a fundamental identification issue: Local amenities, demographics and regulations jointly evolve over time, so their contributions to regulation adoption cannot be separately identified without further assumptions (Bogart, 1993). [9] I provide a novel panel of lot size controls for governments across the country. [10] Panel variation in land use regulations opens up design-based inference that use natural experiments to estimate their determinants. Krimmel (2022) is the only other work I know taking this approach, where he estimates the adoption of residential land use controls before and after California equalized school funding in the 1970s.

In recent work, Song (2021) uses zoning codes reported in administrative data to detect binding minimum lot sizes and estimate market premia of lot size regulations. While we detect and measure minimum lot sizes using different methods, I replicate the aggregate outcomes we share in our work. I use panel variation in regulatory outcomes and historical quasi-experiments to more credibly estimate which historical determinants mattered.

[6] In 1975, the Supreme Court ruled in *Warth v. Seldin* that questions on whether land use regulations are exclusionary lack standing for federal judicial remedies.

[7] Recent work by Sahn (2022) and Shi et al. (2022) also investigate the causal effects of Black migration on postwar urban planning, but on a sample of central cities rather than entire metropolitan areas. My work is complementary, showing the aggregate impact on urban form due to the Great Migration increases once I account for suburban zoning responses.

[8] My work is also related to recent research that shows changes in local jurisdictions' governance structures during the Great Migration (Grumbach, Mickey and Ziblatt (2023), McCasland et al. (2023)). Evolution in governance should be related to the evolution in land use regulations, but governance structures can be reformed later. I show at length that past land use regulations have persistent economic consequences, as restrictive density is costly to reverse.

[9] In the case of land use regulations, the problem of how amenities today evolved jointly with regulation applies to both calibrated structural estimation (Calabrese, Epple and Romano (2007), Parkhomenko (2020)) as well as design-based approaches both past and present (Rolleston (1987), Hilber and Robert-Nicoud (2013)).

[10] The only paper I know of that estimates the adoption of lot sizes from deed records is Zabel and Dalton (2011), which covers one metropolitan area in the U.S. — Boston, MA — from a period starting in the late 1980s. Existing work is therefore reduced in scope along both spatial and time dimensions.

More generally, this paper introduces estimators and statistics that use longitudinal data to recover decision rules that alter incentives — as observed through bunching — and across many decision makers. The estimation approach could also be applied to firms in a labor economics context (Cengiz et al. (2019), Goff (2023)), banks in a finance context (Collier, Ellis and Keys, 2022) or other research areas that have used bunching estimators around a known threshold.

Finally, my results inform a growing literature on urban growth and spatial misallocation, which predicts relaxing urban land use regulations could cause considerable gains in land rents (Turner, Haughwout and van der Klaauw (2014), Anagol, Ferreira and Rexer (2021)) increase real per capita incomes (Hsieh and Moretti (2019), Duranton and Puga (2019)) and reverse economic divergence within nations (Ganong and Shoag, 2017). These empirical findings are puzzling in a model where local actors compete on zoning only to maximize land rents.

I use a natural experiment to understand what motivates local actors to regulate apart from the fiscal dimension of development. I find suggestive evidence that land use regulations adopted decades ago persist today and correlates with present-day disparities. Political decisions reacting to stronger growth could therefore introduce path dependence in today's urban geography. Beyond this paper, the data I employ could further answer questions on American urban dynamics beyond the postwar decades.

The paper proceeds as follows. Section 2 overviews my algorithm, and the assumptions I make, to measure the adoption and restrictiveness of lot size regulations. Section 3 describes the joint evolution of lot size outcome measures and Black migration and what hypotheses the data can test. Section 4 discusses the design and validity of my empirical strategy, with results and additional evidence for racial motives presented in Sections 5 and 6. Section 7 concludes.

2 Measuring Bunching on Lot Sizes

To jointly estimate both regulation levels and their respective adoption dates across local government, I use the CoreLogic Tax Records, a national dataset collecting property-level information from county assessor offices. My algorithm has two steps, with the first step nested within a second model training step.

The first step conducts unsupervised learning over lot size distributions. Of all the possible lot sizes where minimum lot sizes could have been adopted, my algorithm classifies the subset of *bunching bins* of properties built under each government. I show in Section 2.3 the assumptions under which the bunching bins match the lot size regulation distorting housing market activity. In the second step, I refine my definition of "significant bunching" so minimum lot size adoption dates implied by the first step matches actual historical records.

My approach differs from the methods in Zabel and Dalton (2011) and Song (2021), where information on zoning boundaries are first gathered and one lot size in the zone is selected as the effective regulation. My algorithm detects arbitrary numbers of regulations and requires no knowledge of within-jurisdiction zoning geographies. Instead, lot sizes are identified if one assumes they induce bunching behaviour that cannot be explained by buyer or development behaviour in unregulated housing markets.

2.1 Data

From 2009 to 2019, CoreLogic has compiled property records from tax assessor offices across the United States in its Tax Records file. In my analysis, I process 86 million properties classified as single-family homes or as duplexes in 426 U.S. Core-Based Statistical Areas. The data span 1151 counties and the aggregate count of structures in those counties is almost identical to the Census Bureau's 2019 estimates of 1-2 unit structures. I outline the data cleaning process below, while I provide full details in Appendix Section C.3.

I first match each property to a *zoning jurisdiction*, defined as any incorporated city or other local government that have had zoning powers up to 2010. [11] City boundaries shift due to strategic annexation of land, so I use time-consistent boundaries that separate old borders with areas annexed between 1960 and 2010; details are in Appendix Section C.2. I define 10364 separate zoning jurisdictions in my analysis sample. [12]

Five-sixths of the CoreLogic property records have lot square footage and the year the property was built. Both variables are necessary for my algorithm; the latter variable is my proxy for the year the lot first developed. I match a property to the jurisdiction that would have administrated the lot at the year built. That means properties in land annexed by a jurisdiction as of today could be matched to two separate jurisdictions.[13]

The CoreLogic data is a cross-section of properties still standing by the 2010s. Data attrition — observing fewer properties built in a year than I would with contemporary counts — is possible due to data miscoding, but more likely due to redevelopment of housing on a parcel from the 20th Century onwards.

[11] While states can pass laws that curtail jurisdictions' zoning powers, until 1970 there is not much evidence states deviated much from the framework suggested in the 1924 State Standard Zoning Enabling Act. Any state-level reform from 1970-2010 focused on adding review of comprehensive plans that the zoning resolution should enable.

[12] Appendix Section C.1 describes the precise construction of the jurisdictions by state.

[13] Exceptions are properties under jurisdiction of counties or township governments, whose boundaries changed little in the last Century.

To check attrition, I benchmark the CoreLogic records to historical Census estimates of properties built by the end of each decade. I find enough housing from the past endured to the present to conduct analysis. My sample contains 55% of all single-family homes that were built as of 1940 and 86% of all single-family homes that were ever built between 1940 and 1960. Appendix Section C.4 elaborates on the benchmarking procedure.

2.2 Inner Loop: Detecting Bunching Bins

For some definition of "significant bunching," the inner loop recovers lot sizes on which many properties repeatedly bunch across time. My approach first fixes a jurisdiction j and a subset of *adoption times* τ over all observed years built in the data:

Definition 1. *A* ***housing vintage*** *in jurisdiction j, indexed as (j, τ, T) and denoted as $h^j_{\tau,\tau+T}$, is the empirical lot size distribution for homes built in j between years τ and $\tau + T$.*

I discretize $h^j_{\tau,\tau+T}$ into fixed *lot bins* $\{\ell\}$. To see if minimum lot sizes were adopted at τ, the algorithm first checks if any bunching on the distribution is observed between τ and $\tau + T$. The algorithm makes a positive classification based on the following threshold rule:

$$\log \frac{h^j_{\tau,\tau+T}(\ell)}{h^j_{\tau,\tau+T}(\ell-\mu)} - \log \frac{h^j_{\tau',\tau'+T'}(\ell)}{h^j_{\tau',\tau'+T'}(\ell-\mu)} \equiv G^j_\tau(\ell) \geq \alpha,$$

where $\mu > 0$ and $\tau \geq \tau' + T'$. I call $G^j_\tau(\ell)$ a *gradient statistic* aimed to detect a set of *bunching bins* for a vintage (j, τ).

Intuitively, for each ℓ the statistic measures whether the distribution looks like what happens when land is unevenly regulated by minimum lot sizes. Where a lot size regulation binds, there could be *excess mass* on a specific size compared to marginally larger and smaller values. Furthermore, *missing mass* marked by a drop in density for marginally smaller lot sizes reflects substitution to the binding minimum. The more positive this statistic is, the more likely it was caused by a regulation instead of statistical noise. Details of the implemented algorithm are in Appendix Section D.1.

The three panels of Figure 2 visually demonstrate how to construct the gradient statistics G^j_τ with a worked example: lot sizes from the Philadelphia suburb of Lower Merion Township, PA. The algorithm seeks to recover five regulations adopted by the jurisdiction in 1939: minimum lot sizes per unit were set at 5,000, 6,000, 10,000, 18,000 and 30,000 square feet. [14]

[14]These values were recovered from contemporary ordinances and planning documents. The documents themselves highlight some objectives behind adopting lot sizes and are discussed further in Appendix Section C.8.

In Panel (a), I show the histogram of properties after the adoption time $\tau = 1940$. I highlight the 5,000 and 30,000 square feet lot bins in orange, and highlight in blue a bunching region of measure μ to each bin's left. The 30,000 square feet bin has noticeable excess mass, but the 5,000 square feet bin still has missing mass to its left. The log difference of the two densities form the first terms of $G^j_{1940}(\ell)$.

Panel (b) analogously shows statistic calculations over properties built a decade before, the 1930s. If minimum lots were not adopted before 1940, the second term of G^j_{1940} calculated over the older vintage only captures gradient change due to heterogeneous demand, not due to regulations. The gradient statistic is evaluated over both samples and then takes their relative differences.

In Panel (c), I show how I classify the bunching bins. I plot the full gradient statistic for Lower Merion, further standardized by a standard deviation calculated across jurisdictions in its metropolitan area, $\sigma[\{G^j_\tau(\ell)\}_{\ell,j}]$. The median statistic value is close to 0; a few outlier statistics, including the two evaluated at the actual minimum lot sizes of 5,000 and 30,000 square feet, exceed a critical value α. All lot sizes with $G^j_{1940} > \alpha$ are classified to have bunching starting in 1940.

Iterating over the set of vintages (j, τ), the algorithm outputs a collection of bunching bins for each vintage, $\underline{\mathbf{b}}(j, \tau)$. Because the gradient statistic is undefined if no lots were built at values $\ell - \mu$, the algorithm has alternative classification rules accomodating these edge cases. In summary:

Definition 2. *A lot bin ℓ is a* ***bunching bin for a vintage*** (j, τ)*, denoted $\underline{b} \in \underline{\mathbf{b}}(j, \tau)$, if*

1. *ℓ is rejected by the gradient statistic decision rule at threshold level α:*
2. *Either a "minimum observable lot bin" or a "modal value at the left tail" rule holds at ℓ, with both defined in detail in Appendix Section D.1.*

I also enforce a *persistence assumption*, that over multiple time periods market behavior cannot generate the same levels of gradient statistics that lot size regulations do. It is then valid to only keep bunching bins detected across multiple vintages in time: [15]

Definition 3. *A lot bin ℓ is a* ***bunching bin for a jurisdiction****, denoted $\underline{b} \in \underline{\mathbf{b}}(j)$, if*

1. *ℓ was classified as a bunching bin over multiple disjoint vintages indexed by τ;*
2. *In at least one τ, ℓ was classified according to the gradient statistic decision rule.*

[15] A similar condition is used by Song (2021), where the same lot size should be detected for postwar homes and modern homes.

I calculate statistics derived from the $\underline{\mathbf{b}}(j)$ and their characteristics — like the measure of the adoption year for minimum lot sizes — before exiting the inner loop.

2.3 What Bunching on Lot Sizes Measures

Bunching estimation techniques (Kleven (2016), Bertanha et al. (2023)) can be applied to lot size distributions over housing vintages, but it is worth asking how greater measures of bunching reflect more or less exclusionary zoning. This section formalizes the connection using a stylised model of housing development on land regulated by minimum lot sizes.

Consider a jurisdiction divided up into a continuum of parcels p of equal size. each of which can be split up into lot sizes ℓ and developed into residential buildings. I model the decision of a land owner, choosing between selling off the parcel to the highest bidder or keeping ownership.

Housing developers bid competitively for the parcel. If they acquire the parcel, they split up the parcel into lot sizes ℓ and develop residential floor space s on each lot. They act as monopolistic competitors and set house prices according to a demand curve downward sloping in residential density, $n = 1/s$. In this stylised model, I make a parametric assumption that willingness to pay is $P_p(n) = a_p n^{B_p - 1}, B_p \in (0,1)$.[16]

Bids for the parcel are differentiated based on the proposed (ℓ, s), and the land owner captures all profits after development. Profits are net of construction costs common to all developers at $C = n^{\delta}$ per unit of land, $\delta > 1$. The land owner accepts the bid providing the highest profits, as long as it exceeds a reservation value $\underline{R}$ the owner has for continued ownership.

In addition, the parcel may be subject to a minimum lot size per housing unit, $\underline{\ell}$. Given the model assumptions, reasonable levels of $\underline{\ell}$ map to an explicit cap on residential density $\underline{n}$. [17] Denoting L as the developer's proposed building footprint, the owner of parcel p maximizes

$$\max_{n} \quad \left\{ (a_p n^{B_p - 1}) n \cdot L - CL, \underline{R}L \right\}.$$

Factoring out building footprint L and taking first-order conditions yield the revenue-maximizing development density, $n^*(P_p) = \left(\frac{a_p B_p}{\delta}\right)^{1/(\delta - B_p)}$, which is observed if profits per land unit exceeds the reservation value.

[16] Since this model uses a stylized constant-elasticity demand function, the restriction on B_p guarantees there is no inelastic region where no monopolist would price its products.

[17] While a developer could comply with a nontrivial $\underline{\ell}$ by developing floor space in a tall building surrounded by open space, the convexity of construction costs means such bids are less profitable than a bid subdividing the parcel into single-family lots. This argument establishes the connection $n = 1/\ell$, though in practice minimum lot sizes are combined with use restrictions on single-family housing and height restrictions.

Denote three threshold conditions as below, setting $\eta = \frac{B_p}{\delta - B_p}$:

$$a_p^{1+\eta}\left[\left(\frac{B_p}{\delta}\right)^{\eta} - \left(\frac{B_p}{\delta}\right)^{1+\eta}\right] > \underline{R} \tag{1a}$$

$$\underline{\ell}^{B_p - \delta} \geq \frac{a_p B_p}{\delta} \tag{1b}$$

$$a_p \underline{\ell}^{-B_p} - \underline{\ell}^{-\delta} > \underline{R}; \tag{1c}$$

Parcel owners' behavioral responses to minimum lot sizes are therefore defined using the conditions:

$$\ell^*(P_p; \underline{\ell}) = \begin{cases} \left(\frac{a_p B_p}{\delta}\right)^{1/(B_p - \delta)} & \text{if (1a) and (1b)} \\ \underline{\ell} & \text{if (1c) and not (1b)} \\ 0 & \text{otherwise.} \end{cases}$$

The observed share of properties bunching on lot sizes thus should be interpreted as the measure of parcels where condition (1b) does not hold, but condition (1c) does. My estimand of interest is the part of this bunching share made up of owners of "marginal parcels", whose ideal ℓ^* was smaller than $\underline{\ell}$. Such owners still sell parcels, but regulations force an *intensive margin* adjustment on them to build larger units, while aggregate land rents are lower than they would if $\underline{\ell}$ did not bind. Even if we cannot observe counterfactual ℓ^* for individual parcels, a drop in ℓ^* that are marginally smaller than $\underline{\ell}$ relative to counterfactual estimates is the missing mass I use to help detect regulations.

Another way in which owners could respond is on the *extensive margin*, when the profitability condition (1a) holds but condition (1c) fails. In that case, $\underline{\ell}$ is so restrictive that it is no longer individually rational to sell the regulated parcel, so when $\underline{\ell}$ binds the parcel has zero development: the willingness to pay for the minimum lot size is below the average total costs of developing the parcel.

In practice, jurisdictions adopt multiple minimum lot sizes for different neighborhoods across space. To detect all such regulations with observable consequences, the gradient statistic method use three strategies in Bertanha et al. (2023) for accurate estimation in bunching designs. First, the statistic uses only local information without any functional form assumptions. Second, the decision rule verifies if the distribution fails a regularity condition that should hold in the absence of bunching incentives. Finally, the statistic uses a control vintage to flexibly impute counterfactual values for the gradient.

Formally, set identification of lot sizes that induce marginal responses is possible with minimal restrictions on housing demand, but a regularity condition is needed and the pre-period control vintage must be sufficiently close to the counterfactual:

Definition 4. *The **gradient statistic classifier** $\hat{G}$ of threshold α is defined over two distributions $p(z), q(z)$ and a function μ, using the operator $\Delta(m,\mu) = \log(m(x)/m(x-\mu))$ and the rule:*

$$\hat{G}(z^*) = 1 \quad \textit{if} \;\; \Delta(p(z^*), \mu(z^*)) - \Delta(q(z^*), \mu(z^*)) \geq \alpha.$$

Definition 5. *A family of distributions measured over time $\{p^\tau\}$ has **bounded demand adjustment** of degree K if for each ℓ and $\tau \neq \tau'$, $\log p^\tau(\ell) - \log p^{\tau'}(\ell') \leq K|\tau - \tau'|$.*

Proposition ID 1. *Let a jurisdiction have lot size controls $\underline{\ell} \in \underline{\mathbf{L}}$ that induce marginal bunching and were adopted over time. There is a pair $(\overline{K}, \underline{\delta})$ such that if the jurisdiction has bounded demand adjustment of degree $K \leq \overline{K}$, there is a gradient statistic classifier with threshold $\alpha(K)$ that set identifies $\underline{\mathbf{L}}$ with probability $1 - \underline{\delta}$.*

Proof. See Appendix E. □

Importantly, not every minimum lot size set out by a jurisdiction leads to bunching on lot sizes. Bunching measures could underestimate jurisdictions' zoning restrictiveness because it does not increase when parcels are undeveloped due to extensive margin effects.[18]

In the other case, a lot size applied only to parcels where condition (1b) holds causes no observable bunching. (1b) holds when the parcel has low amenity value a_p or when B_p is small, so the price elasticity of density is more elastic. The reasoning in this section, however, suggest such even if those regulations are written down, they are no more "exclusionary" on housing supply than market forces. The regulations also do not force pecuniary losses on land owners, compared to bunching inducing regulations that do so in the short run.

A pre-period control vintage being very close to the true counterfactual for the post-period vintage means demand adjustment is bounded by a small K, preventing the threshold $\alpha(K)$ from blowing up as δ decreases. The additional persistence assumption is a heuristic used when I think K is not small, so set identification happens with low probability $\tilde{\delta}$. If a lot size is classified as a bunching N times rather than 1, the odds of the lot size being a false positive shrinks on the order of $\tilde{\delta}^N$.

[18]A concrete example of an unobservable regulation is the "holding zone" (Babcock and Bosselman, 1973). Vacant land is zoned requiring multiple acres of land per unit. No developer finds it economically feasible to build at those densities, but the jurisdiction lowers it once it finds a specific development project it likes.

The persistence assumption also rules out certain kinds of market behavior. For example, a large-scale subdivision development, where many properties have the same lot size, will not be detected as regulation induced if all units are in one vintage. But if development is "convex" — the developer takes multiple decades to build the full subdivision and is ramping up production throughout — this strategy could be falsely detected as regulation.

2.4 Outer Loop: Training Algorithm with Historical Regulations

In the outer loop, I discipline the parameters that define "significant bunching" by training algorithm classifications to match real historical regulations. The arguments in the previous Section already point out one bias-variance tradeoff: in practical situations where α and K are fixed, the gradient statistic classifier must balance having low power to detect regulations inducing small amounts of bunching or identify the entire set with high error rates. Another bias-variance tradeoff comes from recovering the first adoption date. The threshold α that can detect bunching bins over all possible vintages may be too stringent to detect bunching precisely at the year of adoption.

I employ machine learning models in order to simultaneously refine the parameters for the classifier, as well as further impute statistics like first adoption date to match the data. To train the models, I collect a national sample of past zoning ordinances and reports that span over 380 zoning jurisdictions in 13 states from every Census region. This sample is parsed from both panels of zoning ordinances over time and zoning studies in the 1960s and 1970s, further elaborated in Appendix Section C.5. The lot size outcomes I use for training is existence and adoption timing of highly restrictive lot sizes over 10,000 square feet, which varies in the cross-section more than adoption of any minimum lot size.

To use Lower Merion as an illustration of the training process, I start from the set of bunching bins $\underline{\mathbf{b}}_j(\mathbb{B})$ that vary based on the parameters $\mathbb{B}$ governing the classifier. To each $\underline{\mathbf{b}}$, there is an oldest vintage where the classifer first identifies bunching. In Lower Merion's case, most bunching bins are first detected over the 1940-1949 vintage, with one or two detected later.

I then calculate excess mass relative to a counterfactual density at the bunching bins, [19] not only for the 1940-1949 vintage but for the 1941-1950 and 1939-1948 vintages: I shift the starting years for the vintages to cover the support of years built. I also calculate these statistics where excess mass at bunching bins are counted only for vintages after the first detection of

[19]While the literature tended to use global polynomial methods to get at the density estimate, I follow the suggestions of Bertanha et al. (2023) to use local information. The excess mass is the observed share less a kernel density estimate at the point, exploiting only lot sizes outside a window around the full set of bunching bins.

that bin, as well as where all bins are counted starting from a common adoption date. Details on how this estimate is implemented is in Appendix Section D.1.

This process produces a matrix of features that measure nonlinearities in how excess mass in the lot size density around bunching bins changes over time, under least restrictive or more restrictve assumptions on whether lot sizes were passed over time or all at once. I pass this matrix through structural break models (Bai and Perron, 1998) to get imputed adoption years $\hat{\mathbf{t}}^{1st}$. To aggregate imputed information over different assumptions, my final imputed year is a reweighted ensemble measurement denoted as while allowing $\hat{T}^{1st} = \mathbf{w} \cdot \hat{\mathbf{t}}^{1st}$. Details on modeling imputation are in Appendix Section D.1.

The machine learning problem is to find the vector $\{\hat{\mathbb{B}}, \hat{\mathbf{w}}\}$ that minimizes least squares error between recorded large lot adoption years and predicted large lot adoption years. Table 1 shows performance on two datasets: a 60% sample used for training and a 40% sample used as an out-of-sample test set.

The first row of Panel (a) shows performance on the only outcome on which the model is fit: large lot adoption year in the training set. Moments indicate the classifier does produce unbiased estimates, at the expense of higher noise.[20] The MAE statistics when weighed for population are about 10 years. To the extent that many observations in the training set are small cities or townships, it reflects that first appearance of bunching is still a noisy proxy for first adoption.

The following three rows show predictive performance on other measures of lot size adoption. The estimates' central moments are slightly negative, meaning estimated adoption is a few years after actual adoption. In the test set, which contains more major cities, the algorithm performs better and lead to a smaller population weighted MAE. Over all cities in the data, 63 out of 123, or about half, of all estimated adoption years are within 10 years of historically recorded adoption.

Panel (c) shows a further test out of sample, using lot size control level variation rather than jurisdiction variation. For each actual lot size control, I match to it all bunching bins estimated to have been adopted within 10 years of the date in records. 587 out of 792 lot sizes, or 74%, successfully match under this criterion. I then look at the matched bunching bin which is closest in magnitude to the actual lot size control, and like Song (2021) compare the percentage difference between them.

Across the matched lot sizes, the procedure has a median error of 0. 310 out of 587, or 53%, of actual lot size controls have an estimated bunching bin within 25% of the value. Conditional

[20] The false negative rate – actual large lots going undetected and dropping out of the training set – was around 13%.

on the lot sizes being at most half an acre (21,780 square feet), the statistic is 271 out of 444, or 61%. Along with Figure D.1, which plots a histogram of these variables, it shows that the negative mean and MAE levels are due to larger lot sizes in the ordinances that go undetected with bunching methods — the "holding zones" referenced in Section 2.3.

3 Postwar Exclusionary Zoning: Stylized Facts

3.1 Lot Size Regulation Intensity Over The Last Century

To produce the outcomes of interest, my algorithm iterates over 407 Core-Based Statistical Areas (CBSAs) defined by the Office of Management and Budget. I sample homes with available CoreLogic data, built from 1925 to 2010. For 67.3 million properties across 8422 zoning jurisdictions, I produce a set of bunching bins $\underline{\mathbf{b}}(j)$, lot size regulation adoption years using the procedure in Section 2.4 and restrictiveness measures for each housing vintage built in post-1930 decades. [21]

Returning to Figure 1, I visualize 4,807 jurisdictions that have zoning powers and had at least 5,000 people by the 2010 Census. To create the orange series of lot size adoption, I bin the first adoption year for all jurisdictions with a bunching bin of at least 3,000 square feet.[22] The gray series plots adoption years conditional on detecting a bunching bin of over 7,500 square feet. More generally, over 79% of zoning jurisdictions with detected lot sizes have multiple bunching bins. The mean bin level is 3.7 with a standard deviation of 2.5. This reflects a postwar decision to stratify more and less exclusive residential neighborhoods using menus of minimum lot sizes.

I construct a continuous measure of excess mass restrictiveness, $\text{Excess}_{j,\tau}$, in two steps that address concerns with bunching estimation considered in (Bertanha et al., 2023). First, I produce a truncated lot size distribution for a jurisdiction's housing vintages, dropping properties at each estimated bunching bin $\underline{b}$ and the assumed missing mass region $\underline{b} - \mu$. The data used for excess mass estimation for each (j, τ) are the truncated pre-period vintage distributions, for homes built up to $\tau' + T' \leq \tau$. Second, instead of using a global polynomial, I prioritize local information: I estimate a counterfactual mass at each bunching bin, $\tilde{m}^j_\tau(\underline{b})$ using a kernel

[21] Table A.1 breaks the number down into four types: counties, townships, incorporated cities at constant historic borders, as well as annexed territory for cities growing beyond historic borders. While my algorithm lacks statistical power to confirm lot sizes for small suburbs and rural towns, it does detect bunching bins for jurisdictions where which more than 80% of sample CBSA residents live.

[22] Though lot size regulations below this threshold exist, they appear seldomly in zoning ordinances after World War II and their first adoption is confounded by lot subdivision standards for old homes.

density estimator. The output jurisdiction level measure is

$$\text{Excess}_{j,\tau} = \sum_{\underline{b} \in \underline{\mathbf{b}}(j)} m^j_\tau(\underline{b}) - \tilde{m}^j_\tau(\underline{b}).$$

Figure 3 plots, for each decadal vintage of housing, different measures of restrictiveness due to minimum lot sizes. Excess mass restrictiveness, highlighted in orange, sees rapid growth over the 1940–70 period with a national average of 13.5%. Over all homes built during the period, I conclude 2.71 (= 13.5% × 20.1) million postwar housing units were constrained by lot size controls. The bar plot in blue measures my preferred measure of *lot size restrictiveness*: Excess mass around bunching bins over 7,500 square feet in each housing vintage. This value is around the median lot size for all properties bunching on lot sizes, growing after 1950 to equal half of all properties bunching on lot sizes.

It is worth contrasting the magnitude of excess mass to another measure, the *bunching share*. The sum of densities at bunching bins, not net of counterfactual density, changes from 12% in the 1940-50 period to as high as 23% in the 1960-70 period. The criticism from Section 2.3 applies, where a measure that is almost a quarter of development is likely confounding growing demand for certain lot sizes with regulations that distort supply on the margin.

As suggested by Figure 3, national lot size restrictiveness persists beyond 1970. Over all homes built after 1940, the national rate is 13.9%, or 7.5 million units when multiplied by 54.2 million units built.[23] The same trend appears in the cross-section, as I regress the rank of lot size restrictiveness calculated on 1940–70 development with the rank using development since 1970. Figure B.2 plots the relationship in terms of a scatterplot and binned means. Apart from clear visual evidence of persistence, the rank-rank regression has a slope estimate of 0.72.

The persistence of exclusionary zoning, plus the added costs of redeveloping existing homes, suggest lot size restrictiveness has its own lock-in effects on urban geography. To investigate these relationships, I use tract or ZIP level Census data within each metro m to calculate how much of an outcome is within the borders of the central city $c(m)$, as defined in Section C.2, versus the "suburban share" outside those borders: $N_{m,t} = N^{c(m)}_{m,t} + N^{suburb}_{m,t}$. I use this as the outcome to estimate dynamic effects in the cross-section:

$$\frac{N^{suburb}_{m,t}}{N_{m,t}} = \sum_{\tau=1970}^{2010} \beta_\tau \text{Excess}^{suburb}_{m,40\text{–}70} + \delta_{r(m),t} + \mathbf{X}_{m,pre}\Gamma + \varepsilon_{m,t} \tag{2}$$

[23] These numbers are not far from a bunching rate of 16% calculated in Song (2021). Figure B.1 plots the property-weighted distribution of bunching bins, which has concentrated density at bins that are fractions of an acre as well as further diffuse bunching (Anagol et al. (2022)) at marginally larger lot sizes.

While the β_τ should not be interpreted as causal, I control for the mechanical relationship between the land occupied by the central city and share outcomes by adding the share of population in the central city relative to all metropolitan area residents, as of 1970. I also control for Census division fixed effects, the level of the 1970 Black population in each metro, as well as the Lutz and Sand (2023) share of undevelopable land within the metros' borders. Effects are also reported with 90% confidence intervals, clustered at the metro level.

Panel (a) of Figure 4 first looks at demographic outcomes; in particular, the rate of Black suburbanization recently studied in Bartik and Mast (2023). Those authors highlight greater Black suburbanization occurred alongside "within-Black stratification by income" within metro areas. The estimates in Panel (a) suggest a one standard deviation rise in $\text{Excess}^{suburb}_{m,40-70}$ is associated with a 0.1 lower standard deviation effect over time, or 2.6 percentage points fewer Black residents suburbanize. Lot size restricted suburbs are also where within-Black stratification is more visible. I find lot size restricted suburbs are associated with more suburban sorting of college educated households, a proxy for for high earning households. For college-educated Black households, lot size restrictiveness has insignificant effects on their suburbanization.

I also examine the diversity of suburban land use when they are lot size restricted: the share of multifamily units that are suburban in the decennial Census and the intensity of commercial uses, measured by the share of metro area jobs that are suburban in the 2010 ZIP Code Business Patterns.[24] Negative effects on both outcomes suggest lot size restricted suburbs have lower land use diversity. While $\beta_\tau = 0$ cannot be rejected for suburban apartment share except in 1990, statistical noise could come from apartment developers in some metros leapfrogging exclusionary jurisdictions to build further out — but in areas with lower job access.

Finally, I explore how metros with more postwar lot size restrictiveness have greater social mobility disparities between economically productive central cities and their mostly residential suburbs. I aggregate tract-level mobility estimates from Chetty et al. (2018) to higher geographies using inverse-variance weighting: $\tilde{y}_{m,geo} = \sum_{t \text{in geo of m}} \frac{1}{\sigma_t^2} \overline{y}_t$. With mobility estimates at the central city and suburban level for each m, I run the saturated regression

$$\begin{aligned} \tilde{y}_{m,geo} = \delta_{r(m),t} + \beta_1 \text{Excess}^{suburb}_{m,40-70} + \beta_2 \mathbb{1}[geo = \text{Suburb}] \\ + \beta^{opp} \text{Excess}^{suburb}_{m,40-70} \times \mathbb{1}[geo = \text{Suburb}] + \mathbf{X}_{m,pre}\Gamma + \varepsilon_{m,geo,t}. \end{aligned} \tag{3}$$

Panel (b) of Figure 4 plots the β^{opp} estimates with 90% confidence intervals. For children whose

[24] I calculate my job share variable similarly to Kneebone (2009) with 2010 data, defining "suburban" as "located within 10 to 35 miles of the central city." However, she allows for multiple principal cities in a metro while I impose a single central city.

parents are at the bottom quartile of the national income distribution, the mobility gap by adulthood if they were raised in the central city versus the suburbs is 3.5 percentiles on average. When the suburbs have a one standard deviation rise in postwar lot size restrictiveness, the gap widens by 0.15 standard deviations — or about 20 percent of the baseline gap. A similar gap in elevated incarceration risk in the central city grows by 0.22 SD, or 25 percent of the baseline gap of 1 percentage point, when the suburbs are more lot size restricted.

These associations cannot be explained entirely through the racial sorting channel alone. Additional regressions show lot size restrictiveness has null effects for poorer white children. Unconditional effects are driven by poorer Black children, and especially poorer Black men. These results echo findings in Derenoncourt (2022), who finds for the same Black male cohort an economic mobility penalty where more postwar Black migrants moved. To the extent I find suburban lot size restrictiveness is also caused by the Great Migration, the collected results hint at how the spread of exclusionary zoning spills over to lower social mobility for families who cannot afford suburban living.

3.2 Connections to theoretical frameworks

Given the evidence that lot size outcomes are widespread and have persistent impacts, I return to theories on such regulations' long-run benefits. *Fiscal zoning* theories in the economics literature predicts land use regulations will be adopted as a mechanism that charges a "head tax" to local public goods. Individually rational zoning, when taken together, produces a menu of entry fees over different public goods. This allocation is efficient because residents Tiebout sort based on their demand for public goods, and each locality prices out all free riders (Hamilton (1975), Calabrese, Epple and Romano (2012), Fischel (2015)).

Outside of economics, scholars argue instead local actors assumed above all that postwar suburbanites placed less value on racially integrated neighborhoods. Limiting smaller homes was a means to the end of preserving racially homogenous communities. (Trounstine (2020), Winling and Michney (2021)).

An ideal way to distinguish these alternative narratives with fiscal zoning theories is to conduct the following "marginal outcome test" across cities:

Hypothesis: Denote the average treatment effect on the treated of demographic change in group X on land use controls as ATT $[X]$. Consider demographic change in Black Americans $Black$ with average income level $\overline{y}^{Black}$, then with a subsample of white Americans with the same average income, $Wht \mid \overline{y}^{Black}$. Then

$$\text{ATT}\left[Black\right] - \text{ATT}\left[Wht \mid \overline{y}^{Black}\right] > 0.$$

If income is entirely controlled for, new Black residents are just as able to pay for public goods as new nonblack residents. A *race-neutral* fiscal zoning theory, in this case, could not explain racially differentiated effects. An alternative theory can rationalize racial exclusionary zoning if Black households on the margin of suburban homeownership systematically consume more public goods — or, more plausibly, white incumbents in the suburbs had biased beliefs regarding Black households' schooling needs [25]

Models of race-conscious zoning may also rationalize collective action, where few suburban jurisdiction deviate to allow dense development and capture land rents lost due to distortionary lot sizes. Following arguments developed since Schelling (1971), weak preferences for neighbors of the same race can trigger tipping points and white flight. Even if racial homophily decreases over time, a jurisdiction that deviates from restrictive lot sizes could lose more from fiscal losses related to white flight than gains from dense development. These predictions are a function of incumbent preferences more than of any pivotal interest group.

I do not claim this paper shows results that point identify $\text{LATE}\left[Wht \mid \overline{y}^{Black}\right]$. I first estimate $\text{LATE}\left[Black\right]$, then present effects of lower-class White migration that are much smaller in magnitude. In Section 6, I also use policy variation in boundaries on segregating public goods to offer more direct evidence for race-conscious preferences on local public goods. A collection of causal estimates, put together, offer a case that postwar local actors were in part race conscious when they set up new land use regulations.

4 Empirical Strategy

While early Black population growth in non-Southern cities was driven by migration, Southern migrants did not choose cities at random. Cities that attract more migrants may share multiple factors that also relate to faster suburban growth. This section discusses confounding factors and describes an instrument predicting postwar migration off of exogenous Southern shocks.[26]

[25] For the public good of schooling, Reber (2011) offers evidence on school finance after desegregation. The political outcome of desegregation was additional funding towards integrated districts — closing the spending per pupil gap between majority white districts and integrated districts, but not disproportionately financing integrated districts.

[26] For the remainder of this paper, "Southern" refers to the 14 states in the Census Bureau's South region that are not Delaware, Maryland, or the District of Columbia.

4.1 Variable definitions

My empirical strategy requires panel variation in lot size control outcomes, observed demographic change in destination cities and Southern migration statistics needed to construct instrumental variables. I define *lot size adoption as of decade t* as a dummy variable: if during the time interval $[t-5, t+5)$, the jurisdiction is inferred to adopt lot size or had already adopted it.[27] A jurisdiction adopts a *restrictive lot size* when it is also inferred to have a bunching bin above 7,500 square feet, conforming with my preferred measure of *lot size restrictiveness during decade t*. A jurisdiction in the sample is any zoning jurisdiction that has incorporated by 1980, fixed at the earliest historical boundaries recorded in my data.

Demographic change in destination cities are calculated with decadal Census records, further described in Appendix Section C.6. For each time-consistent central city of a metro, $c(m)$, the identifying variation is central city Black composition change, the difference of Black population shares across decades:

$$\Delta CC_{c,t}^{Black} = \frac{\text{Black}_{c,t}}{\text{Pop}_{c,t}} - \frac{\text{Black}_{c,t-1}}{\text{Pop}_{c,t-1}} \equiv s_{c,t}^{b} - s_{c,t-1}^{b},$$

I calculate ΔCC^{Black} using city and tract-level data, so all calculations follow fixed central city boundaries based on approximations of their 1960 borders. [28]

Panel (a) of Table 2 shows counts and summary statistics for outcomes, both conditional and unconditional on square footage threshold. Panel (b) shows for 228 of 236 non-Southern CBSAs with lot size outcomes, their central cities have demographic change data from 1940–70. About 5,000 suburbs — jurisdictions in a metropolitan area but is not a central city — remain in the analysis sample. The average postwar housing vintage in each jurisdiction has 700 homes, with 5.9% of those homes constrained by lot size regulations.[29] Before the war the average central city had 2% of their households headed by Black Americans, which would grow by 2 percentage points each following decade.

Panel (c) looks at 1940 demographic characteristics for sample jurisdictions, constructed

[27] In calculating adoption, I assume no jurisdictions repeal their minimum lot sizes in full, as well as repeal all large minimum lot sizes in the large lot adoption time series. It is unlikely this assumption significantly biases the time series, given such a reversion is not seen in recent surveys (Gyourko, Hartley and Krimmel (2021)).

[28] ΔCC^{Black} is defined as in Calderon, Fouka and Tabellini (2023) and McCasland et al. (2023). It also relates to the right-hand side variable in Derenoncourt (2022), where it is described as the "1940–70 increase in urban Black population, as a share of the 1940 urban population." Conceptually, the statistic is a product of a Black population growth ratio and of a predetermined share $s_{c,1940}^{b}$. Compared to the Derenoncourt variable, a difference in shares assigns lower values to fast-growing cities in the West and higher values to older cities in the Northeast. A similar statistic to Derenoncourt's is used in Section 5.3 for non-Black migration.

[29] To ensure bunching is not estimated over miniscule vintages before suburbanization reached some jurisdictions, I drop vintages of 30 or fewer properties when looking at lot size restrictiveness outcomes.

from full-count Census data using the Census Place Project crosswalk detailed in Section C.7. Not unlike their central cities, the average suburban zoning jurisdiction had only 1.6% Black residents on average.[30] While suburbs vary little in their Black household share, standard errors show they do vary more in population, homeownership, proximity to the central city and average socioeconomic status.

4.2 Main Specification and Instrument Construction

To estimate causal effects of Black demographic change on land use regulation, I begin with the following linear model on a national panel of suburban jurisdictions j:

$$Reg_{jt} = \beta \Delta CC^{Black}_{c(j),t} + \delta_{r(j),t} + \mathbf{X}_{j,pre}\Gamma + \varepsilon_{j,c(j)t}, \tag{4}$$

where Reg_{jt} are lot size regulation outcomes defined in Section 4.1, $\delta_{r(j),t}$ are Census region–decade fixed effects and $\mathbf{X}_{j,pre}$ are 1940 jurisdiction level characteristics.

Whether the outcome is lot size restrictiveness or adoption, a positive effect means more Black residents in the central city causes suburbs to maintain exclusionary zoning, even when the market can supply dense housing. I note that because jurisdictions do not drop out of the analysis sample after adoption, regressions on adoption with a constant effect across time will estimate more attenuated effects than a model broken out by decade.

Figure 5 visualizes this initial design. Panel (a) maps Northern and West Coast metropolitan areas by their cumulative level of postwar ΔCC^{Black}. Black migration was a fat-tailed phenomenon, with rapid demographic change in the most populous cities and some regional cities. Panel (b) plots a metropolitan level association between central cities' Black demographic change and their suburbs' greater lot size restrictiveness, for 150 metropolitan areas where the algorithm covers at least 10 suburban jurisdictions. There is a visibly positive association with the identifying variation looking across the whole national sample, with blue circles highlighting metropolitan areas near the line of best fit: Seattle, New York or Cleveland. These are not cities covered well with existing historical case studies: my national data can quantify effects that are suggested by case studies of individual cities.

The exogeneity condition, $\mathbb{E}\left[\varepsilon_{j,c(j)t} \middle| \Delta CC^{Black}, \delta_{r(j),t}, \mathbf{X}\right] = 0$, implies no confounding explanations for why some metropolitan area would have both more Black migration to their central cities and suburbs planned a certain way. As standard theory predicts local land use regulation

[30] 83% of sampled prewar suburbanites are native-born whites, so suburbanization had not been widespread for either Black or foreign-born households.

varies according to location demand and incumbents' interests, exogeneity also means urban areas receiving more Black migrants cannot differ in marginal suburban household demand, or in how suburban incumbents interact with the marginal resident.

One of several possible violations of the first condition is if Black migrants disproportionately move to the Northeast or to the West, between which the costs of building on large lots vary due to water availability (Burchfield et al., 2006). The second condition would be violated if cities that offer more economic opportunity for Black migrants also share distinct suburbanization trends. Possibilities include wealthy residents suburbanizing earlier; greater investment in transportation infrastructure that expands the urban area; or early zoning excluding on income causing Black migrants to remain in dense neighbourhoods, a case of reverse causality.

Depending on which theoretical mechanisms dominate, selection into suburbs or reverse causality concerns can bias estimates upwards or downwards. OLS estimates have upward bias if incumbents selecting into suburbs react against growth, such as wealthy incumbents unwilling to share governance with any lower income households and opposing development in advance. OLS has downward bias if those incumbents support growth. For example, if there had already been sorting to the suburbs in reaction to prewar Black demographic change (Shertzer and Walsh, 2019), interest groups in those suburbs may encourage housing growth and having more white neighbors.

Following Boustan (2010) and Derenoncourt (2022), I instrument for Black migration using a shift-share formula. Migration to a central city is predicted putting weight on counties where migrants came from before the war. I use the 1940 full-count Census, the first Census that asks for previous migration, to tabulate the county Black residents in non-Southern central cities lived in in 1935.[31] Because push factors due to the structural transformation of the rural Southern economy can explain rises in outmigration rates, outmigration rates from those counties serve as the identifying variation to predict counterfactual migration.

Across all non-Southern central cities c and potential sending Southern counties k, I define the shift-share instrument for Black migration Z_{ct}^{Black} as:

$$Z_{ct}^{Black} = \sum_{\text{Southern } k} \tilde{\omega}_{c,1940}(k) \times g_k(t), \qquad g_k(t) = \frac{\overrightarrow{\text{Black}}_{k,t}}{\overrightarrow{\text{Black}}_{k,1940}}. \tag{5}$$

$\tilde{\omega}_{c,1940}(k)$ is the share of all 1940 migrants into city c who came from a sending Southern county k and was tabulated in the 1940 Census. The migration "shock" $g_k(t)$ is the ratio of

[31] I count 309 thousand Black Americans in total that moved between county borders, a third of whom moved from the South to the North.

Black migrants leaving a Southern county k in decade t to the count leaving k for the North in 1940. I calculate the shock numerator from additional postwar Census records compiled by Boustan (2016) and used in Derenoncourt (2022).

In the literature, the shift-share instrument is usually expressed as a linear combination of $\omega_{k,1940}(c)$, the share of migrants out of county k who move to city c, multiplied by the level of predicted migrants out of k. Appendix F.2 shows the two instruments are numerically equivalent up to a scaling factor, central city Black residents $\text{Black}_{c,1940}$.

4.3 Addressing Threats to Identification

Recent literature on shift-share instruments similar to Z_{ct}^{Black} separate cases where it is exogeneity of shares (Goldsmith-Pinkham, Sorkin and Swift (2020)) or the exogeneity of shocks (Borusyak, Hull and Jaravel (2022), hereafter BHJ) that implies the exclusion restriction. I argue the shift-share instrument identifies causal effects through exogenous shocks: the shares $\tilde{\omega}_{c,1940}(k)$ represent differential selection of Black migrants to cities before 1940, so city-level factors that drove selection may be correlated with later urban growth patterns.

In Appendix Section F.2, I state the formal requirements on shock exogeneity based on the results in Borusyak, Hull and Jaravel (2022). Intuitively, two conditions are sufficient for the shift-share IV to satisfy the exclusion restriction. First, each county-level migration rate shock $g_k(t)$ must be *idiosyncratic*, the sum of a random component and of a fixed effect at some cluster of counties. Second, as the number of observations in the sample grows, the sample size of shocks also grows instead of staying fixed.

The unadjusted shock term $g_k(t)$ may still be correlated with Northern destination cities, because durable pull factors in destination cities can induce serial correlation in migration patterns. Social networks between Black migrants can cascade high migration *levels* into even higher levels tomorrow (Stuart and Taylor, 2021). Measurement error in the shock denominator could also lead to small rural counties having outlier growth in the shock term, reducing instrument relevance.

As in Derenoncourt (2022), I replace actual outmigration levels from counties with projected levels $\widehat{\text{mig rate}}_{kt}$, based on cross-sectional relationships between migration push factors at the county level and actual migration. The resulting simulated migration shock $\hat{g}_k(t) = g_k(t) \times \frac{\widehat{\text{mig rate}}_k}{\text{mig rate}_k}$ varies not on county-specific ties to Northern cities, but only on structural transformation in the South to which counties had varying exposure. I confirm the previous papers' interpretation that outmigration can be explained in part by agricultural mechanization following World War II, which rolled out at differential rates across the South. Appendix Section

F.1 describes the procedure in more detail.

To reduce influence of outlier shocks, I transform the projected $\hat{g}_k(t)$ based on its quantile in the empirical CDF of the within-decade shock distribution. The resulting shift-share instrument ranges from zero to one, i.e. a city whose migrants in 1940 were all from the county that had the most projected migration. Table A.2 offers summary statistics for the untransformed and transformed shocks, while also confirming migration is not excessively concentrated in a few county-decades that would violate identifying assumptions shown in BHJ.

The two-stage model that augments the OLS model in this section is

$$\begin{aligned} Reg_{jt} &= \beta \Delta CC^{Black}_{c(j),t} + \delta_{r(j),t} + \mathbf{X}_{j,pre}\Gamma + \varepsilon_{j,c(j)t} \\ \Delta CC^{black}_{c(j)t} &= \gamma \tilde{Z}^{Black}_{c(j)t} + \delta_{r(j),t} + \mathbf{X}_{c(j),pre}\Gamma + \nu_{c(j),t}, \end{aligned} \tag{6}$$

where the new variable $\tilde{Z}$ is the shift-share using the quantile transformed shocks. As noted in BHJ, regressions using shift-shares defined at the city level are numerically equivalent to one defined over county-level shocks, but where outcomes are transformed by shock exposure weights. This shock-level transformation, paired with standard error clustering, implements the standard error correction for shift-share IV suggested by Adão, Kolesár and Morales (2019). In my context, the transformation is $\overline{y}_k = (\sum_c \tilde{\omega}_{ck} y_c)/\sum_c \tilde{\omega}_{ck}$.

In Appendix Section F.3, I visualize the first-stage relationship between ΔCC^{Black} and $\tilde{Z}^{Black}$ at the central city–decade level, suggesting instrument relevance holds both within and across decades. When it comes to exogeneity, any significant association between the instrument and jurisdiction characteristics indicates a causal path that must be closed through added controls.

I conduct balance tests between the shift-share instrument and jurisdiction characteristics as of 1940, before the analysis period. Figure 6 visualizes those balance tests over two specifications, relative to a standard deviation rise in the percentile-transformed shock over the 1940s. Confidence intervals come out of shock-level regression, so they have correct coverage. Two significant associations stand out: a strong association between migration shocks and cities with more manufacturing, and a weaker association with cities whose suburbs have more Black residents, more professional workers and more homeownership. When the same tests are run but controlling flexibly for the share of 1940 central city workers in manufacturing, the manufacturing association becomes insignificant and demographic selection on the suburbs is attenuated.[32]

The sum of evidence suggests I include the 1940 central city manufacturing share as a

[32] Appendix Section F.3 provides further balance tests for central cities and 1930 Census outcomes.

control, plus additional suburban controls to close confounding paths and improve precision. Furthermore, because not all 1940 Black migrants from the South report the county they are from, I follow BHJ in controlling for "incomplete" county-decade shares interacted with decade fixed effects.

A final concern is that regulatory outcomes are measured separately across jurisdictions in the same metropolitan area, so the outcomes must be grouped to a level that matches that for the endogenous variable. My final model used to estimate causal effects of the Great Migration on lot size restrictions is therefore:

$$\begin{aligned} \overline{Reg}_{kt}^{\models} &= \beta \overline{\Delta CC^{Black}}_{kt}^{\perp} + \overline{\varepsilon}_{kt}^{\models} \\ \overline{\Delta CC^{Black}}_{kt}^{\perp} &= \tilde{\gamma}\hat{g}_k(t) + \xi_{kt}, \end{aligned} \tag{7}$$

where $Y^{\perp}$ is a variable defined over metropolitan areas and residualized by metro-level controls. $Y^{\models}$ refers to a variable additionally residualized by fixed effects and jurisdiction-level controls before taking the grouped mean at the metropolitan area level.[33] All standard errors are clustered at the county of origin level k and cluster-adjusted first stage $F-$statistics are around 20, around the threshold for 10% weak instruments bias recommended by Olea and Pflueger (2013).

5 Results

5.1 The Great Migration's Impact on Lot Size Restrictions

Figure 7 shows how suburban lot size outcomes respond to within-decade variation in the migration shift-share instrument. As mentioned in Section 4.3, I transform both Black composition change and the shift-share instrument into within-decade percentiles, varying from 0 to 1 for each decade. Pooled over all three decades from 1940–70, there is a clear positive relationship with predicted Black migration to central cities and lot size restrictiveness defined in Section 3.1. An analogous graph for adoption of restrictive minimum lots above 7,500 square feet is given in Figure B.3.

The slope of the relationships, reflecting the reduced form estimate of causal effects

$$Reg_{ct}^{\models} = \alpha + \tilde{\beta} Z_{ct}^{Black} + \mathbf{X}_{c,pre}\Gamma + \nu_{ct} \tag{8}$$

[33]Section 5.1 will also present models estimated directly with Equation 6 to verify grouping does not significantly alter point estimates.

imply a move from the metro with the lowest predicted Black migration to the highest one explains a rise in 1.84 percentage points of lot size restrictiveness.

Table 3 presents regression estimates of causal effects of central city Black composition change on lot size outcomes, across specifications based on the OLS model (4) and the IV model (7). My preferred causal estimates are in Column (2) of the IV models, which presents two-stage least squares estimates of β where the outcome is residualized on central city and suburban controls mentioned in Section 4.3.

Under constant effects between decades, a move from the metro with the lowest to the highest Black composition change every decade explains a 40 percentage point rise in the propensity to adopt restrictive lot size controls above 7,500 square feet. On my preferred measure of restrictiveness, the same move across metropolitan areas explains a 4.66 percentage increase in excess mass. The result on adoption can be rejected from the null at 10% significance, while the result on restrictiveness can be rejected at 5%.

To better interpret these estimates, a one standard deviation rise in Black composition change across sampled central cities is 6.9 percentage points over three decades, or a shift up 65 percentiles in each decade's distribution. Such growth then causes restrictive lot size adoption propensity to rise by 26 percentage points and the rate of lot size restrictiveness rises by 3.0 percentage points over housing built. Relative to the postwar lot size restrictiveness measure used in Section 3.1, the impact of Black composition change is 0.7 (= 3.0/4.3) standard deviations.

The OLS models exhibit downward bias relative to Column (1) of the IV model. One explanation is that the IV model corrects for reverse causality in the construction of Black composition change: less restrictive suburbs will have more population growth, pushing down central city population even if Black migration levels were flat. Column (1) is also smaller in magnitude relative to Column (2). Controlling for perceived disamenities already in place in 1940 means I adjust for suburbs of Black migration destinations being pro-growth, which would lower restrictiveness even if population growth did not cause racial change.

To calculate a nationwide "Great Migration effect" effect on the treated analysis sample, I aggregate up decadal variation to derive a cumulative effect on postwar lot size outcomes. I calculate for each decade the mean ΔCC^{Black} over non-Southern central cities, which increases from 1.3 to 2.4 percentage points a year: over three decades, the average central city sees a rise of 6 percentage points in central city Black composition. Then, the difference is converted to percentiles through simulating from bootstrap samples of decadal empirical CDFs for ΔCC^{Black}.

The Great Migration effect is measured relative to a counterfactual where Black migration does not accelerate after 1940, so the Black population share in central city stays constant and $\mathbb{E}[\Delta CC_t^{Black}] = 0$. Effects are relative to baseline outcome levels over non-Southern metros:

$$Excl_t = \frac{\hat{\beta} \times [F^{CC}(s_{\Delta CC_t^{black}}) - F^{CC}(0)]}{E[Reg_t]}$$

Panel (a) of 8 visualizes the Great Migration effect for two primary outcomes: restrictive lot size adoption and lot size restrictiveness, pooling decade-specific effects by weighing each decade by units built. I plot point estimates along with simulations assigning influence weights across CBSA-decades to derive a Bayesian bootstrap confidence interval (Rubin (1981)). Figures B.4 visualize the same for different restrictiveness measures plotted in Figure 3.

For the adoption outcome of lot sizes, I derive a point estimate of share explained at 28.5% with a 95% bootstrapped confidence interval of $[9.0\%, 51.3\%]$. The share explained of lot size restrictiveness — which measures the extent to where local governments designed lot size controls to limit development — is significantly greater at 56.6%. The 95% bootstrapped CI is $[25.7\%, 98.4\%]$.

5.2 Effect Interpretation and Dynamics

In level terms of total postwar units constrained by lot size regulations, I use estimates from Panel (c) of Table 3 that imply a Great Migration Effect of 38.5%. Migration effect expressed in shares can be scaled by all suburban units recorded in my postwar sample:

$$\begin{aligned}\text{Units Affected} &= \left(12.6M \text{ units} \times 13.4\%\right) \times 38.5\% \text{ explained by GM} \\ &= 1.68M \times 38.5\% = \text{ 650,000 units.}\end{aligned}$$

I predict without local zoning evolving in reacting to the Great Migration, 650 thousand units built during the postwar years would have been built more densely. Just how many additional units would have been provided is difficult to predict, absent a model of housing demand and location choice fit on historical prices and quantities. Instead, I consider simpler back-of-the-envelope calculations.

One counterfactual is that without lot size restrictiveness, developers would subdivide their lots to be twice as dense as what was restricted, which would lead to 650,000 more units built. In this counterfactual, developers would supply few large lot units but many more rowhouses or “missing middle” properties.

Another counterfactual is densification in neighborhoods zoned for restricted lot sizes above 7,500 square feet: the average postwar density per acre among properties bunching on lot sizes is 5.1 units per acre, but it is 3.3 units per acre filtered to just restrictive lots. An upper bound on additional units that can be provided on marginal restricted parcels can be derived from replacing actual units built with a density of 15 dwelling units per acre.[34] In that case, the housing capacity of the marginal restricted parcels is 1.3 million ($= (15/3.3-1) \times 640,000$ units $\times$ 56.6% explained) units.

It is also plausible that as some cities were disproportionate destinations of Black migration, the effect would be concentrated in cities that Black migrants moved to for economic opportunity. To take two urban areas close in space, Cleveland, Ohio had 20 percentage points higher postwar Black demographic change than Columbus, Ohio — about 1.75 standard deviations. Using the scaling in the last section, this is a 1.23 standard deviation effect on lot size restrictiveness. Results from 3.1 suggest racial exclusionary zoning would have a meaningful impact on mobility gains for families suburbanizing in Cleveland relative to those suburbanizing in Columbus.

Though I have assumed so far effects do not vary over time, the postwar period also saw gradual bans on private actors' discriminatory practices against Black homebuyers. Up to 1950, developers specialized in suburban housing with racial restrictive covenants, forbidding any individual on a street to deviate and sell to a Black homeowner (Sood and Ehrman-Solberg (2023)). Commercial banks practiced redlining beyond 1950, denying mortgages at high rates to Black renters capable of ownership (Ross and Tootell (2004)).

As mentioned in Section 1.1, the *Shelley v. Kramer* ruling banned racial restrictive covenants nationwide going into the 1950s. However, racial exclusionary zoning could function as a substitute for these covenants, provided by local governments under race-neutral language instead of by private actors. If localities practiced racial exclusionary zoning, the national ban on racial covenants should not lead to a decline in Great Migration effects.

Panel (b) of Figure 8 estimates reduced-form effects like in Specification 8 separately across each decade, then express dynamic effects in standard deviation terms like in Equation 2. For both restrictive lot size adoption and lot size restrictiveness, effects are the largest in the 1940s: a one standard deviation rise in the shift-share instrument causes about a one standard deviation change in both outcomes. In the next two decades, though, lot size adoption effects turn null and then negative. Lot size restrictiveness effects stay steady in the 1950s, then turn

[34]A degree of density that corresponds to suburban apartments or townhomes of up to 3 stories. Massachusetts' 2022 upzoning laws bans zoning for densities below this level around transportation corridors (Kulka, Sood and Chiumenti (2022)).

null in the 1960s.

The drop off in adoption effects indicate that the Great Migration accelerated restrictive lot size adoption, but by the 1960s low Black migration suburbs had caught up and adopted similar regulations. The trend is different for lot size restrictiveness, which measures the *degree* to which exclusionary zoning was binding on development. The lack of an effect size drop off before and after 1950 would be consistent with lot size restrictiveness serving as a substitute for private discrimination. With the 1960s being a decade of both major advances in civil rights as well as backlash , the drop off in the 1960s could be the result of conflicting social trends. I discuss the drop off further when estimating the contribution of one such trend— school desegregation — in Section 6.

5.3 Opposite Sign Effects with White Demographic Change

Even though they were less likely to leave the South from 1940–1970, the *level* of Southern White migration to non-Southern cities was higher than those of Black migrants. Pooling Census records by migrants' state of birth, I confirm estimates by Gregory (2005) that around 8.2 million Southern White residents left their region over my analysis period.

Estimating Southern white migrants' effects on lot size outcomes checks against two race-neutral explanations for the Great Migration effect. First, the Great Migration effect could be a reaction against neighborhood congestion caused by faster white flight from the central city. If this is the main mechanism, metro areas whose population growth was mostly driven by Southern whites should see more restrictive suburban lot size outcomes.

Second, Southern white migrants are a group not differentiated from incumbents by race but is differentiated through income. Figure B.5 contrasts household income distributions for three groups in the 1960 Census data from Ruggles et al. (2022), conditional on living in metropolitan areas: incumbent whites not living in an identified central city, self-reported white migrants from Southern states since 1955, and self-reported Black migrants from Southern states. While recent Southern white migrants earn more on average than Black migrants, their mean annual household income is \$6,500 in 1960 dollars. When compared to an average of \$8,400 for incumbent whites, there is a 23% gap in mean household incomes. Anticipation for these lower earners to purchase homes beyond the central city, and not pay their share of local public goods should still cause a less pronounced increase in restrictive suburban lot sizes.

To credibly estimate how non-Black demographic change impacts lot size outcomes, I repeat the shift-share identification strategy for Black migrants, but constructed with 1940 shares and outmigration rates for Southern whites. The shift-share, $Z^{\text{S–white}}$, instruments for an endogenous variable reflecting metropolitan changes in the Southern white population when metropolitan population is ambiguously defined, analogous to the composition change variable in Derenoncourt (2022):

$$\Delta MSA^{\text{S–white}}_{m(j),t} = \left(\log \text{S–white}_{m,t} - \log \text{S–white}_{m,t-1}\right) \times s^{w}_{m,1940}.$$

Appendix F.3 details variable construction, which uses 1940 full count Census data and postwar Census tabulations on migration rates.

I therefore run the instrumental variables regression

$$\begin{aligned} Reg_{jt} &= \beta^{W} \Delta MSA^{\text{S–white}}_{m(j),t} + \delta_t + \mathbf{X}_{j,pre}\Gamma + \varepsilon_{j,m(j)t} \\ \Delta MSA^{\text{S–white}}_{m(j),t} &= \gamma^{W} Z^{S-white}_{c(j)t} + \delta_t + \mathbf{X}_{c(j),pre}\Gamma + \nu_{c(j),t}, \end{aligned} \tag{9}$$

which is defined over a smaller analysis sample than in Section 5.1 due to incomplete coverage of the right-hand side variables.

Tables 4 and A.3 compares reduced form and TSLS estimates across demographic changes by race, using each group's definition of demographic change. The main finding of Panel (a) is that, unlike effects of Black migration on lot size outcomes, results over both restrictiveness and adoption using instrumented Southern white migration has effects that are null at the 90% confidence level, whose point estimates are somewhat negative.

While units differ between different demographic change variables, the estimates imply the migration of Southern whites to non-Southern metros explains 7% of restrictive lot size adoption and -30% of lot size restrictiveness. A negative effect on lot size restrictiveness implies that as Southern white migration grows in a metro, jurisdictions segregate neighborhoods less with minimum lot sizes or allow for denser housing than there was built earlier in the analysis time period. When lot size regulations adapt this way to market conditions, the urban area densifies in response to a growing population in a way consistent with standard Alonso-Muth-Mills theory.

One concern with this finding is that enough Southern White migrants were integrated in labor markets that their incomes would quickly converge with suburban incumbents. Tables 4 and A.3 show an alternative specification, testing if greater flight from the central city by foreign-born residents correlates with higher lot size outcomes. Foreign-born workers at the time were working class residents who could take up homeownership subsidies with average incomes similar to Southern white migrants in the 1960 Census. However, a one percentage point fall in the central city foreign-born resident share explains a minimal amount of lot size outcomes, even less than the share predicted from an OLS specification using Black demographic change.

Another concern is that Southern white migrants moved more to the West while Black migrants moved to industrial cities, so there are differential trends between cities receiving the most migrants in each group. To address this, Panel (b) of Tables 4 and A.3 limits all analyses to a common sample: the top 100 central cities by Black composition change for which Southern white migration data were recorded. A weak first stage in this sample leads to noisy estimates of Black composition effects, but for Southern white migration the first-stage F-stat remains strong at around 27. Nonetheless, Southern white composition change the effects remain null and negative.

This section's empirical strategies ultimately exploit variation in white migrant subsamples who still have large wage premia over Black migrants, so I lack an ideal test as described in Section 3.2. However, across multiple specifications no measure of Nonblack migration can produce effects in the same direction as the Great Migration effect. Fiscal zoning theories would predict a smaller, but positive response in land use regulation, unlike negative effects that suggest zoning remained loose to meet demand from white homeowners belonging to the working class.

5.4 Effects Robust Across Specifications

This section summarizes additional checks displayed in Table 5 and discuss alternate approaches to inference in my shift-share design. Appendix G.1 offers further robustness checks on how sensitive the results are to the estimation procedure, the definition of restrictive lot sizes, outcome mismeasurement and outlier cities.

Method of inference. In the baseline specification, I follow suggested procedures for shift-share instruments in Adão, Kolesár and Morales (2019) and Borusyak, Hull and Jaravel (2022) by clustering standard errors at the level of origin counties. Standard errors are similar if I use heteroskedastic-robust errors. Standard errors decrease if I cluster at the level of state-decades, which could be caused by negative spatial correlation between urban counties with high shock values versus low shock value neighbors.

Varying controls. Columns (2) and (3) of Table 5 offer evidence that main results on lot size adoption and restrictiveness are not sensitive to further controls for differential suburban demand across cities. Column (2) controls the shift-share instrument for Southern White migration defined in Section 5.3, which are larger in magnitude compared to the main estimates in Column (1).

Column (3) controls for 1940 density of each metro area's central city, calculated from Census CCDB data: one path closed by this control is if residents living in denser units before the war were more willing to suburbanize and consume more housing structure. Point estimates increase for the adoption result and attenuates somewhat for the restrictiveness result, though they are noisy because the instrument is weak over the sample for which density can be calculated.

Accounting for outcome mismeasurement. To address concerns that the bunching bins I estimate may reflect private developer decisions instead of persistent zoning, Column (4) of Table 5 filters on jurisdictions with a maximum gradient statistic of at least 1.5, before any grouping or estimation. The remaining jurisdictions have more significant bunching as a share of all development: the average jurisdiction in this subsample has lot size restrictiveness of 5.7%, versus 2.9% for the average jurisdiction overall. Point estimates for adoption and restrictiveness remain similar relative to Column (1), though estimates are noisier.

Additionally, Appendix G.1 uses collected historical zoning information used in Section 2.4 and a measurement error correction from Meyer and Mittag (2017) to argue adoption effects may be downward biased.

Alternative specifications. Column (5) of Table 5 shows results where jurisdiction-level controls are replaced with jurisdiction fixed effects calculated across three decades, so outcomes are residualized by those fixed effects before grouping and estimation.[35]When results are estimated using such within-jurisdiction variation, estimates are larger in magnitude and more precise.

Column (6) shows the timing of migration shocks relative to outcomes does not drive results: when I run regressions of Black demographic change over $[t, t+10]$ over outcomes calculated between $[t+5, t+15]$, effect estimates are close to Column (1) in magnitude and precision. Finally, Column (7) runs models where neither Black composition change nor the shift-share shocks were percentile transformed. The one standard deviation effect on restrictiveness using the untransformed specification is 3.1 percentage points. The magnitude is very similar to the main specification, even if the percentile transform makes results more robust to outlier metros.

[35] In other words, this specification replaces all controls except the missing shares control with jurisdiction fixed effects. This specification is not intended to be a two-way fixed effects regression with decade fixed effects, as TWFE models may differ just due to contaminated control groups

6 Exclusionary Zoning Circumvents Desegregation Policies

The Great Migration effect on restrictive zoning in Section 5 is robust to several race-neutral explanations, but also declines in importance over time. These results could be consistent with restrictive zoning satisfying suburbanites' homophily preferences, which became less salient as opinions on racial integration shifted in the North. Another explanation is that fiscal zoning — the process of zoning suburbs to prevent free-riding of local public goods — had greater salience when racial change in cities happened jointly with income change.

This section explores the importance of the latter mechanism in explaining the Great Migration effect. While my results cannot decompose how much different mechanisms contributed to the aggregate effect, they offer further evidence that postwar zoning evolved as contemporary demand for segregated public goods increased. Cities where desegregation policy were the most controversial may thus have seen further divergence in their suburban zoning than if those policies were not pursued.

6.1 School Desegregation In Times of Postwar Suburbanization

While the history of how the U.S. governments and courts reversed legal racial segregation in the South is well documented, the case is less true for "de facto" segregation of public goods in the North. The Supreme Court's *Brown v. Board* ruling offered a clear precedent to rulings in the 1960s that desegregated Southern schhols. Outside of Southern states, political action to desegregate schools is better interpreted as a series of localized political struggles.

Activists, often members of the local branch of the National Association for the Advancement of Colored People (NAACP), lobbied for remedies to overcrowded schools with predominantly Black student bodies, like busing or school consolidation. Their legal strategies faced not only opposition from representatives of incumbent white homeowners, but also early pushback from moderate NAACP members (Burkholder, 2021).

That said, by the late 1950s the NAACP's national committee had agreed to litigate and lobby for policy solutions to school desegregation in the North (Burkholder, 2021). One of the early examples of a court-ordered remedy for school segregation is in the 1963 decision *Taylor v. Bd. of Ed. of New Rochelle*: the most predominantly Black school "should be razed, with its children dispersed to other elementary schools." (Feron, 1986)

Against this evolution of school desegregation policy, other historians document that postwar suburbs planned development to mitigate what they expected were the consequences of desegregation. Erickson (2016) describes how Nashville, after an early court order to desegregate schools, barely changed attendance boundaries for schools in wealthy neighborhoods. Dougherty (2012) argues for Connecticut suburbs — and Erickson as well for Nashville suburbs — that local actors planned land use while consulting developers wishing to market exclusive schools as a neighborhood amenity.

Postwar suburbanization therefore occurred in tandem with two trends. First, non-land use planning strategies to maintain racial homogeneity within schools were gradually removed. Second, all the while that urban incumbents reacted with threats or violence to new Black students, suburbs were also building exclusive schools to attract desirable residents. A simple model connecting these trends is to assume metros have a measure of residents who have high income and whose marginal utility for local public goods is higher when those goods service a racially homogenous populations. Anticipating a fall in homogeneity due to policy changes if they do not move, they choose to sort into and pay a premium for new suburban neighborhoods that have large lot sizes. Greater lot size restrictiveness not only captures more of households' willingness to pay for homogenous public goods, but also excludes entry by Black households who are far from affording large lot size homes.

6.2 School Desegregation's Interactions With Great Migration Effects

I hypothesize that when a locality observes significant school desegregation, incumbents who see this as a disamenity on the margin relocate to lot restricted suburbs. Even if mobility is unobserved, local actors adopting restrictive lot sizes to meet such demand should show up as increased lot size restrictiveness in jurisdictions surrounding the locality. I consider "significant" school desegregation to be either individual Black families sending children to entirely White schools or changing school assignment following desegregation plans, but I limit my focus to cities with the most Black composition change, where the scope of desegregation should lead to wider metropolitan impacts.

My first estimand of interest therefore is defined over metros with above median predicted levels of Black migration, using the shift-share instrument calculated up to 1950. In this sample, I take the difference of means of lot size outcomes between metros with desegregated schools and more segregated ones:

$$E[Reg_{jt} \,|\text{desegregated}_{m(j)t}, F(\tilde{Z}) \geq 0.5] > E[Reg_{jt} \,|\text{segregated}_{m(j)t}, F(\tilde{Z}) \geq 0.5]. \qquad (10)$$

If the above inequality holds in the data, it suggests desegregated metropolitan school systems also had more residential sorting into lot size restricted neighborhoods. The earliest data on metropolitan school segregation indices I have are 1968 data surveyed by the Department of Education, tabulated in Logan, Oakley and Stowell (2008). I define a metro as desegregated based on it having below-median values of Black/White dissimilarity between schools across the metro. The variation in school desegregation is mapped in Panel (a) of Figure B.6. I also calculate outcomes separately every 5 years in this section compared to 10 years earlier. The main results are presented grouped into three separate time periods, though Appendix Figure G.3 show event study plots by each 5-year period.

Panel (a) of Figure 9 plots whether the difference of means in Equation 10 is positive, with 95% confidence intervals generated from a wild bootstrap sampled over metropolitan areas. I find that, by the end of my sample, a 0.9 percentage point increase in average jurisdiction lot size restrictiveness when predicted Black migrant destinations are undergoing more school desegregation. This accounts for 30% of the total Great Migration effect outside the South estimated in Section 5.1.

Because school segregation measures predetermined before 1950 are unavailable, these results could come out of reverse causality: postwar lot size restrictiveness that were not a reaction to school desegregation restricted housing options to white incumbents as well, lowering their likelihood to self-segregate despite racial change in schools. However, Panel (a) also compares metropolitan areas with above and below median residentially desegregation, calculated using changes in segregation indices from Cutler, Glaeser and Vigdor (1999). More desegregated metros see *lower* lot size restrictiveness, an opposite sign effect. Appendix Section G.2 also shows states with more anti-discrimination laws in 1950 did not have positive lot size restrictiveness responses. I conclude the school desegregation measure is unlikely to be selected on larger-scale factors that drive lot size restrictiveness, and more dependent on local policy decisions.

6.3 Evidence From Local Policy Case Studies

The results in Section 6.2 suggest lot size outcomes respond to school desegregation, but lack external validity as a treatment effect unless school desegregation dynamics are unconfounded (conditionally independent) of $\tilde{Z}$. [36] I aim for a more internally valid research design by focusing on a set of counties where desegregation plans were legally contested before lot size adoption slowed in the 1970s.

Court-ordered desegregation in the North, such as busing plans with a city's suburbs, is

[36] This follows from the usual arguments that without unconfoundedness, school segregated metro averages are unequal to counterfactual averages for school desegregated metros if they were not desegregated.

mostly a 1970s phenomenon (Baum-Snow and Lutz, 2011). However, I use data compiled by Brown University (2005) to identify 18 counties that are "early adopters" of desegregation plans, because a jurisdiction in it had a plan imposed by or affirmed by a court challenge. I consider the whole county to be an adopter to account for spatial spillovers of a plan on the local housing market: i.e. a desegregation plan in New Haven, CT should trigger a lot size response in surrounding suburbs. The counties are mapped in Panel (b) of Figure B.6.

The earliest cases in the adopter data are the landmark *Brown v. Board* and *Bolling v. Sharpe* cases, decided in 1954; the last cases were decided in 1966. While most cases were argued in the 1960s, treating the decision date as an adoption date neglects that many cases were also affirmations of a plan passed by local officials, which themselves follow sustained debate and minor concessions to activists.[37] I make an assumption that the "event" of a school desegregation plan had a common beginning in the early 1950s, so they all share a common pre-period of the 1940s.

With these additional analysis decisions in mind, I run a dynamic event study at the jurisdiction level to derive a variant of Equation 10. Instead of conditional means based on more or less school segregation, that variable is replaced by the jurisdiction being in an early adopter county or not having any desegregation at all: I filter out of the control group any county where failed desegregation cases occurred, and also counties where the early U.S. Department of Education were investigating discriminatory school segregation.

Panel (b) of Figure 9 show the results binned similarly as in Panel (a)[38] The event study estimates on lot size restrictiveness are also positive and statistically significant. Effect sizes indicate that when a nearby locality considers and adopts a school desegregation plan, surrounding suburbs zone more restrictively at a rate that is 75% to 90% of the total Great Migration effect.[39]

When comparing just the difference of means, there is a significant difference in outcomes between early adopter and control counties in the 1940s pre-period. To ensure results are not driven by a violation of parallel trends, I further estimate a triple difference strategy: I subtract the difference-in-differences estimate over high $\tilde{Z}$ metros with the difference-in-differences estimate calculated over below median $\tilde{Z}$ metros. The identification assumption in this case

[37] Appendix Section C.8 summarizes a successful and a failed court case. Another interpretation is that successful cases are proxies for localities where desegregation plans were comprehensive and adopted despite polarized political support. Failed cases reflect localities where either desegregation advocates or opponents were marginal, unable to articulate why the court should alter existing consensus.

[38]Apart from the change to being in an early adopter county, standard error bootstrapping and specifications are identical to what was used to produce Panel (a).

[39] Appendix Figure G.4 breaks down effects by observations separated every 5 years to show the displayed effects grow steadily over time.

is that any diverging trends between $\tilde{Z}$ stratified metros that have desegregation plans are parallel to divergent trends between $\tilde{Z}$ stratified metros that lack desegregation plans (Olden and Møen, 2022). The relatively larger group of jurisdictions in the latter category that lack plans means I can estimate the trends to a degree of precision.

The triple difference results are also presented in Panel (b), where the assumed reference group is the 1940s pre-period. The triple difference point estimates are somewhat attenuated but not significantly so from earlier point estimates.

Appendix Section G.2 also shows null effects if the same triple difference design is used not on early adopter counties, but counties where desegregation cases did not lead to plan adoption. I also show in the same Section that the same design cannot produce significant effects in unconditional excess mass or growth in housing supply in the same jurisdictions. The collection of results suggest effects are not driven by speculation over unsuccessful cases, or that metros with desegregation plans had substantially greater white flight. Rather, local actors had learned to plan both land use and investment in education. The minority in Northern metropolitan areas who disapproved public good integration were nevertheless customers to whom jurisdictions can advertise the amenity of exclusive public goods.

This Section's results, in the context of effect drop offs over time in Section 5.2, suggests exclusionary zoning surrounding Great Migration destinations would have further declined if not for household demand to avoid schools showing racial integration. Another interpretation is that further divergence in lot size restrictiveness after the 1950s was contingent on the intensity with which localities pursued integration policy. Without state or federal intervention to limit the regulatory powers of jurisdictions, ambitious integration policy faced an equity-efficiency tradeoff: competing jurisdictions zoned more exclusively to meet short-run demand, a planning decision that had persistent consequences over multiple ensuing decades.

7 Conclusion

In this paper, I develop a novel statistical algorithm that measures when and how U.S. suburban land use controls emerged over time. While legal tools local actors can use to regulate land use continue to evolve, my algorithm focuses on the intensity with which suburbs have used minimum lot size controls since 1940. Beyond real estate, the algorithm has applications for numerous economic questions where bunching estimators are used.

I use time-varying outcomes in lot size adoption and restrictiveness, the first in the literature, to estimate how much the adoption of restrictive lot size controls can be explained by postwar Black migration to non-Southern cities. While the literature has focused on how regulation distorts housing supply in recent time, from 1940–1970 around 600 thousand units could have been developed more densely absent local government land use controls. Additional results suggest these effects come from suburban local actors willing to supply the demand for exclusive neighborhoods in the housing market, especially when cities debated how to racially integrate access to local public goods.

My outcomes of interest are most relevant for suburban jurisdictions and not for the urban core of metropolitan areas.[40] However, recent policy proposals like California's Senate Bills 9 and 10 in 2021 or Massachusetts's 2022 upzoning regulations specifically override density controls beyond the urban core to improve housing affordability (MilNeil (2022)). My results indicate these policies are defensible from an equity angle — reversing planning decisions attributable to racial prejudice — and are targeted toward neighborhoods that had and likely can still support dense development.

Policy reforms grow in urgency as America anticipates higher internal mobility. In recent years, households from the largest cities were on the move in the wake of remote work (Ramani and Bloom (2021)). Increased busing of immigrants to large American cities, and resistance to temporary housing for those migrants, formed what the cities' mayors call a "migrant crisis." (Montoya-Galvez (2023)) What I document in the past may speak to tomorrow's planning concerns: if incumbents view migrants as part of an out-group, they can use the law to exclude and limit residential choice for the newcomers. Decisionmaking at higher levels of government should consider how to align local governments' incentives away from restrictive land use to ensure growing cities remain integrated and equitable.

[40]In the urban core, zoning parameters are based much less on minimum lot sizes and more on floor-to-area ratios (FAR), as studied by Brueckner and Singh (2020) and Anagol, Ferreira and Rexer (2021).

References

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. "Shift-Share Designs: Theory and Inference." *The Quarterly Journal of Economics*, 134(4): 1949–2010.

Albright, Alex, Jeremy A. Cook, James J. Feigenbaum, Laura Kincaide, Jason Long, and Nathan Nunn. 2021. "After the Burning: The Economic Effects of the 1921 Tulsa Race Massacre."

Anagol, Santosh, Benjamin B Lockwood, Allan Davids, and Tarun Ramadorai. 2022. "Diffuse Bunching with Frictions: Theory and Estimation."

Anagol, Santosh, Fernando V. Ferreira, and Jonah M. Rexer. 2021. "Estimating the Economic Value of Zoning Reform."

Babcock, Richard F. 1966. *The Zoning Game: Municipal Practices and Policies.* University of Wisconsin Press.

Babcock, Richard F., and Fred P. Bosselman. 1973. *Exclusionary Zoning. Praeger Special Studies in U.S. Economic, Social and Political Issues*, Praeger.

Bai, Jushan, and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica*, 66(1): 47.

Bai, Jushan, and Pierre Perron. 2003. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics*, 18(1): 1–22.

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. 2020. "Urban Transport Expansions and Changes in the Spatial Structure of U.S. Cities: Implications for Productivity and Welfare." *The Review of Economics and Statistics*, 102(5): 929–945.

Baum-Snow, Nathaniel, and Byron F. Lutz. 2011. "School Desegregation, School Choice, and Changes in Residential Location Patterns by Race." *American Economic Review*, 101(7): 3019–3046.

Bazzi, Samuel, Andreas Ferrara, Martin Fiszbein, Thomas Pearson, and Patrick A Testa. 2023. "The Other Great Migration: Southern Whites and the New Right*." *The Quarterly Journal of Economics*, 138(3): 1577–1647.

Berkes, Enrico, Ezra Karger, and Peter Nencka. 2022. "The Census Place Project: A Method for Geolocating Unstructured Place Names."

Bertanha, Marinho, Carolina Caetano, Hugo Jales, and Nathan Seegert. 2023. "Bunching Estimation Methods." *Handbook of Labor Economics*, 5.

Bogart, William T. 1993. "'What Big Teeth You Have!': Identifying the Motivations for Exclusionary Zoning." *Urban Studies*, 30(10): 1669–1681.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-Experimental Shift-Share Research Designs." *The Review of Economic Studies*, 89(1): 181–213.

Boustan, Leah. 2016. "Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets." National Bureau of Economic Research, Inc NBER Books.

Boustan, Leah Platt. 2010. "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration*." *The Quarterly Journal of Economics*, 125(1): 417–443.

Bowles, Gladys K., James D. Tarver, Calvin L. Beale, and Everette S. Lee. 2016. "Net Migration of the Population by Age, Sex, and Race, 1950-1970."

Broderick, Tamara, Ryan Giordano, and Rachael Meager. 2021. "An Automatic Finite-Sample Robustness Metric: When Can Dropping a Little Data Make a Big Difference?"

Bronin, Sara C., and Ilya Ilyankou. 2021. "How to Make a Zoning Atlas: A Methodology for Translating and Standardizing District-Specific Regulations."

Brown University, American Communities Project. 2005. "Desegregation Court Cases & School Demographic Data."

Brueckner, Jan K., and Ruchi Singh. 2020. "Stringency of Land-Use Regulation: Building Heights in US Cities." *Journal of Urban Economics*, 116: 103239.

Bucholtz, Shawn, Emily Molfino, and Jed Kolko. 2020. "The Urbanization Perceptions Small Area Index: An Application of Machine Learning and Small Area Estimation to Household Survey Data."

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew A. Turner. 2006. "Causes of Sprawl: A Portrait from Space*." *The Quarterly Journal of Economics*, 121(2): 587–633.

Burkholder, Zoë. 2021. *An African American Dilemma: A History of School Integration and Civil Rights in the North.* Oxford University Press.

Calabrese, Stephen, Dennis Epple, and Richard Romano. 2007. "On the Political Economy of Zoning." *Journal of Public Economics*, 91(1): 25–49.

Calabrese, Stephen M., Dennis N. Epple, and Richard E. Romano. 2012. "Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition." *Review of Economic Studies*, 79(3): 1081–1111.

Calderon, Alvaro, Vasiliki Fouka, and Marco Tabellini. 2023. "Racial Diversity and Racial Policy Preferences: The Great Migration and Civil Rights." *The Review of Economic Studies*, 90(1): 165–200.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs*." *The Quarterly Journal of Economics*, 134(3): 1405–1454.

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.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility."

Collier, Benjamin, Cameron Ellis, and Benjamin J. Keys. 2022. "The Cost of Consumer Collateral: Evidence from Bunching."

Collins, William J. 2021. "The Great Migration of Black Americans from the US South: A Guide and Interpretation." *Explorations in Economic History*, 80: 101382.

Cook, Lisa D, Maggie E C Jones, Trevon D Logan, and David Rosé. 2022. "The Evolution of Access to Public Accommodations in the United States*." *The Quarterly Journal of Economics*, qjac035.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy*, 107(3): 455–506.

Derenoncourt, Ellora. 2022. "Can You Move to Opportunity? Evidence from the Great Migration." *American Economic Review*, 112(2): 369–408.

Ditzen, Jan, Yiannis Karavias, and Joakin Westerlund. 2021. "Xtbreak: Estimating and Testing for Structural Breaks in Stata". Working Paper,."

Dougherty, Jack. 2012. "Shopping for Schools: How Public Education and Private Housing Shaped Suburban Connecticut." *https://journals.sagepub.com/doi/full/10.1177/0096144211427112*.

Duranton, Gilles, and Diego Puga. 2019. "Urban Growth and Its Aggregate Implications." National Bureau of Economic Research w26591, Cambridge, MA.

Erickson, Ansley T. 2016. "Making the Unequal Metropolis: School Desegregation and Its Limits." In *Making the Unequal Metropolis*. University of Chicago Press.

Feron, James. 1986. "New Rochelle Recalls Landmark Bias Ruling." *The New York Times*.

Fischel, William A. 2015. *Zoning Rules! : The Economics of Land Use Regulation.* Lincoln Institute of Land Policy.

Freund, David M. P. 2007. *Colored Property: State Policy and White Racial Politics in Suburban America. Historical Studies of Urban America*, University of Chicago Press.

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

Goff, Leonard. 2023. "Treatment Effects in Bunching Designs: The Impact of Mandatory Overtime Pay on Hours."

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review*, 110(8): 2586–2624.

Gordon, Colin. 2008. *Mapping Decline: St. Louis and the Fate of the American City.* University of Pennsylvania Press.

Gregory, James N. 2005. *The Southern Diaspora: How the Great Migrations of Black and White Southerners Transformed America.* Chapel Hill ::University of North Carolina Press,.

Grumbach, Jacob M., Robert Mickey, and Daniel Ziblatt. 2023. "The Insulation of Local Governance from Black Electoral Power: Northern Cities and the Great Migration."

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

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

Hilber, Christian A. L., and Frédéric Robert-Nicoud. 2013. "On the Origins of Land Use Regulations: Theory and Evidence from US Metro Areas." *Journal of Urban Economics*, 75: 29–43.

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

Infranca, John. 2023. "Singling Out Single-Family Zoning." *THE GEORGETOWN LAW JOURNAL*, 111.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics*, 8(1): 435–464.

Klimek, Amanda, Christopher Mazur, William Chapin, and Ellen Wilson. 2018. "Housing Administrative Records Simulation."

Kneebone, Elizabeth. 2009. "Job Sprawl Revisited: The Changing Geography of Metropolitan Employment."

Krimmel, Jacob. 2022. "Reclaiming Local Control: School Finance Reforms and Housing Supply Restrictions."

Kulka, Amrita, Aradhya Sood, and Nicholas Chiumenti. 2022. "How to Increase Housing Affordability? Understanding Local Deterrents to Building Multifamily Housing."

Leys, Christophe, Christophe Ley, Olivier Klein, Philippe Bernard, and Laurent Licata. 2013. "Detecting Outliers: Do Not Use Standard Deviation around the Mean, Use Absolute Deviation around the Median." *Journal of Experimental Social Psychology*, 49(4): 764–766.

Logan, John R., Deirdre Oakley, and Jacob Stowell. 2008. "School Segregation in Metropolitan Regions, 1970–2000: The Impacts of Policy Choices on Public Education." *American Journal of Sociology*, 113(6): 1611–1644.

Lower Merion Township, PA Planning Commission. 1937. *A Plan for Lower Merion Township.*

Lutz, Chandler, and Ben Sand. 2023. "Highly Disaggregated Land Unavailability."

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2021. *IPUMS National Historical Geographic Information System: Version 16.0 [Dataset].* Minneapolis, MN:IPUMS.

McCasland, Jamie, Tomàs Monarrez, David Schönholzer, and Everett Stamm. 2023. "Shattered Metropolis: The Great Migration and The Fragmentation of Political Jurisdictions."

McMichael, Stanley L. 1949. *Real Estate Subdivisions. Prentice-Hall Real Estate Series,* New York:Prentice-Hall.

Meyer, Bruce D., and Nikolas Mittag. 2017. "Misclassification in Binary Choice Models." *Journal of Econometrics*, 200(2): 295–311.

MilNeil, Christian. 2022. "New State Rule Would Force Suburbs to Legalize Thousands of New Apartments Near T Stops."

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

Montoya-Galvez, Camilo. 2023. "As Cities Struggle to House Migrants, Biden Administration Resists Proposals That Officials Say Could Help - CBS News." *https://www.cbsnews.com/news/migrant-crisis-nyc-chicago-biden-administration-proposals/.*

Murray, Pauli. 1950. *States' Laws on Race and Color.* [Cincinnati,:Woman's Division of Christian Service, Board of Missions and Church Extension, Methodist Church].

Olden, Andreas, and Jarle Møen. 2022. "The Triple Difference Estimator." *The Econometrics Journal*, 25(3): 531–553.

Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics*, 31(3): 358–369.

Parkhomenko, Andrii. 2020. "Local Causes and Aggregate Implications of Land Use Regulation."

Ramani, Arjun, and Nicholas Bloom. 2021. "The Donut Effect of Covid-19 on Cities."

Reber, Sarah J. 2011. "From Separate and Unequal to Integrated and Equal? School Desegregation and School Finance in Louisiana." *The Review of Economics and Statistics*, 93(2): 404–415.

Reynolds, Conor Dwyer. 2019. "The Motives for Exclusionary Zoning."

Rolleston, Barbara Sherman. 1987. "Determinants of Restrictive Suburban Zoning: An Empirical Analysis." *Journal of Urban Economics*, 21(1): 1–21.

Ross, Stephen L., and Geoffrey M. B. Tootell. 2004. "Redlining, the Community Reinvestment Act, and Private Mortgage Insurance." *Journal of Urban Economics*, 55(2): 278–297.

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

Rubin, Donald B. 1981. "The Bayesian Bootstrap." *The Annals of Statistics*, 9(1): 130–134.

Ruggles, Steven, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Matt A. Nelson, Evan Roberts, Megan Schouweiler, and Matthew Sobek. 2021. *IPUMS Ancestry Full Count Data: Version 3.0 [Dataset].* Minneapolis, MN:IPUMS.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Megan Schouweiler, and Matthew Sobek. 2022. *IPUMS USA: Version 12.0 [Dataset].* Minneapolis, MN:IPUMS.

Saavedra, Martin, and Tate Twinam. 2020. "A Machine Learning Approach to Improving Occupational Income Scores." *Explorations in Economic History*, 75: 101304.

Sahn, Alexander. 2022. "Racial Diversity and Exclusionary Zoning: Evidence from the Great Migration."

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *The Journal of Mathematical Sociology*, 1(2): 143–186.

Shertzer, Allison, and Randall P. Walsh. 2019. "Racial Sorting and the Emergence of Segregation in American Cities." *The Review of Economics and Statistics*, 101(3): 415–427.

Shi, Ying, Daniel Hartley, Bhash Mazumder, and Aastha Rajan. 2022. "The Effects of the Great Migration on Urban Renewal." *Journal of Public Economics*, 209: 104647.

Song, Jaehee. 2021. "The Effects of Residential Zoning in U.S. Housing Markets."

Sood, Aradhya, and Kevin Ehrman-Solberg. 2023. "Long Shadow of Racial Discrimination: Evidence from Housing Racial Covenants."

Stuart, Bryan A., and Evan J. Taylor. 2021. "Migration Networks and Location Decisions: Evidence from US Mass Migration." *American Economic Journal: Applied Economics*, 13(3): 134–175.

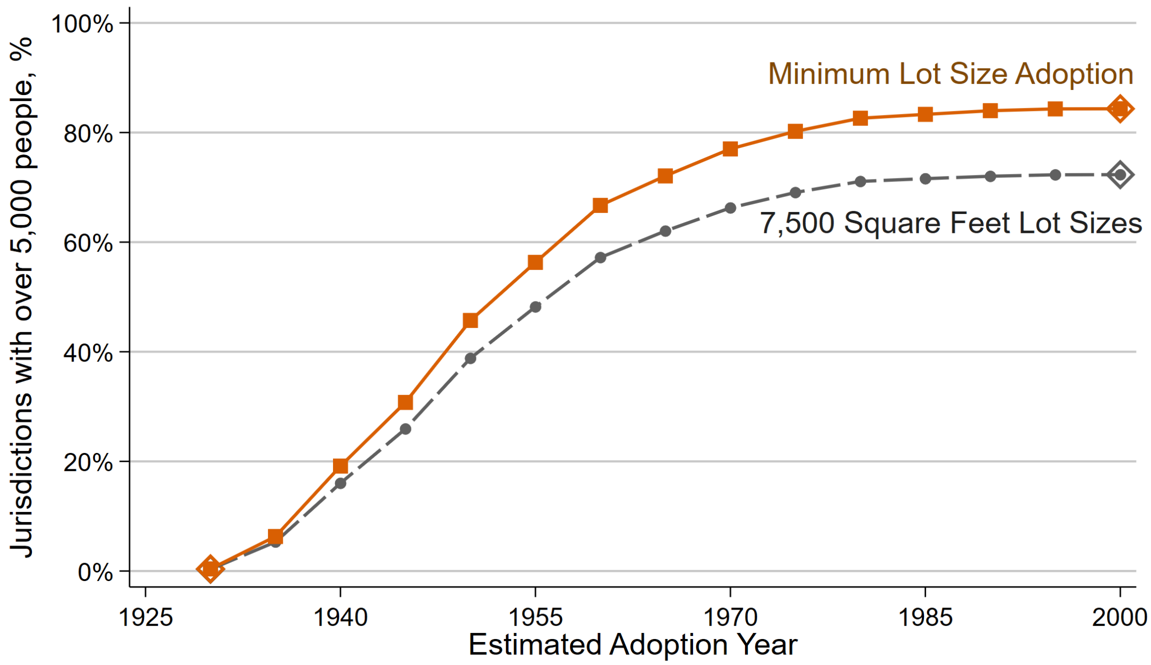
Trounstine, Jessica. 2020. "The Geography of Inequality: How Land Use Regulation Produces Segregation." *American Political Science Review*, 114(2): 443–455.

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

Winling, LaDale C, and Todd M Michney. 2021. "The Roots of Redlining: Academic, Governmental, and Professional Networks in the Making of the New Deal Lending Regime." *Journal of American History*, 108(1): 42–69.

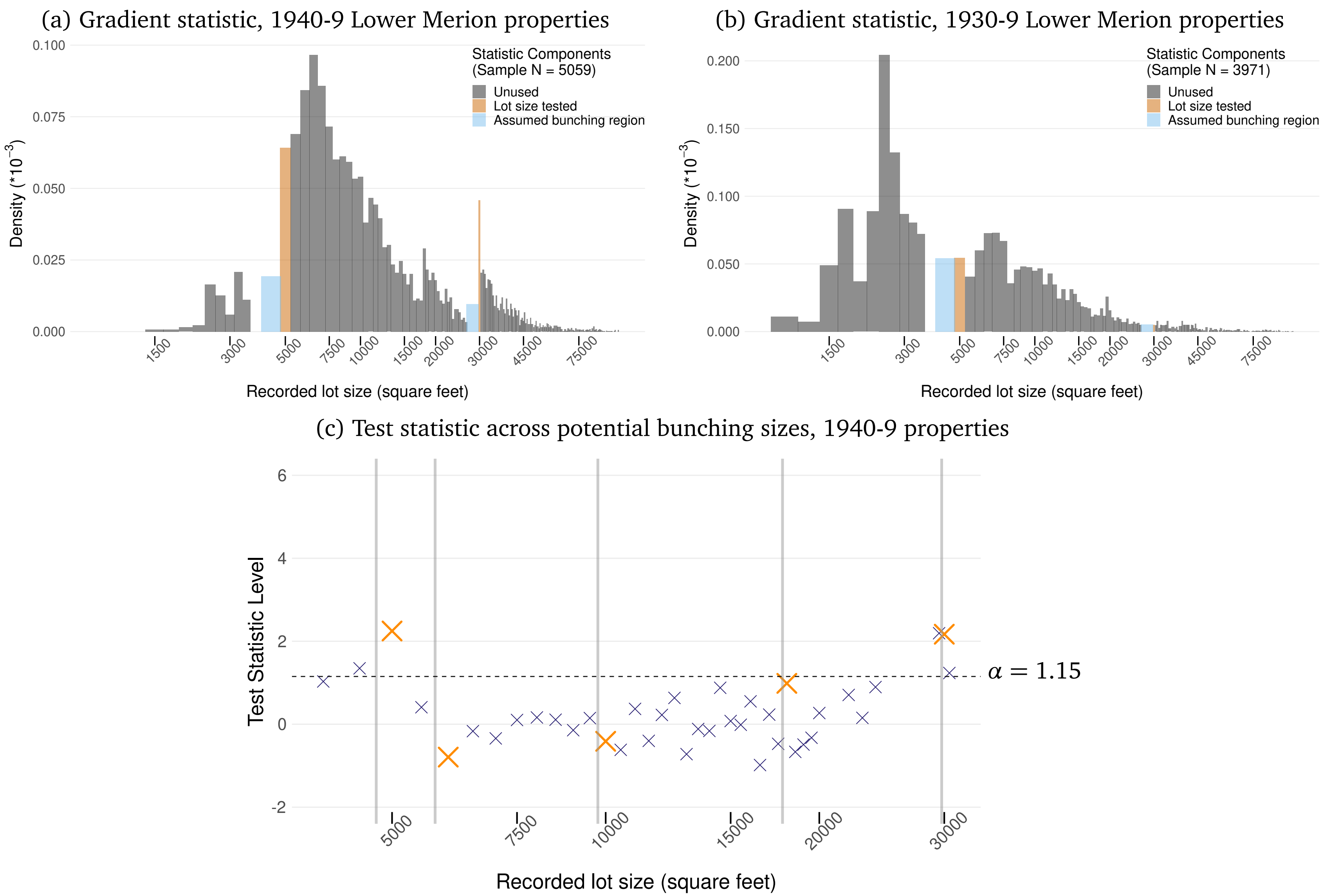
Zabel, Jeffrey, and Maurice Dalton. 2011. "The Impact of Minimum Lot Size Regulations on House Prices in Eastern Massachusetts." *Regional Science and Urban Economics*, 41(6): 571–583.

Figure 1: Estimated Trends in Minimum Lot Size Adoption



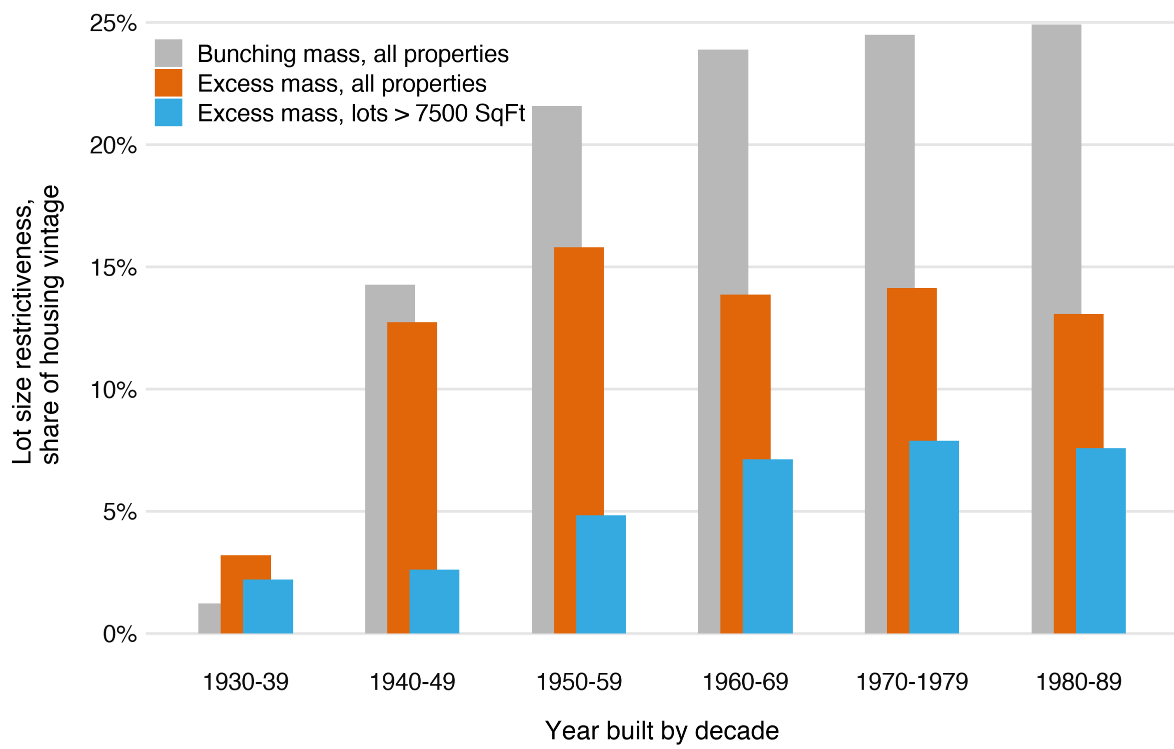
Notes: This figure plots the first adoption estimates defined in Section 3.1 based on the lot size detection algorithm in Section 2. Adoption rates are calculated over all cities and counties with zoning powers with at least 5,000 residents as of the 2010 census. The main series considers adoption of any minimum lot size, while the dashed series considers adoption of large minimum lots greater than 7,500 square feet ($\sim 1/6$ acre). Displayed rates are based on values at the end of five-year time bins. Diamond markers note the start and end of the time series in 1930 and 2000, respectively.
Sources: Calculations from CoreLogic Tax Records, 2010 Census.

Figure 2: Visualization of Locally Estimated Bunching Model



Notes: This figure visualizes the construction of the gradient statistic $\hat{G}$, defined in Section 2.2, in steps. For two known minimum lot sizes, Panel (a) highlights which histogram densities are used to calculate $\hat{G}$ terms based on the post-period sample. Panel (b) does the same but for the remaining two terms based on the pre-period sample. Panel (c) plots $\hat{G}$ for every lot size for which the statistic is defined. Lot sizes minimums mandated by the jurisdiction's regulations are highlighted in orange. Values are compared to the critical value, marked by the dashed line.
Sources: Calculations from CoreLogic Tax Records.

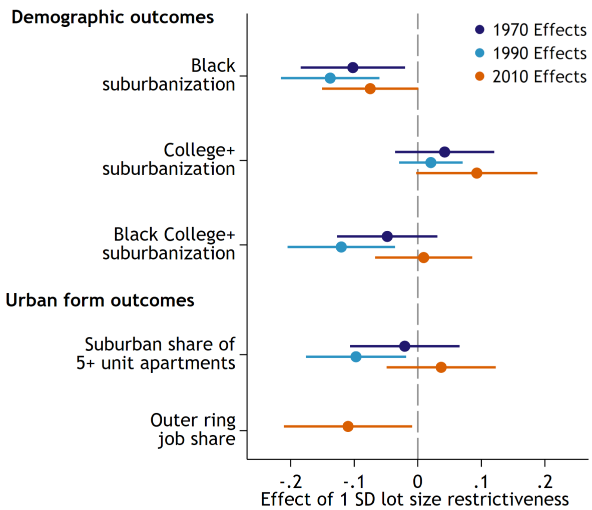
Figure 3: National Trends in Lot Size Restrictiveness



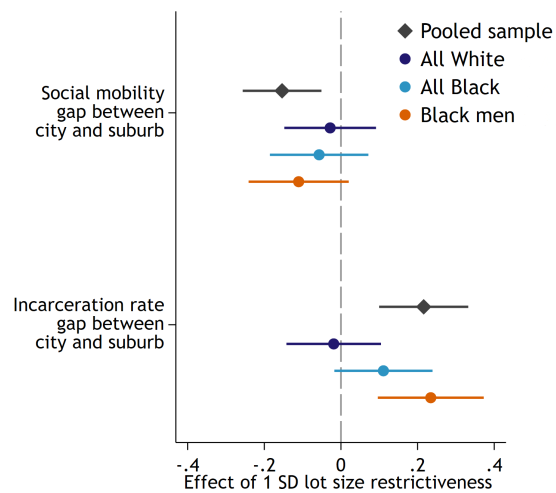
Notes: This figure plots how three measures of lot size restrictiveness, described in Section 3.1, change across homes built in different decades. Over a decadal vintage of single-family and duplex homes, the bunching mass and excess mass measures over all properties sums up the respecive measures over all detected bunching bins. The bunching mass measure for lots over 7,500 square feet sums up mass around bunching bins over 7,500 square feet, then express it as a ratio of all properties that decade.
Sources: Calculations from CoreLogic Tax Records.

Figure 4: Persistent Impacts of Postwar Exclusionary Zoning

(a) Persistence on urban stratification



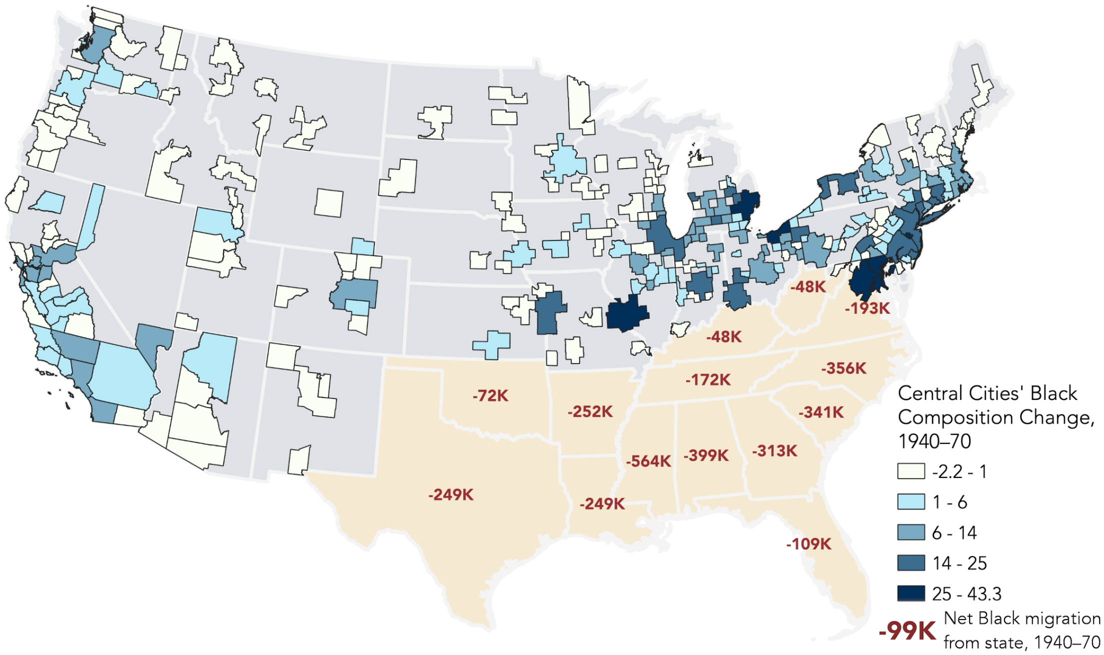
(b) Persistence on social mobility



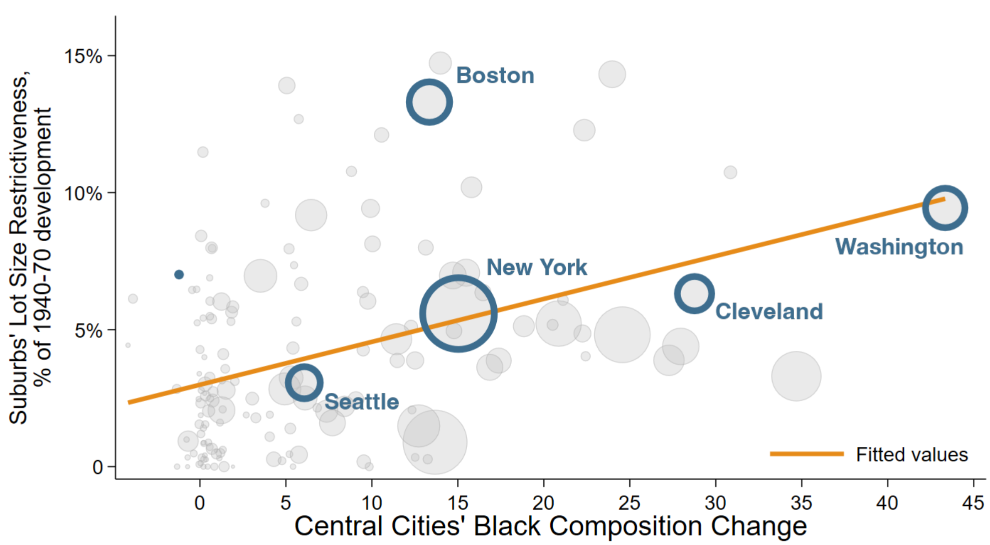
Notes: Both panels plot metropolitan area level variation in large lot restrictiveness, calculated from excess mass bunching at large lots as described in Section 3.1. Panel (a) visualizes a rank-rank regression of metro areas based on postwar intensity of large lot zoning with their ranked intensity in the 40 years onwards. A bubble plot scaled by postwar development levels is paired with a binscatter. Panel (b) visualizes coefficients from a stacked regression of predetermined postwar large lot restrictiveness, as described in Section 3.1. Each point estimate is paired with 90% confidence intervals. Both the predetermined variable and outcomes are standardized separately for each decade, so coefficients represent how a 1 SD national rise in large lot restrictiveness associates with SD changes in the outcome.
Sources: Calculations from NHGIS Tables (Manson et al. (2021)), Ruggles et al. (2022)), Census ZIP Code Business Patterns and CoreLogic Tax Records.

Figure 5: The Geography of Postwar Minimum Lot Sizes and Black Migration

(a) Black demographic change during the Second Great Migration

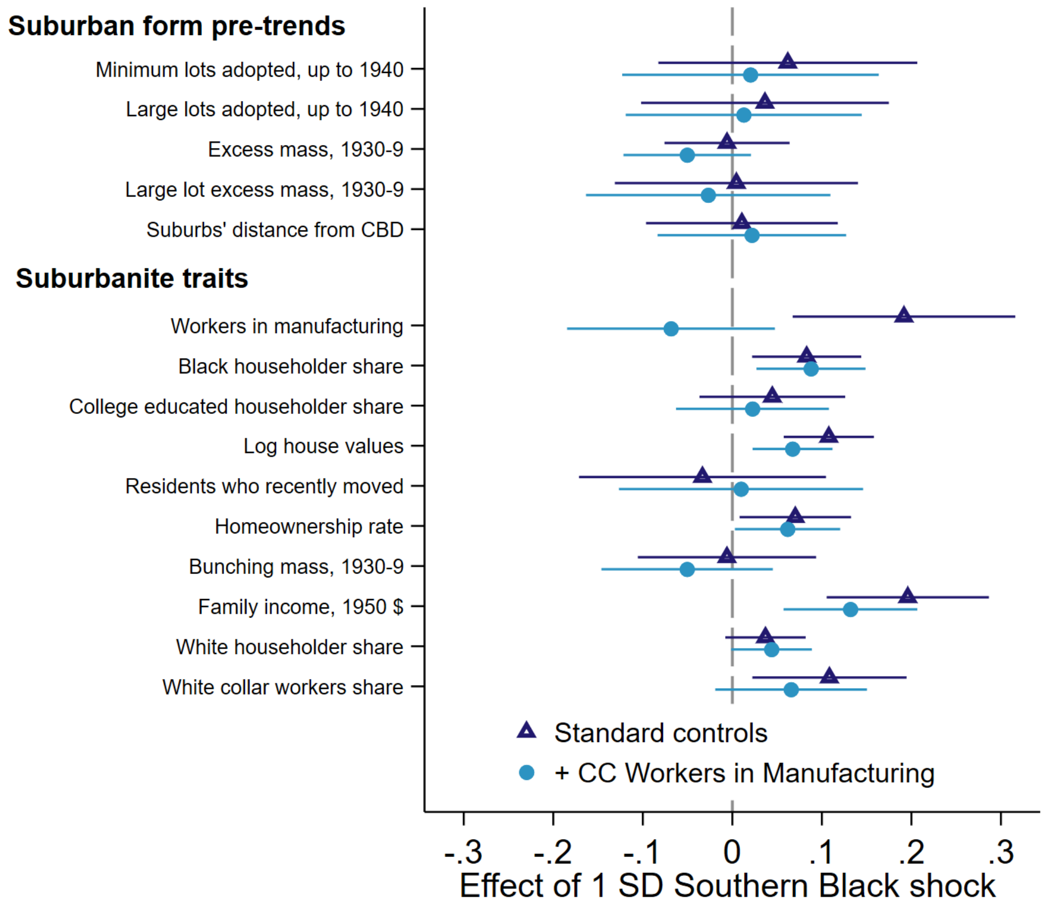


(b) Black migration associates with lot size restrictiveness



Notes: Panel (a) of this figure plots the cumulated migration out of 14 Southern states, as well as the variation across CBSA central cities in Black migration growth. Each non-Southern CBSA is marked by its central city's Black composition change variable ΔCC^{Black}_{ct} as defined in Section 4.1, transformed to the percentile in the decade distribution and averaged over 1940–70. State-level outmigration estimates are based off of methods in Gregory (2005) and are further explained in Appendix Section C.6.
Panel (b) of this figure aggregates the excess mass measure of lot size restrictiveness, as defined in Section 3.1, over 1940–70 and all detected jurisdictions in CBSA borders. Bubble sizes are based on levels of postwar housing in each CBSA. CBSAs are filtered to where minimum lot sizes were detected for a nontrivial number of jurisdictions.
Sources: Calculations from NHGIS Tables (Manson et al. (2021)), Ruggles et al. (2022), Gregory (2005) and CoreLogic Tax Records.

Figure 6: Balancing Checks for Southern Black Migration Shocks

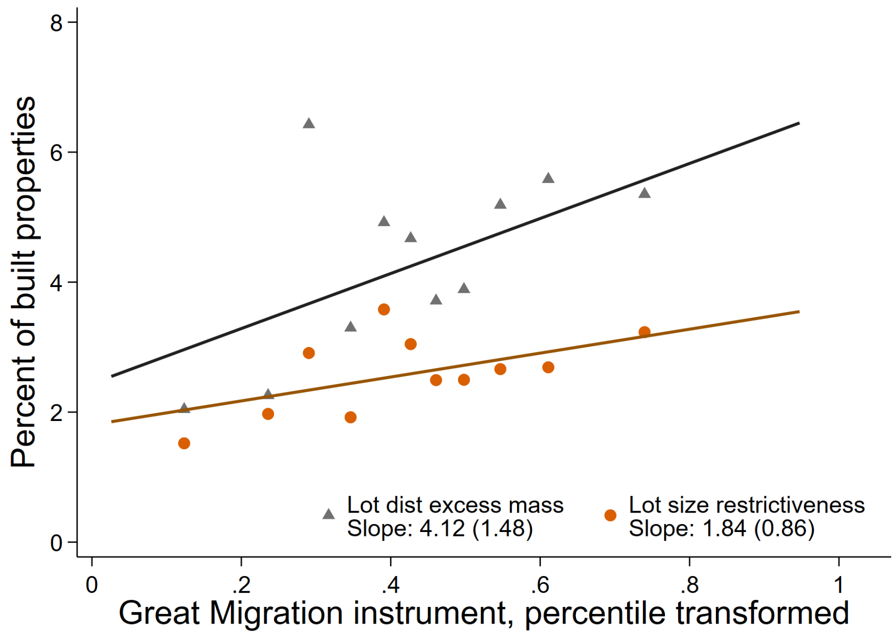


Notes: This plot uses the specification suggested by Borusyak, Hull and Jaravel (2022), projecting Northern city-level covariates $x_{j,pre}$ onto Southern counties and running the shock-level regressions

$$x^{\perp}_{k,pre} = \hat{g}(k)_{pre} + \xi_k,$$

where $\hat{g}(k)_{pre}$ are predicted 1950 Black migration shocks defined in Section 4.2. Pre-period covariates are tabulated from the 1940 Full count census, or derived from Corelogic assessor records for homes built up to 1940. I show results for two specifications, one without controls and another with transformed Northern central city manufacturing on the right hand side. 95% confidence intervals are shown with standard errors clustered at the county level.
Sources: Calculations from IPUMS 1940 full count Census (Ruggles et al. (2021)), Boustan (2016), Derenoncourt (2022) and CoreLogic Tax Records.

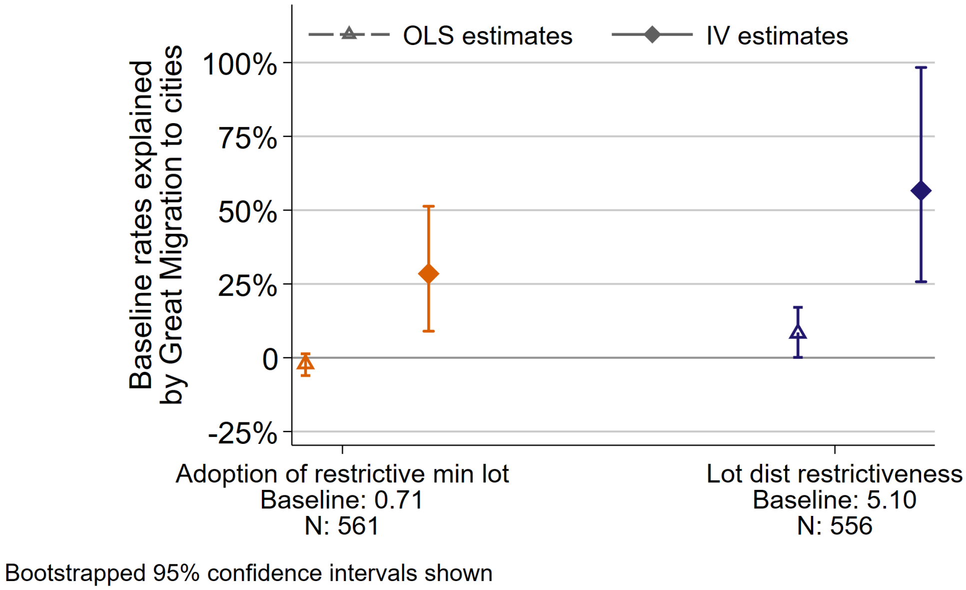
Figure 7: Lot Size Restrictiveness Linked with Black Composition Change



Notes: This figure plots reduced form, nonparametric relationships between the shift-share instrument described in Section 4.2 and two outcomes of interest: measures of lot size restrictiveness from 1940–1970. The source data is a panel of non-central city jurisdictions in CBSAs outside of 14 Southern states. The instrument is first transformed from levels to percentiles of each decade's distribution, as in Derenoncourt (2022), then residualized on share exposure variables as described in Section 4.2 and additional controls. Control variables include the CBSA central city's manufacturing share, and analysis sample cities' 1940 black share, homeownership rates and distance to CBD, interacted by period. Reported standard errors are clustered at the CBSA-decade level.
Sources: Calculations from NHGIS Tables (Manson et al. (2021)), Ruggles et al. (2022), CCDB, IPUMS 1940 full count Census (Ruggles et al. (2021)), Boustan (2016), Derenoncourt (2022) and CoreLogic Tax Records.

Figure 8: Calculating The Impact of Racial Exclusionary Zoning

(a) Lot Size Outcomes Explained by Black Migration, Constant Effects

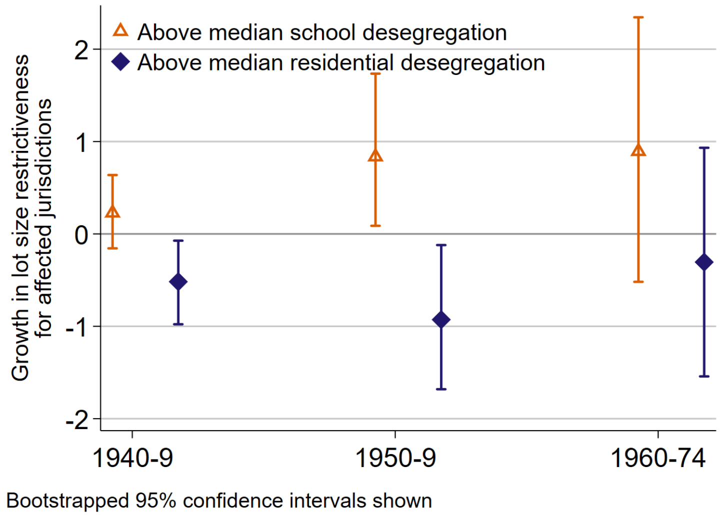


(a) Outcome Explained by Black Migration, Decade Varying Effects

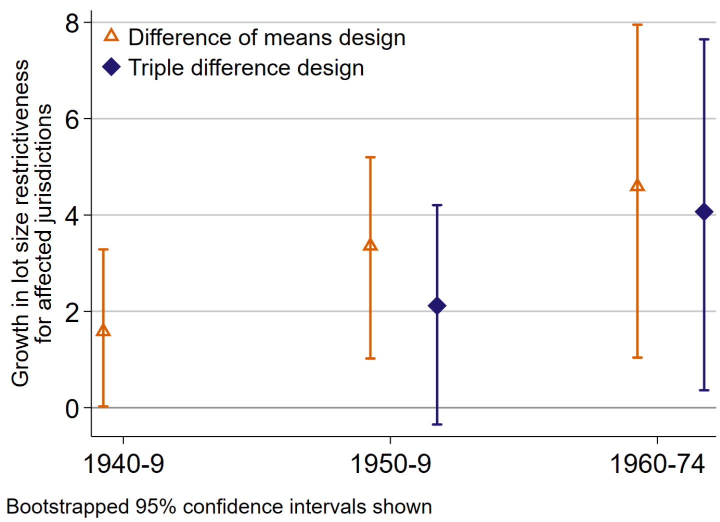
Notes: This figure presents an aggregation exercise, converting the regression coefficients estimated in Table 3 into how much the Second Great Migration explained lot size outcomes in non-Southern metropolitan areas. The aggregation converts Northern Black demographic change into percentile terms and is detailed in Section 5.1. Both panels plot two outcomes: a binary variable for lot size adoption and a ratio for lot size restrictiveness. Panel (a) assumes constant effects across the 1940–70 analysis period, pooling Black migration over three decades to estimate the share of lot size outcomes explained by the Great Migration. Panel (b) estimates a separate causal effect of Black migration for each decade, then conducts the conversion to share of outcome explained for each decade by itself. 95% confidence intervals are bootstrapped using random weights on CBSA-decade clusters, following the Bayesian bootstrap of Rubin (1981).
Sources: Calculations from NHGIS Tables (Manson et al. (2021)), Ruggles et al. (2022), CCDB, IPUMS 1940 full count Census (Ruggles et al. (2021)), Boustan (2016), Derenoncourt (2022) and CoreLogic Tax Records.

Figure 9: Great Migration Effects Strongest With Desegregating Schools

(a) Differences of means across metro area suburbs



(b) Event study estimates using school desegregation cases



Notes: Both panels show estimands measuring the importance of interactions between larger Black migration and metropolitan school desegregation on lot size outcome effects. The "difference of means" estimates are estimated over non-Southern metropolitan areas with above median values of the Great Migration instrument, and is the difference between groups in the subsample split according to the panel labels. Panel (b) splits the subsample between whether the jurisdiction is neighboring a city enacting a court-ordered school desegregation plan during the analysis period. The "triple difference design" estimates further control for common trends among below median Great Migration instrument metros. Further details on the design are in Section 6.
Sources: Desegregation data from Logan, Oakley and Stowell (2008) and Brown University (2005). Calculations from NHGIS Tables (Manson et al. (2021)), Ruggles et al. (2022), CCDB, IPUMS 1940 full count Census (Ruggles et al. (2021)), Boustan (2016), Derenoncourt (2022) and CoreLogic Tax Records.

Table 1: Evaluation of Algorithm on Training and Test Data

	N	Mean	Median	MAE	Pop-weighted MAE
Panel A: Training Set					
Difference in first large lot adoption year	55	0.17	0	10.9	9.89
Difference in first adoption year	76	-0.30	0	9.74	12.0
Panel B: Test Set					
Difference in first large lot adoption year	15	-2.10	-3.0	8.14	5.43
Difference in first adoption year	32	-2.48	-3.5	10.6	5.68
Panel C: Actual Minimum Lot Sizes Matched to Estimates					
Log difference between matched lot sizes	587	-0.195	0	0.42	0.47

Notes: This table reports metrics for model evaluation over the three outcomes matched through optimizing tuning parameters. The mean and median statistics measure unbiasedness. Both are taken over differences between estimated values and values coded in historical zoning records. The mean absolute error (MAE) terms measure fit. The population-weighted MAE uses 2010 population weights on each jurisdiction when calculated the weighted mean of absolute errors. The sample size varies between outcomes, as not all historical zoning records record all three outcomes.

Table 2: Summary Statistics

Variable Name	Analysis Sample				Full Data	
	N	Juri.	Mean	SD	Juri.	Mean
Panel A: Outcome variables						
Minimum lots adopted near end of decade	16282	5435	0.721	0.448	8164	0.708
Restrictive lots adopted near end of decade	16282	5435	0.553	0.497	8164	0.552
Units of housing in decade-long vintage	16282	5435	694.6	2202.7	8164	825.5
Excess mass around minimum lot sizes	13592	5435	5.879	10.13	8164	5.555
Lot size restrictiveness (excess > 7,500 Sq.Ft)	13592	5435	2.982	6.701	8164	2.879
Panel B: Measures of demographic change						
End of decade population, 000s	680	228	146.0	356.8	389	128.8
CC Black composition change	680	228	2.613	4.618	389	4.494
GM shift-share, black migrants	680	228	0.339	0.229	389	0.401
1940 worker share in manufacturing	680	228	0.219	0.0650	389	0.212
1940 Black household head share	680	228	0.023	0.0374	389	0.113
Panel C: 1940 demographic variables						
Census household count		3814	2088.7	4965.3	5190	2069.1
Native born white household head share		4587	0.831	0.114	6686	0.810
Black household head share		4587	0.016	0.046	6686	0.074
Distance from central city CBD, km		4958	29.47	19.21	6851	30.66
Household homeownership rate, %		4587	58.16	11.05	6686	54.32
Median household income, est. 1950 000$		4587	2.633	0.452	6686	2.366
Workers in white-collar industries		4587	0.373	0.134	6686	0.353

Notes: This table reports summary statistics of all variables used in the regression specification of Section 5.1, plus additional demographic variables for context. I report the number of observations over the panel dataset, as well as the number of time-invarying jurisdictions that are included in the data. For variables defined at a central city level, observations are counted at a central city–time period level.

Table 3: Black Demographic Change Causes Lot Size Controls

	OLS		IV	
	(1)	(2)	(1)	(2)
Panel A: Effects on Restrictive Lot Size Adoption				
Percentile of $\Delta CC^{Black}_{c(j)t}$	0.0108	-0.0293	0.0990	0.402*
	(0.0341)	(0.0380)	(0.152)	(0.237)
First-stage F-stat			26.39	20.46
Anderson-Rubin p-value			0.519	0.085
Baseline mean	0.534	0.534	0.534	0.534
Panel B: Effects on Lot Size Restrictiveness				
Percentile of $\Delta CC^{Black}_{c(j)t}$	0.912*	0.682	2.651	4.663**
	(0.524)	(0.543)	(1.731)	(2.278)
First-stage F-stat			26.60	19.51
Anderson-Rubin p-value			0.136	0.042
Baseline mean	2.410	2.410	2.410	2.410
Panel C: Effects on Lot Size Excess Mass				
Percentile of $\Delta CC^{Black}_{c(j)t}$	2.133***	2.864***	6.093**	10.47***
	(0.814)	(0.875)	(2.831)	(3.853)
First-stage F-stat			26.60	19.51
Anderson-Rubin p-value			0.040	0.0076
Baseline mean	5.008	5.008	5.008	5.008
Panel Units	Metro	Metro	Shock	Shock
Panel N	561	561	3418	3418
Region–Decade FE	X	X	X	X
Incomplete shares	X	X	X	X
Pre-period controls		X		X

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

Notes: This table presents regressions of central city Black composition change, as defined in Section 4.1, to lot size outcomes defined in Section 3,

$$Reg_{jt} = \beta \Delta CC^{Black}_{c(j),t} + \delta_{r(j),t} + \mathbf{X}_{j,pre}\Gamma + \varepsilon_{j,c(j)t},$$

using both OLS specifications and instrumenting for Black composition change with shift-share Black migration instruments. In addition, Black composition changes and the migration instruments are normalized to their quantile within each decade's distributions. Across specifications, I add Census region–decade fixed effects and a central city's 1940 exposure to Black migrants over different Southern states, constructing the shares as described in Section 4.3. I also control for decade-specific trends based on cities' 1940 rates of Black households, distance from the CBD, homeownership rates and central city manufacturing worker share. Standard errors are calculated clustering at the CBSA-decade level.

Table 4: Causal Effects of Demographic Changes on Lot Size Restrictiveness

	Δ Black		Δ South. White		Δ Foreigner
	Reduce.	IV	Reduce.	IV	OLS
Panel A: Sample varies by specification					
Central City Demographic Change	1.836**	4.663**	-0.153	-0.753	-0.0542
	(0.858)	(2.278)	(0.0821)	(0.434)	(0.060)
Panel N	3418	3418	3952	3952	556
First-stage F-stat		19.51		14.82	
Anderson-Rubin p-value	0.042	0.042	0.053	0.053	
Baseline mean	2.410	2.410	2.410	2.410	2.410
Panel B: Common Black destination sample					
Central City Demographic Change	3.494*		-0.071		0.0399
	(1.823)		(0.125)		(0.095)
Panel N	3344	3344	3948	3948	236
Baseline mean	2.410	2.410	2.410	2.410	2.410
Panel Units	Shock	Shock	Shock	Shock	Metro
Region–Decade FE	X	X	X	X	X
Incomplete shares	X	X	X	X	X
Pre-period controls	X	X	X	X	X

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

Notes: This table presents regressions of central city composition change on lot size restrictiveness, defined in Section 3, across measures of composition change for multiple demographic groups,

$$Excess_{jt} = \beta \Delta CC^{Dem}_{c(j),t} + \delta_{r(j),t} + \mathbf{X}_{j,pre}\Gamma + \varepsilon_{j,c(j)t}.$$

Columns (1) and (2) repeat reduced form and instrument variable estimates of Black composition change on lot size restrictiveness. Columns (3) and (4) show specifications using Southern white demographic change in central cities, defined by and instrumented with shift-share instruments as defined in Section 5.3. Column (5) define central city composition change on foreign-born White residents in central cities analogously to what I do for Black residents. In addition, I add Census region–decade fixed effects and normalize demographic composition change variables and the migration instruments to their percentile within each decade's distributions. For Black composition change, I add a central city's 1940 exposure to Black migrants over different Southern states, constructing the shares as described in Section 4.3. I also control for decade-specific trends based on cities' 1940 rates of Black households, distance from the CBD, homeownership rates and central city manufacturing worker share. Standard errors are calculated clustering at the CBSA-decade level.

Table 5: Robustness of Migration's Effects on Lot Size Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Effects on Restrictive Lot Size Adoption							
Pctile. $\Delta CC^{Black}_{c(j)t}$ (IV)	0.402*	0.431*	0.683**	0.331**	0.978***	0.414*	
	(0.237)	(0.239)	(0.349)	(0.156)	(0.258)	(0.236)	
$\Delta CC^{Black}_{c(j)t}$ (IV)							0.116*
							(0.699)
First-stage F-stat	20.46	20.04	9.475	22.29	20.67	20.46	23.94
Anderson-Rubin p-value	0.0853	0.0681	0.0410	0.0138	0.000	0.0747	0.0853
Panel B: Effects on Lot Size Restrictiveness							
Pctile. $\Delta CC^{Black}_{c(j)t}$ (IV)	4.663**	5.115**	3.680	4.760	6.170***	4.249*	
	(2.278)	(2.308)	(4.006)	(3.355)	(1.874)	(2.539)	
$\Delta CC^{Black}_{c(j)t}$ (IV)							1.358**
							(0.641)
First-stage F-stat	19.51	19.09	8.770	22.16	19.94	19.63	22.56
Anderson-Rubin p-value	0.0417	0.0273	0.353	0.166	0.000	0.0862	0.0417
Panel N	3418	3418	3403	3402	3418	3418	3418
Metros in sample	561	561	468	495	561	561	561
Main specification controls	X	X	X	X		X	X
Southern white IV control		X					
CC density control			X				
Most anomalous bunching				X			
Jurisdiction FE model					X		
5-year lagged demographics						X	
No pctile transformation							X

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

Notes: This table presents alternate specifications showing results in Table 3 are robust across regression specifications as well as outcome misspecification. Column (1) displays again the OLS and IV results of Black composition change on lot size restrictiveness,

$$Excess_{jt} = \beta \Delta CC^{Black}_{c(j),t} + \delta_{r(j),t} + \mathbf{X}_{j,pre}\Gamma + \varepsilon_{j,c(j)t}.$$

In addition, Black composition changes and the migration instruments are normalized to their percentile within each decade's distributions. Columns (2) and (3) report robustness to changes in right-hand side controls in the specifications. Columns (4) to (6) report robustness to definitions and scale of observation for the lot size restrictiveness outcome. Column (7) reports a specification using levels of Black composition change not normalized by the within-decade percentile function. Details of the specifications are described in Section 5.4. Main specification controls include a central city's 1940 exposure to Black migrants over different Southern states. They also include decade-specific trends based on cities' 1940 rates of Black households, distance from the CBD, homeownership rates and central city manufacturing worker share. Standard errors are calculated clustering at the CBSA-decade level.