NBER WORKING PAPER SERIES

THE DISTASTE FOR HOUSING DENSITY

Joseph Gyourko
Sean E. McCulloch

Working Paper 33078
http://www.nber.org/papers/w33078

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2024, Revised July 2026

We thank Braydon Neiszner and Vivian Wang for helpful research assistance. We also appreciate the comments of Hector Blanco, Sophie Calder-Wang, Jonathan Dingel, Gilles Duranton, Fernando Ferreira, Ed Glaeser, Peter Hull, Brad Ross, Evan Soltas, Matthew Turner, and seminar participants at the University of Wisconsin, the 2024 UEA North American conference, the Urban Lunch Seminar at Wharton, Brown University, and the 2025 NBER Summer Institute. Naturally, we remain responsible for the final product. Gyourko thanks the Research Sponsor Program of the Zell/Lurie Real Estate Center for financial support, and McCulloch thanks the Population Studies and Training Center at Brown University, which receives funding from the NIH, for training support (T32 HD007338) and for general support (P2C HD041020). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Distaste for Housing Density
Joseph Gyourko and Sean E. McCulloch
NBER Working Paper No. 33078
October 2024, Revised July 2026
JEL No. R0, R31

ABSTRACT

Research has established local regulations limiting housing unit density have contributed to growing affordability problems across many housing markets. Less well understood is the value existing residents place on low neighborhood density. We estimate suburban homeowners' preferences for density by regressing house prices on a localized measure of density exposure. For identification, we derive a formula instrument based on a house's distance to a border with a discontinuity in minimum lot sizes, drawing on recent advances in applied econometrics. We find adding 100 houses within 500 meters decreases a house's value by 1 to 2% on average, though effects vary widely across neighborhoods. Linking to a hedonic model, these estimates suggest the typical preference is moderately negative, but tastes differ substantially across households. These results provide an important foundation for understanding the demand for density regulation across U.S. suburbs.

Joseph Gyourko
University of Pennsylvania
and NBER
gyourko@wharton.upenn.edu

Sean E. McCulloch
Brown University
sean_mcculloch@brown.edu

A data appendix is available at http://www.nber.org/data-appendix/w33078

1 Introduction

Real constant-quality house prices are at all-time highs in many markets, sparking widespread concern about housing affordability.[1] A large literature points to the impact of local government land use restrictions on building activity to explain this phenomenon (Glaeser and Gyourko, 2018). Even with the growing salience of the affordability issue, relatively few jurisdictions have changed zoning or building permit approval rules to allow more development and the rate of increase in new supply has been tepid at best in many markets. That there are not many more communities implementing such obvious policy responses begs the question of why not. A straightforward answer is that local homeowners do not favor such a policy and can stop new construction and densification through the political process. This paper's primary goal is to estimate how much suburban homeowners are willing to pay to avoid density in their neighborhood.

One possible approach to estimate preferences for neighborhood density is a hedonic regression. To fix ideas, consider estimating the average marginal willingness to pay (MWTP) for neighborhood density by the following equation:

$$p = \alpha + \beta D + \epsilon, \tag{1}$$

where p is the house price, D is neighborhood density, and ϵ is the error term. Under ideal conditions, a regression estimate of β recovers the average homeowner MWTP.

However, we typically suspect a hedonic regression might suffer from several identification issues. One is the standard problem of omitted variable bias, with a typical concern being that wealthy suburbs have both low density and other desirable amenities that would contaminate an OLS estimate. A standard solution is to use instrumental variables or some other quasi-experimental method. In an early application, Chay and Greenstone (2005) raise a second issue of external validity in the context of hedonic IV. If there is heterogeneity in preferences[2] and people sort across neighborhoods in a way correlated with the instrument, then an IV estimate need not be representative of the average MWTP. For example, the instrument could reflect policy variation in zoning regulation and the people who chose to live in neighborhoods induced into low density by binding zoning regulation may dislike density more than the overall population.

[1] Real constant quality prices are about 15 percent above their peak just price to the Global Financial Crisis (GFC). The same holds for many major metropolitan areas across the country. See Figure 5 and Appendix Figure 1 in Glaeser and Gyourko (2025) for more detail.

[2] In a Pew Research Center survey, 55% of Americans said they would prefer to live in a community where "houses are larger and farther apart, but schools, stores and restaurants are several miles away", while 44% said they would not. PewSurvey. Other survey-based research also suggests wide variation in preferences regarding density, as well as how to compensate local residents for allowing its increase (Davis et al., 2026).

To address the omitted variable bias issue, we derive a formula instrument that predicts exposure to housing density based on distance to a border with a discontinuity in minimum lot sizes. Intuitively, density exposure falls approaching a neighborhood with a larger minimum lot size. We show variation in density exposure due to a regulatory discontinuity follows a known formula, determined by a house's distance to the border and the magnitude of the discontinuity. This reveals a previously unrecognized connection between similar identification strategies for zoning externalities in the prior literature (Turner et al., 2014) and Borusyak and Hull (2023)'s econometric framework for formula instruments. Other instrument constructions using distance to the border and the magnitude of the regulatory discontinuity, i.e. the external effect in Turner et al. (2014), fit in this class of instruments. These instruments are valid if, within a small geographic area, distance to the border is uncorrelated with unobservables, conditional on fixed effects for each side of a municipal border (and other controls).

We then develop another approach that does not require distance to the border to be exogenous by using a "recentering" procedure following Borusyak and Hull (2023). For intuition, imagine two towns, A and B, share a border and set different minimum lot sizes on each side of the border. For simplicity, assume these minimums fully determine houses' actual density exposures and town B always sets a larger minimum. Hence, density exposure decreases approaching town B. For the recentering strategy to be appropriate, for a house in town A, the minimum set by town B must be an exogenous shock (conditional on controls), with variation in the exact size of the minimum. For example, say town B could set either a 1/2 acre minimum or a 1 acre minimum. Knowing these counterfactuals, instrument validity can be restored by adjusting for the density exposure we would typically expect at a house's distance to town B. In this example, if a house in town A close to the border happened to be next to a 1/2 acre minimum instead of a 1 acre minimum, it should have a high instrument value because it has higher exposure than we would expect given how close it is to town B—despite the actual density exposure decreasing as we approach town B. As we will show, the density exposure recentered by the expected exposure is a valid instrument even when distance to the border is correlated with unobservables.

We find conventional TSLS/IV estimates vary substantially across constructions of the formula instrument and are more negative than a corresponding OLS estimate. TSLS/IV estimates range from -3 to -7% of house value, compared to an OLS estimate of -1%. Interpreted as a constant effect, these results yield the potentially puzzling conclusion that the OLS estimate actually understates the effect of density exposure on price. While this could suggest a failure in instrument validity, in a LATE framework this pattern could also be consistent with each instrument generating a different compliant subpopulation, whose LATEs are internally valid but not externally valid to the sample ATE. For example, the instrument has more power in neighborhoods where minimum lot size regulations are more binding. If the residents who sort into neighborhoods with binding zoning regulation also tend to dislike density exposure more, then we might expect an IV estimate to be more negative than the average effect.

In response, we consider the issue of external validity by providing two other types of IV estimators that target the sample ATE under additional assumptions. The first estimator imposes the assumption that price effects are heterogeneous across neighborhoods, but homogeneous within neighborhoods. This is consistent with a model where sorting determines price functions at the neighborhood level. Given this assumption, we estimate the ATE by partitioning the sample into neighborhoods, calculating a TSLS/IV estimate for each neighborhood, and then averaging the neighborhood-specific estimates weighted by the number of observations (Angrist and Fernández-Val, 2013). Under this approach, estimates for the different formula instruments range from -1 to -2%, a narrower range than the analogous conventional TSLS/IV estimates. Neighborhood-specific estimates suggest a wide distribution of price effects across neighborhoods. The interquartile range of price effect estimates is -7% to +4%, with a median that is also around -1%.

The second IV estimator assumes a model of selection into treatment and estimates the ATE by a control function approach. We follow the estimation routine outlined in Masten and Torgovitsky (2016). Under this approach, we again find estimates for the ATE that range from -1 to -2%. Thus, using two distinct estimators that impose different assumptions but target the same causal parameter, the sample ATE, we find similar results.

We use a conservative estimate of -1% to consider how much a homeowner would be willing to pay to avoid density. Several studies have noted the challenge of disentangling whether a price decrease from higher density is due to a localized "supply effect" from increasing housing unit quantity or a "disamenity effect" from lowering neighborhood amenities (e.g. Pennington, 2021, Asquith et al., 2023, Blanco and Sportiche, 2026). Although the heterogeneity in impacts across neighborhoods appears more consistent with the latter, under the conservative assumption that the entire estimated effect is due to a supply effect, an incumbent homeowner would still be willing to pay $4,000 to avoid the associated property value loss. On the other hand, interpreting the full -1% as a disamenity effect in a hedonic model, we estimate the average combined asset and amenity value loss to a homeowner to be at least -$8,000. Nearby density is a disamenity to most owners and should be taken into account in any welfare analysis. While we leave that complex exercise for future research, we note in the penultimate section that a -1% price externality is small enough for extant density still to be below the level that maximizes land value in the typical American suburb, not just in high-priced coastal markets such as Boston and the Bay Area that have long been thought to be supply constrained (Glaeser and Ward, 2009).

In sum, our empirical strategy effectively isolates price effects due to changes in the “character" of a neighborhood by forming comparisons within the same town, thereby holding fixed public goods provision that changes across towns and the direct effects of zoning on house size. This total effect of density exposure on price within a neighborhood is the number that explains why homeowners typically oppose even marginal increases in local density. As such, it captures the policy-relevant variable that is vital to understanding NIMBYism. What is does not do, of course, is identify the many possible factors that mediate how density is experienced. Obviously, density could affect neighborhood character through a variety of channels. Historically, zoning regulations were often intentionally adopted to induce sorting on desired demographic characteristics—racial segregation being a prominent example (Rothstein, 2017). Other possible mediating factors include increased traffic, noise, crime, and access to local businesses. While we briefly explore two of these potential mechanisms in the penultimate section of the paper, decomposing the average price effect across these channels would require additional instruments and assumptions that are beyond the scope of this already lengthy paper.[3] Identifying and estimating the various potential mediating factors clearly is a next step for research in this area.

A *Relation to the literature*

Our paper adds to a large literature on zoning and housing regulation by providing estimates of how much homeowners value lower neighborhood density. Recent research has focused on how land use restrictions increase the cost of housing by restricting local market housing supply (Glaeser and Gyourko, 2018, Gyourko and Krimmel, 2022) or decrease aggregate output at the national level by limiting the size of the overall housing stock in a market (Herkenhoff et al., 2018, Duranton and Puga, 2023). While there is now significant evidence on the cost of land use regulation at the market or national levels, Baum-Snow (2023) notes there is much less evidence about how local amenities would be affected if these neighborhoods did experience densification. This paper directly addresses that critique.

[3]See Huber (2019) or Roth and Kwon (2026) for more discussion on causal mediation analysis.

A growing literature measures localized price effects from new construction (Diamond and McQuade, 2019, Pennington, 2021, Davidoff et al., 2022, Asquith et al., 2023, Blanco and Sportiche, 2026), with Davidoff et al. (2022) and Blanco and Sportiche (2026) considering new development that explicitly raises the existing density level. In contemporaneous work, Kashner and Ross (2025) study the hedonic value of traffic exposure, emphasizing the close connection between traffic and density. Relative to these studies, we evaluate the hedonic value for a direct measure of density exposure for a national sample of suburban homeowners.

We make a methodological contribution to boundary discontinuity designs by studying identification of spillovers away from the boundary. Originating with Black (1999), many studies employ boundary discontinuities to evaluate the impact of a policy on house price or other outcomes. Within the zoning literature itself, there are numerous applications (Turner et al., 2014, Severen and Plantinga, 2018, Song, 2025, Kulka et al., 2026). Most studies focus on the effect measured by the discontinuity at the boundary rather than a policy's possible spillovers, perhaps because identification moving farther away from the boundary is more tenuous, especially in the context of house sales where the nearby land is already developed. Most closely related, Turner et al. (2014) derive a model-implied expression for the spillovers from a discontinuity in land use regulation, which they term the *external effect*. We connect this type of model-implied expression to Borusyak and Hull (2023)'s econometric framework for formula instruments. This link clarifies the assumptions needed to identify the spillover effect and suggests a way to weaken the identifying assumptions on distance to the boundary. More generally, this connection might be useful in other RD settings when there are multiple cutoffs where the spillover strength depends on the discontinuity's magnitude.

We also revisit an idea introduced by Chay and Greenstone (2005) that sorting into neighborhoods according to heterogeneous preferences can lead to a difference between an IV estimand and the average marginal willingness to pay. Those authors also report estimates from both a conventional IV estimator and a control function estimator. In contrast, we also operationalize an approach to estimate the average treatment effect following Angrist and Fernández-Val (2013). In Chay and Greenstone (2005)'s application to pollution, they find omitted variable bias is a more important consideration than treatment effect heterogeneity. For the valuation of density exposure, we find the opposite. In light of renewed interest in the IV estimand's interpretation in the presence of heterogeneous treatment effects (Mogstad and Torgovitsky, 2024, Blandhol et al., 2025), this suggests applied urban researchers, who frequently study settings with sorting, should carefully consider how conventional quasi-experimental estimators aggregate individual causal effects.

Heterogeneity in willingness to pay seems likely to be relevant for how building requests are handled on the ground. Our $8,000 average impact may not be the most relevant figure if sorting is associated with people in low density communities having a much higher distaste for density. Compensating those residents may be cost prohibitive, and could help explain why such places never densify. It also could be that having just a few of those types in higher density communities is enough to generate costly opposition to further densification in those places, too. Consequently, we view the full spectrum of results as potentially important inputs into future research that examines the political economy of the building approval process. Our results also should be a helpful input into broader welfare analyses that take account of the losses of those who would have benefited from new housing had it been built.

The plan of this paper is as follows. Section 2 posits a border model that describes variation in density exposure around a municipal border where minimum lot size restrictions shift. Section 3 lays out the econometric framework. Section 4 describes the data. Section 5 presents results. Section 6 discusses and Section 7 concludes.

2 Theory

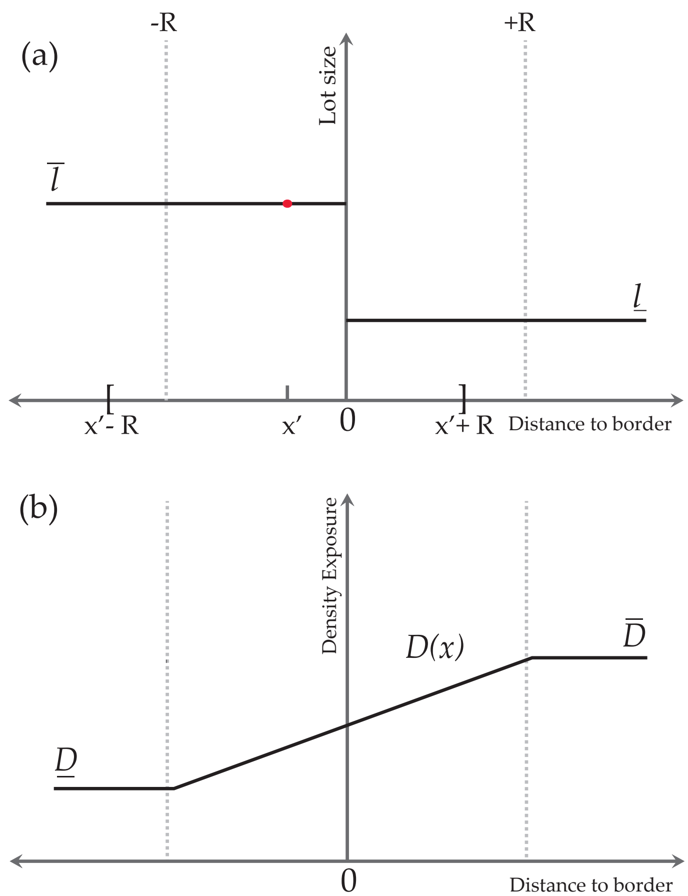
We begin by positing a simple model of how density exposure varies in relation to a municipal border with a change in minimum lot size restrictions. This leads to a prediction of a house's density exposure as a function of the lot size restrictions and distance to the municipal border. In Section 3, we then discuss how we use this predictable variation to identify the hedonic relationship between price and density exposure introduced in (1).

A *Border model*

We start with a simple model similar to Turner et al. (2014). There is one border between two municipalities in 1-dimensional space x where the boundary is at $x = 0$. The minimum lot size on the left side is $\bar{l}$ and the minimum on the right is $\underline{l}$. The left side of the border is more highly regulated than the right side, $\bar{l} > \underline{l}$. For this example, the minimum lot sizes are binding, so at each point lot size equals the minimum on that side of the border. Figure 1(a) plots houses' lot sizes as a function of distance to the border where lot size drops discontinuously at the border.

We depart from Turner et al. (2014) by constructing an explicit parametrization of the external effect, the spillovers from the discontinuity in land use regulation at the boundary. We define a measure for how exposed a house is to neighborhood housing density at any location x', which we term the density exposure. Let R be a constant bounding the spatial extent at which spillovers from housing density are relevant to individual homeowners. Figure 1(a) shows the spillover extent for the location x'.

Figure 1: Parametrizing density spillovers near a lot size discontinuity



Note: This figure illustrates the parametrization of the spillover, the density exposure, based on the lot size discontinuity. Panel (a) plots lot size as a step function that decreases discontinuously at the boundary entering the less regulated municipality. Panel (b) plots the density exposure from the lot size discontinuity based on the formula (2).

The density exposure is the number of housing units within R distance. For example, $D(x')$ could be the number of housing units within $R = 500$ meters of the house at location x'. For a point x' that is exposed to both sides of the border, the density exposure is determined by how much of the interval $[x' - R, x' + R]$ is on the left versus the right side of the border:

$$D(x') = \frac{R - x'}{\overline{l}} + \frac{x' + R}{\underline{l}}$$

where the first term on the right hand side reflects the length of the exposure interval that falls in the left municipality and the second term reflects the length in the right municipality. The density exposure as a function of x is

$$D(x') = \begin{cases} 2R/\bar{l} = \underline{D} & \text{if } x' < -R, \\ \frac{R-x'}{\bar{l}} + \frac{x'+R}{\underline{l}} & \text{if } x' \in [-R, R], \\ 2R/\underline{l} = \bar{D} & \text{if } x' > R. \end{cases} \tag{2}$$

Panel (b) of Figure 1 plots the relationship of density exposure on x. Density exposure is mechanically related to the lot size, and the discontinuity at $x = 0$ induces changes in $D(x)$. Density exposure is flat until $x \in (-R, R)$ where exposure to the other side of the border induces a slope change. Moving from left to right, density exposure increases continuously through the border as houses are located closer to and then further into the municipality with the smaller minimum lot size.

B *Moving to 2-d space*

In reality, houses' geographic locations exist in 2-dimensional space. We can still write a density exposure formula in terms of the distance to the border and the minimum lot sizes, but we need to be more explicit about how x is defined. Let x^{rad} be the radial distance to the border on either side of the border. Set $x = -x^{rad}$ for a point in space in the more regulated municipality and $x = x^{rad}$ for a point in space in the less regulated municipality. Thus, x is the signed distance as density increases moving left to right along the x-axis as in Figure 1.

A similar expression to (2) holds in 2-d space based on the area of a circle rather than the length of an interval. Define $a(x)$ as the area of the circle within the left side of the border. For a given x on the right side of the border in $(0,R)$, the density exposure is

$$D(x) = \frac{a(x)}{\bar{l}} + \frac{\pi R^2 - a(x)}{\underline{l}} \tag{3}$$

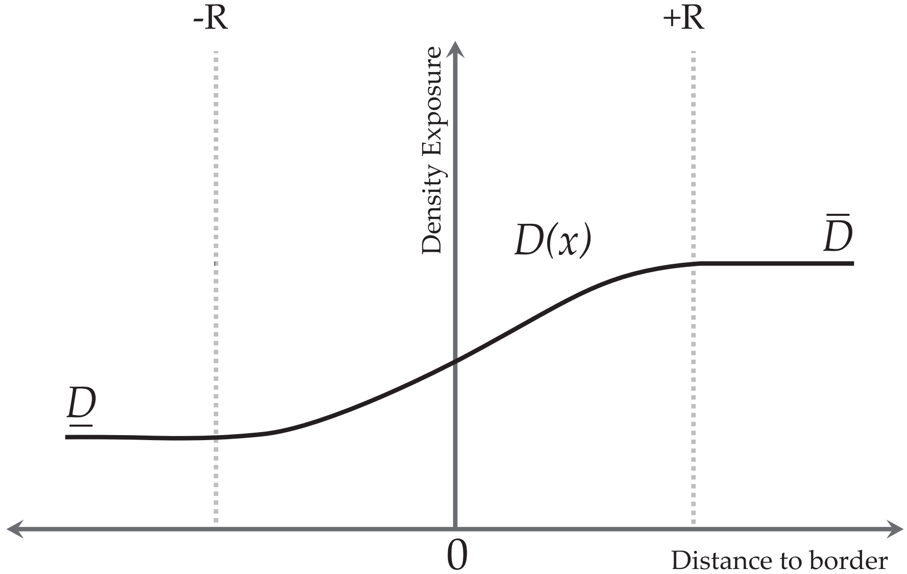
where, using the formula for the area of a circular segment,

$$a(x) = R^2 \arccos(\frac{x}{R}) - x\sqrt{2R(R-x) - (R-x)^2}. \tag{4}$$

Thus, the density exposure formula is similar to the 1-dimensional case, increasing continuously through the border, but is now a nonlinear function of distance to the border and the minimum lot sizes.[4] Figure 2 illustrates the density exposure gradient on x predicted by formula (3).

[4]Appendix D1 provides more details on deriving (4) from the area of a circular segment.

Figure 2: 2-d density exposure on x



Note: This diagram plots density exposure on x based on the formula (3).

C *Multiple borders and border segments*

To map the model to the empirical setting, we introduce notation that allows for multiple municipal borders. Let $b(j)$ index a municipal border area that a house j is nearby. Also, denote $bm(j)$ as the border area$\times$municipality in which house j is located. This is one side of a border in the same municipality.

Further, we will allow there to be multiple zoning districts with different lot size restrictions within the same municipality. Next define $s(j)$ as a segment of the border to which house j is closest. Finally, we will distinguish between each side of a border segment pair. Let $sm(j)$ denote the border segment$\times$municipality in which house j is located. This is one side of a border in the same municipality along a portion of the border. Similarly, define $-sm(j)$ to refer to the border segment$\times$municipality on the opposite side of house j's border segment. Thus, $\underline{l}_{sm(j)}$ refers to the minimum lot size restriction in the municipality on house j's side of a border segment pair and $\underline{l}_{-sm(j)}$ refers to the lot size restriction on the opposite side. These within municipality zoning districts create border "segments" along the same border where the discontinuities in lot size across the border vary in size.

The border model and equation (3) is symmetric with respect to the two sides of the border, so we can rewrite the expression relative to each house j's side of the border $sm(j)$ by using the radial distance to the municipal border x_j^{rad}. For $x_j^{rad} \in (0, R]$, we have

$$f(x_j^{rad}, \underline{l}_{sm(j)}, \underline{l}_{-sm(j)}) = \frac{a(x_j^{rad})}{\underline{l}_{-sm(j)}} + \frac{\pi R^2 - a(x_j^{rad})}{\underline{l}_{sm(j)}} \tag{5}$$

where the area on the other side of the border $a(.)$ is defined as in (4).

D *The formula as a prediction of the actual density exposure*

Thus far, we have imagined there is only variation in lot sizes and densities from the minimum lot size restrictions and distance to the border. Now, we allow houses to have individual level variation in lot sizes and for there to be variation in density exposure not due to the lot size restrictions.

A house $j \in \mathcal{J}$ has a geographic location s_j that maps to its radial distance to the border x_j^{rad}. An individual house also has an individual lot size l_j and density exposure D_j. Lot sizes can be any value but are constrained by the restriction that $l_j \geq \underline{l}_{sm(j)}$. Define the buffer with radius R around s_j by $B(s_j)$. The number of houses in a buffer $B(s_j)$ is $N_{B(s_j)}$. Finally, denote $U_{B(s_j)}$ as the share of the buffer that is covered in housing with the remainder of the buffer not covered for whatever reason—e.g., roads, open spaces, commercial development, and so on. We can write the actual density exposure from single family housing for house j as

$$D_j = \frac{\pi R^2}{\frac{1}{N_{B(s_j)}} \sum_{k \in B(s_j)} l_k} U_{B(s_j)} = \frac{\pi R^2}{\bar{l}_{B(s_j)}} U_{B(s_j)} \tag{6}$$

where the denominator is the average lot size among houses within a buffer with radius R.

We can decompose this expression into the two sides of the border. Denote $B_{opp}(s_j)$ as the portion of the density exposure buffer overlapping the opposing side of the border relative to the location s_j, and $B_{own}(s_j)$ as the portion contained on the same side of the border. Denote analogous average lot sizes $\bar{l}_{B_{own}(s_j)}, \bar{l}_{B_{opp}(s_j)}$ and shares developed $U_{B_{own}(s_j)}, U_{B_{opp}(s_j)}$ for each side of the border. (6) can be rewritten so the actual density exposure is in terms of the two sides of the border:

$$D_j = \frac{a(x_j^{rad})}{\bar{l}_{B_{opp}(s_j)}} U_{B_{opp}(s_j)} + \frac{\pi R^2 - a(x_j^{rad})}{\bar{l}_{B_{own}(s_j)}} U_{B_{own}(s_j)}. \tag{7}$$

Replacing the denominators with the minimum lot size restrictions and setting $U_{B_{own}(x_j)}$ and $U_{B_{opp}(x_j)}$ to 1 yields Equation (5). Thus, Equation (5) can be viewed as a simulated instrument for the actual density exposure, forming a prediction of the density exposure that only uses the variation from radial distance to the border and the minimum lot sizes. In the next section, we discuss when Equation (5) is a valid instrument to identify the relationship between price and density exposure.

3 Econometric Framework

A *Setup*

Consider a linear hedonic regression of price on density exposure. A house j's price p_j is determined by its density exposure D_j and its lot size l_j:

$$p_j = \beta D_j + \tau l_j + A_{bm(j)} + W_j' \delta + \epsilon_j. \tag{8}$$

Here, β is the effect from a higher density exposure on price and τ is the effect from a larger lot size. To start, we consider β as a constant effect. We relax the treatment effect homogeneity assumption below. $A_{bm(j)}$ are geographic fixed effects specific to one side of a border area and W_j is a vector of observable housing characteristics. To account for the direct effect of a lot size restriction on price, we include fixed effects for each side of the border in all regression results, removing the level shift in prices from crossing municipal boundaries, and control for lot and house size.[5] Finally, ϵ_j is an error term.

The estimated coefficient for β in an OLS regression only identifies the causal effect on price if density exposure is uncorrelated with ϵ_j. However, suburbs with lower density exposure may also be wealthy and happen to have other unobservable amenities. While we can (and do) include granular geographic fixed effects as controls, by construction it is impossible to control perfectly for geography without absorbing all variation in density exposure.

We want to consider the formula in Equation (5) derived in the previous section as an instrument z to identify β. For a house $j \in \{1, 2, \ldots, J\}$, define $z_j := f(x_j^{rad}, \underline{l}_{sm(j)}, \underline{l}_{-sm(j)})$, which takes as arguments a house j's radial distance to the border and the minimum lot sizes on either side of its border segment. Note also that other simpler possible instruments, including those from the prior zoning literature, take the same variables as arguments. One natural choice is to imagine the x-axis in Figure 1 as an instrument. While x may appear to be just one variable, notice that x implicitly uses the minimum lot sizes to assign directionality to the x-axis; that is, for a house j, we have $x_j = x_j^{rad}\,\text{sgn}(\underline{l}_{sm(j)} - l_{-sm(j)})$. Hence, $z_j^{sgn} := x_j$ is also a formula instrument taking the arguments $x_j^{rad}, \underline{l}_{sm(j)}$, and $l_{-sm(j)}$. Another instrument is suggested by Turner et al. (2014). For their external effect regressions, those authors construct a

[5]While the direct effect is not this paper's primary interest, an interesting note is that this structure is analogous to spillovers on a social network. For example, Miguel and Kremer (2004) study deworming medication's effect on children's health outcomes. In this setting, a child experiences an indirect effect from the number of their peers who get dewormed and a direct effect from that child herself being dewormed. In our setting, a housing unit experiences an indirect effect from the lot sizes of its neighbors and a direct effect from the housing unit's own lot size. Instead of a social network, geographic space links houses together.

variable interacting a distance dummy $\mathcal{X}_j$ with the regulation discontinuity's magnitude, $z_j^{ext} := \mathcal{X}_j(\underline{l}_{sm(j)} - l_{-sm(j)})$. Here, $\mathcal{X}_j$ equals 1 if a house is far away from the border and 0 if a house is close to the border. However, because these formula instruments combine multiple sources of variation it is not obvious when they will be valid instruments.

To clarify identification, we relate z to the recent literature on shift-share and formula instruments (Goldsmith-Pinkham et al., 2020, Borusyak et al., 2022, Borusyak and Hull, 2023). In this context, the size of the discontinuity at a border is similar to the "shift" and a house's distance to that border the "share" in a shift-share design. More formally, we will consider z in the context of Borusyak and Hull (2023)'s econometric framework for formula instruments. The starting point is partitioning $f(.)$'s arguments into exogenous shocks g versus the other variables w that govern exposure to the shocks. The underlying intuition in this design, and in similar designs in the prior zoning literature, is that while zoning regulation within a house's own town is presumably endogenous, the restriction set by an adjacent town is plausibly exogenous (conditional on controls). Below, we formalize this assumption precisely. To emphasize $\underline{l}_{-sm(j)}$ as the exogenous shock, henceforth we denote $g_{-sm(j)} := \underline{l}_{-sm(j)}$. In this application, the other variables collected in w are x_j^{rad},$\underline{l}_{sm(j)}$, and $sm(j)$.

Viewed from the perspective of Borusyak and Hull (2023)'s framework, many of specification and sampling decisions in the prior literature are ways to address the potential importance of nonrandom exposure due to either $\underline{l}$ or x_j^{rad}. For example, Turner et al. (2014) bolster the credibility of their estimates in a number of ways. They saturate the geographic fixed effects at the level of their regulatory measure, effectively removing level shifts due to $\underline{l}_{sm(j)}$. They restrict to straight line boundaries whose location may be drawn due to bureaucratic happenstance rather than other geographic factors. In addition, they assess sensitivity to including various sets of controls for physical geography that we may expect to be correlated with distance to the border x_j^{rad}.[6]

B *Identification and instrument recentering*

We depart from the prior zoning literature by estimating β using an identification argument relying on exogeneity of the shocks as defined by the minimum lot sizes set by a different town. We consider the observed lot size discontinuities across different segments of the same border as a realization of conditionally exogenous shocks. Instead

[6]Hence, the prior zoning literature has implicitly argued for identification by maintaining that exposure is exogenous, subject to sample restrictions and conditioning on observable covariates. In the empirical results, we also control for $\underline{l}_{sm(j)}$, restrict to straight line boundaries, and test sensitivity to including controls for physical geography.

of assuming distance to the border is orthogonal with the error term, we assume we can model the shock assignment process using nearby minimum lot size restrictions as counterfactuals. Given a correct specification of counterfactual shocks, we can restore instrument validity by an appropriate "recentering" of the candidate instrument.

We formalize our main assumption for identification following Borusyak and Hull (2023).

Assumption 1 *(Shock exogeneity):* $\mathrm{E}[\epsilon_j|g,w] = \mathrm{E}[\epsilon_j|w]$ *a.s. for all* j.

This says that, conditional on the observed variables w, the outcome errors are mean-independent from the lot size restriction set by an adjacent town, $g_{-sm(j)}$. Effectively, this imposes two restrictions. First, the minimum lot size set by a different, adjacent town only impacts a house's price through the model-implied spillover, i.e., an exclusion restriction. Second, this minimum is drawn as-good-as-randomly from some set of possible minimums. Note, however that we will not assume that any minimum lot size could arise; for instance, this pool may include as potential policies only the minimums set within the same zoning authority, as we formalize below. In addition, we require instrument relevance, i.e., density exposure is affected by minimum lot size restrictions.

Candidate Instrument: The formula instrument in Equation (5) is the candidate instrument. Repeating Equation (5),

$$z_j = \frac{a(x_j^{rad})}{g_{-sm(j)}} + \frac{\pi R^2 - a(x_j^{rad})}{\underline{l}_{sm(j)}} \quad \text{for } x_j^{rad} \in (0,R) \tag{9}$$

where $a(x_j^{rad})$ is the area of a house's density exposure circle contained within the other side of the municipal border.

Expected instrument: We define the *expected instrument* $\mu_j = \mathrm{E}[z_j|x_j^{rad},\underline{l}_{sm(j)},sm(j)]$ as the average value of z_j across different realizations of the shocks, conditional on a house's radial distance to the border x_j^{rad}, the lot size restriction set in its own zoning district $\underline{l}_{sm(j)}$, and its border segment and side of the border $sm(j)$. By the definition of μ and Assumption 1, we have

$$\mathrm{E}[\frac{1}{N}\sum_j z_j\epsilon_j] = \mathrm{E}[\frac{1}{N}\sum_j \mu_j\epsilon_j].$$

If we knew μ_j, we could adjust z_j by recentering: $\tilde{z}_j = z_j - \mu_j$. Then, by construction, the recentered instrument $\tilde{z}_j$ satisfies an IV moment condition:

$$\mathrm{E}[\frac{1}{N}\sum_j \tilde{z}_j\epsilon_j] = \mathrm{E}[\frac{1}{N}\sum_j z_j\epsilon_j] - \mathrm{E}[\frac{1}{N}\sum_j \mu_j\epsilon_j] = 0.$$

If the recentered instrument $\tilde{z}_j$ is also relevant, we can identify β using $\tilde{z}_j$ in an IV regression. Alternatively, controlling for μ_j in an IV regression that uses the unadjusted z_j as the instrument would also satisfy the orthogonality condition because controlling for μ_j implicitly recenters z_j. Hence, our goal is to construct μ_j.

Borusyak and Hull (2023) propose computing the expected instrument by (i) specifying an assignment process, (ii) drawing sets of counterfactual shocks from this process, recomputing the candidate instrument each time, and (iii) averaging the instrument across the counterfactuals. In the next subsection, we describe a thought experiment to motivate counterfactual minimum lot sizes.

C *Counterfactual shocks*

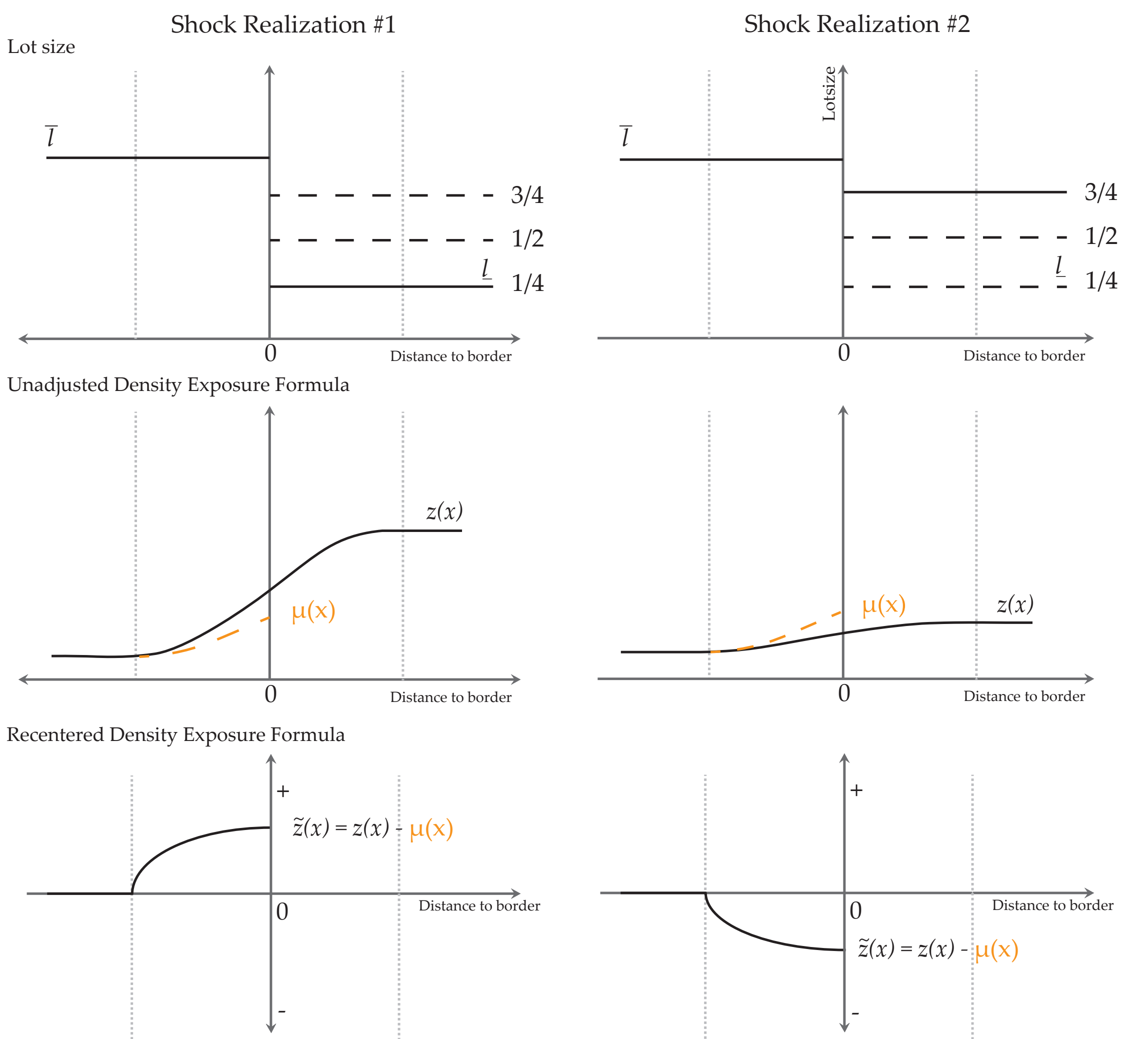
In our context, we have variation in the magnitude of the lot size discontinuities along borders. To identify the density exposure's price effect, we will imagine that for a house on one side of a border there is some as-good-as-random variation in the exact lot size restrictions set by the municipality on the opposite side of a border.

We provide a stylized example to provide intuition for the identifying variation. As before, consider a border where the community on the left side has a larger minimum lot size restriction as compared to the one on the right side. Imagine there are two types of towns: L and R. The left municipality is type L and the right municipality is type R. A type L town always sets the minimum lot size at 1 acre. A type R town sets smaller lot size restrictions but chooses randomly between $\frac{3}{4}$, $\frac{1}{2}$ and $\frac{1}{4}$ acre restrictions with equal probability.

Figure 3 illustrates this scenario under two counterfactual shock realizations. The left column shows a shock realization where the R town set its minimum to its smallest option of $\frac{1}{4}$ acre. The right column shows a realization where the R town set its minimum to its largest option of $\frac{3}{4}$ acre. In this figure, we define distance to the border x as in Figure 1 where we assign houses in the more regulated municipalities to have negative distance. The three rows or panels in the figure each plot an outcome on distance to the border: The top one is for lot size restrictions. The middle or second one is the formula prediction of the density exposure based on Equation (5). The third and final plot is of the recentered instrument.

More specifically, the top panel plots lot size restrictions on x. The solid lines are realized lot sizes. In each shock realization scenario, the gray dotted lines show other counterfactual lot size restrictions that the R-type town could have set but did not. As noted above, a L town always sets the minimum lot size at 1 acre. In the left column under Shock Realization #1, the R town set its minimum lot size to $\frac{1}{4}$ acre. In the right column under Shock Realization #2, the R town set its minimum to a larger lot size of $\frac{3}{4}$ acre.

Figure 3: Stylized example for the recentered instrument



Note: This figure is a stylized scenario to illustrate instrument recentering. The left column shows a shock realization where the right municipality set its lot size restriction to its smallest option at $\frac{1}{4}$ acre. The right column shows a realization where the right municipality set its lot size restriction to its largest option at $\frac{3}{4}$ acre. The three rows each plot an outcome on distance to the border. First, the lot size restrictions. Second, the formula prediction of the density exposure based on Equation (5). Third, the recentered instrument.

The second row or panel plots the density exposure predicted by the formula on x. The solid lines are the exposures predicted as a function of the realized minimum lot sizes and x. In either scenario, the density exposure determined by the lot size restrictions increases moving left to right. However, the gradient is much flatter under Shock Realization #2. We also plot the expected instrument on x for houses on the left side of the border, which is shown by the orange dotted lines. At a given x, the expected instrument μ is the mean value of $z(x)$ averaged over the R town's set of counterfactual minimum lot sizes. By definition, the expected instrument μ is the same for the left and right columns, and lies between the two scenarios' gradients. Because the right town tends to set smaller minimums compared to the left town, we expect the formula prediction of the density exposure to increase moving left to right.

The third row plots the recentered density exposure on x calculated by subtracting the orange dotted line from the solid line in the middle panel, i.e. $\tilde{z}_j = z_j - \mu_j$. Under Shock Realization #1, houses in the left municipality would have a higher than expected density exposure instrument value. As a result, the gradient moving left to right for the recentered instrument is similar to the gradient for the unadjusted instrument in the middle row. In contrast, under Shock Realization #2, the recentered instrument becomes more negative closer to the border. In this scenario, the houses close to the border actually have lower density exposure than we would expect given their close proximity to a R town. For a house on the left side of the border, the recentered instrument would only use variation in density exposure that comes from its proximity to the right town's minimum lot size being higher or lower than we would expect for the right town.

In this stylized example, we assumed knowledge of a R town's distribution of counterfactual minimum lot sizes. More generally, let $G(g|w)$ denote the distribution of minimums conditional on w. The following assumption formalizes this knowledge of the assignment process.

Assumption 2 *(Known assignment process): $G(g|w)$ is known in the support of w.*

Assumption 2 may be satisfied if one is willing to view minimum lot sizes set by an opposing town as *i.i.d.* within some set of minimums in the opposing town. In Section 4 after introducing the data sources, we provide an empirical example with variation analogous to Figure 3 and illustrate how we use within-jurisdiction variation in minimum lot sizes to estimate counterfactuals from the data.

D *Heterogeneous treatment effects*

In this context, a textbook linear IV model like that in (8) that assumes homogeneous treatment effects, also assumes homogeneous preferences over density exposure. In particular, the hedonic model relates a homeowner i's MWTP to the price derivative:

$$\text{MWTP}_i = \frac{\partial p(D)}{\partial D} = \beta \ \ \forall \ i.$$

Here, the marginal effect of density exposure on price, β, is a constant.[7] In this case, using a conventional TSLS/IV estimator with a valid instrument identifies the average MWTP. However, this also leads to the unnatural conclusion that all homeowners have the same marginal willingness to pay for density exposure.

In the presence of heterogeneous treatment effects, we must consider averages of individual effects. Consider taking the average over homeowners on both sides of the first order condition:

$$\mathrm{E}[\mathrm{MWTP}] = \mathrm{E}[\frac{\partial p(D)}{\partial D}] = \mathrm{E}[\beta]. \tag{10}$$

The left hand side is the average MWTP and the right hand side is the average price derivative, which we term the average treatment effect (ATE) for familiarity, though more precisely it is the average partial effect of a structural function. Hence, if we had an estimator that identified the ATE, we could identify the average MWTP for density exposure. However, a conventional TSLS/IV estimand generally will not identify the ATE, although it may identify a convex weighted average of marginal effects under certain conditions.[8]

We pursue two approaches to extrapolate from an IV estimand to the sample ATE under additional but distinct assumptions. First, we consider identification under the assumption that price effects are heterogeneous across neighborhoods but homogeneous within neighborhoods. With this assumption, neighborhood-specific ATEs are identified and can then be reaggregated to the sample ATE (Angrist and Fernández-Val, 2013). Second, we consider identification assuming a first stage selection equation where the unobserved heterogeneity is fully characterized by a one-dimensional latent index. Given this assumption, the sample ATE can be estimated using a control function approach. We then follow the estimation procedure outlined in Masten and Torgovitsky (2016).

LATE-REWEIGHT: Here, we take the outcome equation in (8) and allow a limited form of heterogeneity as random coefficients β. In particular, we group houses into neighborhoods by geographic location and allow price effect heterogeneity only by neighborhood.

Neighborhoods partition a sample. Specifically, denote a house j's neighborhood as $n(j) \in \{1,2,\ldots,N\}$, with neighborhoods forming an exclusive and exhaustive set of geographically adjacent groups. To emphasize price effect heterogeneity, denote $\beta_{n(j)}$ as

[7] Appendix C derives the relation between MWTP and the price derivative in the standard hedonic model.

[8] Appendix C.1 of Borusyak and Hull (2023) shows under IV assumptions and a first stage monotonicity assumption that the recentered IV identifies a convex weighted average of marginal effects without negative weighting.

the price effect of density exposure specific to neighborhood $n(j)$. This generates a set of TSLS models with second stage price equations:

$$p_j = \beta_{n(j)} D_j + \tau_{n(j)} l_j + A_{bm(j)} + W_j' \delta_{n(j)} + \epsilon_j. \tag{11}$$

Assuming a constant β_n within neighborhood, we can construct the sample ATE from the within-neighborhood ATES. Each β_n is a conditional ATE and we can reaggregate to the unconditional ATE by averaging over the marginal distribution of houses across neighborhoods:

$$\mathrm{E}[\beta] = \mathrm{E}[\mathrm{E}[\beta|n]] = \sum_{n=1}^{N} \beta_{\text{TSLS},n} P(n(j) = n).$$

This result is an application of the LATE-REWEIGHT theorem in Angrist and Fernández-Val (2013), in which the observable covariate used to characterize treatment effect heterogeneity is a discretization of geographic space.

There is an economic rationale for a within-neighborhood homogeneity assumption. In the hedonic model, homeowners sort into neighborhoods based on their preferences and one price function is determined within a market. Presumably, a housing market is at least the size of a neighborhood. Still, it is possible to imagine violations and practical issues with this assumption. For example, if larger houses suffer larger price declines from an increase in density exposure, this nonseparability could create heterogeneity at a level more granular than the neighborhood. As we discuss more fully below, we must choose a geographic partition coarse enough that numerical estimates are stable. Consequently, we cannot characterize heterogeneity for an arbitrarily fine geographic partition. Hence, we next turn to a second approach to extrapolate to the sample ATE.

IVCRC: We now allow for unit-level heterogeneity in the random coefficients. For emphasis, denote β_j as the slope coefficient in a linear correlated random coefficient model and rewrite the econometric model as

$$p_j = \beta_j D_j + \tau_j l_j + A_{bm(j)} + W_j' \delta_j + \epsilon_j. \tag{12}$$

For a linear correlated random coefficient model, Masten and Torgovitsky (2016) show a consistent estimator for the ATE can be constructed under additional assumptions. The most substantive assumption is in the form of a first stage selection equation. Let C_j be a vector collecting all control variables. In particular, assume selection into treatment takes the following form.

Assumption 3 (First stage equation) *$D_j = \nu(\tilde{z}_j, C_j, v_j)$ where ν is an unknown function that is strictly decreasing in v_j, which is a scalar unobservable.*

This assumption is sometimes referred to as a rank invariance assumption because it requires that the ordinal ranking of any two housing units would be the same in terms of the density exposure if both houses received the same instrument value, for any instrument value. In addition, although the following assumption is relatively standard for IV, it is worth noting that we also need to strengthen the instrument exogeneity assumption.

Assumption 4 *(Full Exogeneity):* $\tilde{z} \perp\!\!\!\perp (\epsilon, v)|C$

Masten and Torgovitsky (2016) show that, under these additional assumptions, $\mathrm{E}[\beta]$ is identified. They also propose an estimator and provide an estimation procedure (Benson et al., 2022). Following those authors, we call this estimator the IVCRC estimator.

E *Estimation*

We briefly outline the estimation procedures for the LATE-REWEIGHT and IVCRC approaches. We face a challenge with the LATE-REWEIGHT approach in how to choose a partition into neighborhoods. Finer partitions better account for heterogeneity across space, but arbitrarily fine partitions will have low first stage power in practice. We use a simple algorithm to form a partition where no individual neighborhood has low power. To start, we form an initial partition by setting a house's neighborhood to its side of a border area, $n(j) := bm(j)$. For each neighborhood, we then run a separate first stage regression using only the observations within that neighborhood and retrieve the F statistic. Next, we find the neighborhood with the lowest F and merge it with the geographically closest other neighborhood in the current partition. We then rerun the first stage regression on the new neighborhood formed from merging together the two original neighborhoods. We repeat this process of coarsening the partition until every neighborhood has $F > 10$. Given this final partition, we can then calculate standard TSLS estimates separately for each neighborhood.

The IVCRC estimation routine has two stages. In the first stage, we use a series of quantile regressions to estimate conditional ranks $\hat{r}_j$ for each observation. In practice, the large number of observations and fixed effects in our regression specification poses a computational problem for estimating many quantile regressions. To deal with this issue, we adopt a correlated random effects specification for the first stage equation using a Mundlak device (Wooldridge, 2010, Arkhangelsky and Imbens, 2024, Kashner and Ross, 2025) and use recent software packages for quantile regression with large-scale data (He et al., 2023). Under Assumptions 3 and 4, the conditional ranks are a valid control variable. In the second stage, OLS estimation of (12) conditional on a rank r' therefore identifies $\mathrm{E}[\beta|\text{rank} = r']$ and then averaging over ranks identifies $\mathrm{E}[\beta]$, the sample ATE. The conditioning on r' is incorporated by applying kernel weights, where the weights are based on the distance of $\hat{r}_j$ from r'.

More details on the estimation routines for both approaches are in Appendix F.

4 Data

We construct empirical analogs to the variation in Figures 1 and 3 by leveraging data from several sources. The three key data sets are tax assessment files on single family houses from Cotality (formerly CoreLogic), census block housing unit counts, and shapefiles of administrative borders. For one border, we illustrate how we combine these three data sources to form the outcome, treatment, and instrument in our IV design. We then recenter the instrument in this example using other nearby minimum lot sizes within the same town as counterfactuals. The last part of this section provides a description of the main analysis sample, which includes thousands of borders across the United States.

A *Data sources*

Primary data sources: Our analysis has three main data sources.

1. *Cotality Single Family Houses.* The primary dataset is parcel-level data on single family homes from the Cotality tax assessment file. For the near universe of single family houses, the Cotality data has information on a house's most recent sale price, its precise location by geographic coordinates, and other observable individual characteristics including lot size, living area, and the age of the house. This data serves several purposes in our analysis. We use Cotality to characterize the outcome and treatment, and we also infer minimum lot restrictions following Song (2025). More detail is provided below.

2. *Census Block Housing Unit Counts.* To count the number of housing units close to an individual house's location in Cotality, we retrieve census block housing unit counts from the 2010 U.S. Census.

3. *Census TIGER/Line Shapefiles.* To characterize administrative borders between zoning authorities, we retrieve shapefiles for Census Places and Census County Subdivisions from the 2019 U.S. Census TIGER/Line Shapefiles.

Supplementary data sources: We use several supplementary data sources to assess sensitivity to inclusion of geographic controls, perform falsification tests, and explore mechanisms.

4. *OpenStreetMap Points of Interest.* We calculate each house's distance to a highway, body of water, and school using point of interest data from OpenStreetMap.

5. *CBSA centers.* We measure the CBSA center as the coordinate location reported by searching the central cities of CBSAs on Google Maps.

6. *NASA Digital Elevation Raster Data.* To measure local topography, we download digital elevation maps (DEMs) from NASA. Specifically, we use maps derived from the Shuttle Radar Topography Mission (SRTM) Global 1-arc second data collection. These data report the elevation for most of the Earth's surface at a spatial resolution of about 30 meters. We also construct slope maps by translating these elevation maps using standard GIS tools.

7. *Census Block Demographics.* We collect granular geographic information on neighborhood race counts available at the census block level in the 2010 U.S. Census.

8. *Gun Violence Archive Incident Data.* We measure exposure to violent crime by counting nearby incidents of violent crime over roughly a decade. We use all incidents reported in the Gun Violence Archive from January 1st, 2014 to July 5th, 2024.

Minimum lot size measures: Unfortunately, direct data on minimum lot sizes at a fine spatial scale is typically unavailable. However, a recent literature has shown considerable accuracy in inferring minimum lot sizes from the observed distribution of lot sizes in Cotality (Song, 2025, Cui, 2024, Macek, 2024). The general idea for any of these methods is that the lot size distribution will have bunching at the minimum, and estimating this bunching location therefore reveals the minimum.

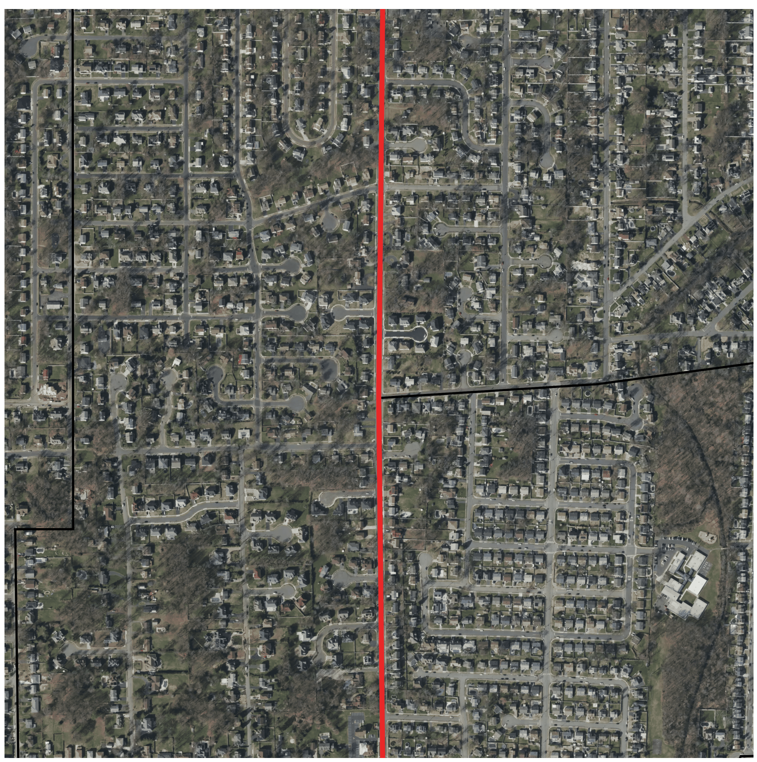
We apply the minimum lot size detection algorithm in Song (2025)'s replication package.[9] Because our research design requires minimum lot sizes attached to a location, we run the algorithm on census block groups. In contrast, Song (2025) groups houses according to the zoning district code reported by Cotality, and then block group if the code is missing. Except for running the algorithm on block groups, we otherwise perform the same exercise as Song (2025). While block groups do not follow zoning districts exactly, when we compare the inferred minimums with actual minimums in Song (2025)'s validation set, we attain correlations and error rates that are only moderately worse than those attained by Song (2025)'s measure. For example, matching the block groups in our sample to the Boston zoning districts in Song (2025)'s validation set, our block group-level measures attain a correlation of 0.8 and a median error rate of 25%, while Song (2025)'s measure has a correlation of 0.9 and a median error rate of 15% among the same zoning districts. Appendix E provides more details on Song (2025)'s minimum lot size detection algorithm and the validation exercise.

[9] `https://www.openicpsr.org/openicpsr/project/231447/version/V1/view`.

We combine the first three data sources listed above to construct the outcome (p_j), treatment (D_j), and formula instrument (z_j) following the design derived from the theoretical framework. Cotality directly reports the outcome–a house's sale price, p_j. For the treatment, by taking a house's location reported in Cotality and counting the number of housing units in nearby census blocks, we construct a density exposure measure, D_j. For the instrument, first we infer minimum lot size restrictions using the distribution of single family lot sizes in Cotality following Song (2025), yielding $\underline{l}_{sm(j)}$ and $g_{-sm(j)}$. By then combining these minimum lot size measures with a house's distance to the administrative border in the Census Shapefiles, we can characterize each argument that determines the formula instrument, z_j. To illustrate the construction of these three variables, we next walk through an example for one border in our estimation sample.

B *A border example*

Figure 4: Edison-Woodbridge Border, NJ



Note: This figure shows part of the border between the towns of Edison and Woodbridge in New Jersey. The red line is the border between the towns. The left side is Edison and the right side is Woodbridge. The black horizontal line in the middle of the image delineates census block groups, forming two "segments" along the same municipal border.

The red line in Figure 4 is the municipal border between the towns of Edison and Woodbridge in the suburbs of New York City. The left side is Edison and the right side is Woodbridge. Edison is modestly higher income than Woodbridge and also sets relatively larger minimum lot sizes across most of its jurisdiction. At the border, lot and physical structure sizes are immediately smaller on the Woodbridge side, particularly for houses south of West Clark Place, the road indicated by the black horizontal line toward the center of the image. The decrease in minimum lot size at the border generates variation in density exposure moving left to right within a small geographic area.

We can consider a regression that includes geographic fixed effects down to a side of a border area, thereby isolating only comparisons between housing sales that are in a small geographic area within the same municipality. We could also instrument by distance to the border, thereby using only variation in the house price gradient moving left to right. Still, it is at least possible to imagine threats to identification. For example, we can observe a school and green spaces lying in the interior of Woodbridge. These amenities could be correlated with distance to the border and could affect house prices. While one could of course control for specific amenities, it is impossible to control perfectly for all relevant geographic factors without removing all variation in the density exposure.

However, there is also variation in the shift in lot sizes along the border. Within the Woodbridge side, most houses above West Clark Place are zoned at a higher minimum lot size than those below West Clark Place. The black line shows block group divisions from the census shapefiles, forming two "segments" along the same municipal border where the block groups are divided at West Clark Place. As described more fully below, we will exploit variation in the magnitude of the minimum lot size discontinuities at the border to construct a recentered instrument.

We construct a house-level density exposure measure by counting the number of housing units in nearby census blocks. Figure 5 illustrates for one parcel in the Edison-Woodbridge border area. The blue dot is the parcel's location and the blue circle shows the area within 500 meters of that parcel. The dotted black lines delineate U.S. census blocks. The black numbers are the housing unit counts within a given census block plotted at its interior point as reported by the U.S. Census. To construct the density exposure measure, we total the counts of all census blocks with internal points within the blue circle. We repeat this process separately for each parcel in our sample yielding a measure of density exposure for each individual house.

Figure 5: A measure of a border parcel's density exposure

Note: This figure shows the construction of a density exposure measure for one parcel close to the Edison-Woodbridge border. The blue dot is the parcel's location and the blue circle shows the area within 500 meters of that parcel. The dotted black lines are U.S. census blocks. The black numbers are the housing unit counts in a given census block plotted at its interior point as reported by the U.S. Census. To construct the density exposure measure, we total the counts of all census blocks with internal points within the blue circle.

Figure 6 then depicts parcel-level density exposure measures and the associated instruments on the left side of the border in Edison. For all four panels, yellow indicates low density values and purple indicates high density values. In each panel, the right side of the border is colored by the lot size restrictions in Woodbridge measured following Song (2025). The block group above West Clark Place has a minimum lot size measure of 10,000 ft², while the minimum lot size below West Clark Place is 5,900 ft².

Panel (a) shows all the parcels' density exposures calculated as in Figure 5. As anticipated, density exposure tends to increase moving left to right, but there is also variation in exposure due to other factors. For example, parcels in the bottom left tend to have lower density exposure because there is a church within their 500m buffers.

Panel (b) plots the density exposure predicted by the formula in Equation (5). For each border segment, the formula instrument uses only the variation predicted by the measured lot size restrictions and radial distance to the municipal border. In contrast to the actual density exposure in panel (a), there is no variation due to other geographic factors.

Panel (c) plots each parcel's expected density exposure. This is the average density exposure over the counterfactual restrictions that Woodbridge might have set in the block group on the opposite side of a parcel's border segment. To estimate counterfactual restrictions for the two block groups on the Woodbridge side of the border, we use the 8 closest minimum lot size restrictions within Woodbridge.[10] Figure 7 shows an example of the 8 closest minimum lot sizes for the block group south of West Clark Place. The increasing gradient moving left to right suggests that for either border segment between the two municipalities we should expect density to increase going from Edison to Woodbridge because Edison tends to set higher minimum lot sizes than Woodbridge.

Panel (d) plots the recentered density exposure, which is the difference between panel (b) and panel (c). The effect of recentering can be seen by contrasting the predicted density exposure in panel (b) with the recentered density exposure in panel (d). This contrast is especially pronounced for parcels above West Clark Place close to the municipal border. The recentered instrument no longer places these parcels in the medium to high density exposure group because their formula density exposure is actually lower than we would expect given their close proximity to Woodbridge. On the other hand, parcels below West Clark Place are close to a lot size restriction on the opposite side of the border segment that is low relative to the typical restriction set in Woodbridge. This implies a similar gradient in the recentered instrument as the predicted density exposure that increases moving left to right.

This border example forms an empirical analog similar to the two scenarios in Figure 3. Although we can not observe multiple counterfactuals for the same border segment, the relationship between the recentered instrument and distance to the border for parcels above West Clark Place is similar to Shock Realization #2 in Figure 3, while for parcels below West Clark Place the relationship is similar to Shock Realization #1.

[10]Using the $k = 8$ nearest minimum lot sizes is an arbitrary bandwidth selection. In Appendix Table B2, we assess sensitivity to every choice of k between 2 and 15.

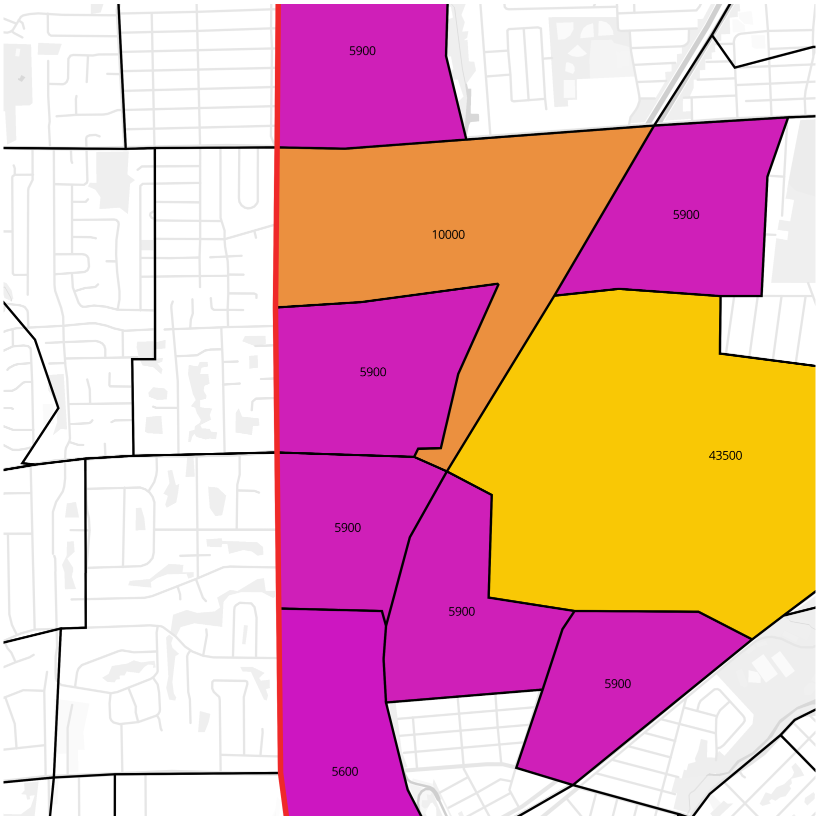
Figure 6: Recentering example

(a) Density Exposure (b) Predicted Density Exposure

(c) Expected Density Exposure (d) Recentered Density Exposure

Note: This figure shows the density exposure and instruments, with darker shading indicating higher density values. The Edison side of the border plots house-level variation in the density exposure and instrument values. The Woodbridge side reports the minimum lot size in each block group measured following Song (2025). Panel (a) shows each parcel's density exposure measure calculated as in Figure 5. Panel (b) is the density exposure instrument predicted by the formula in Equation (5). Panel (c) is the expected density exposure using the 8 closest lot size restrictions in Woodbridge to each block group as counterfactuals. Panel (d) is the recentered density exposure: the difference between panel (b) and panel (c).

Figure 7: Minimum lot size cluster example



Note: This figure illustrates the 8 closest minimum lot size measures to the block group south of West Clark Place in Figures 4 to 6. The black numbers report the minimum lot size restriction measure in each block group in ft². We use the 8 nearest minimum lot size measures within Woodbridge as counterfactuals for minimum lot sizes that could have occurred in the block group adjacent to the border but did not. We then calculate expected instrument values for border parcels in the town on the opposing side of the border, Edison.

C *Estimation sample*

To construct our sample, we follow Turner et al. (2014) and select borders that do not follow natural features. We start by finding every municipal border between zoning authorities in the conterminous United States using the shapefiles for Census Places and County Subdivisions and find any single family house in Cotality within 500 meters of that border. We then drop borders that follow a highway, river, or a ridge line and implement the main sample restriction in Turner et al. (2014) by restricting to parcels close to municipal borders that are straight lines, which are more likely to follow old federal survey lines. To identify parcels associated with a straight line border, we apply the same algorithm as in Turner et al. (2014).[11] This yields the final analysis sample that still contains over 1.4 million single family houses around 2,432 borders and 2,783 municipalities, which reflects the fact that some municipalities have multiple borders with different communities.

The resulting estimation sample is representative for a typical U.S. suburb, but may not be representative for the typical U.S. household. The average lot size is around one-quarter acre, the average density exposure is 625 housing units within 500 meters, and the average distance from the CBSA center is 24.5 kilometers, implying our sample is predominantly suburban. The average sale price is comparable to the national average for a single family house at $408,376 in 2023 dollars. More detailed summary statistics on the estimation sample and boundary selection are provided in Appendix Table B1.

While suburban, predominantly single family neighborhoods are the primary neighborhood type of interest concerning strong density restrictions, our results may not be generalizable to the typical American who lives closer to the dense core of a metropolitan area, especially given the attention we pay to preference heterogeneity. We discuss external validity and its relation to preference heterogeneity at length in Section 6.

5 Results

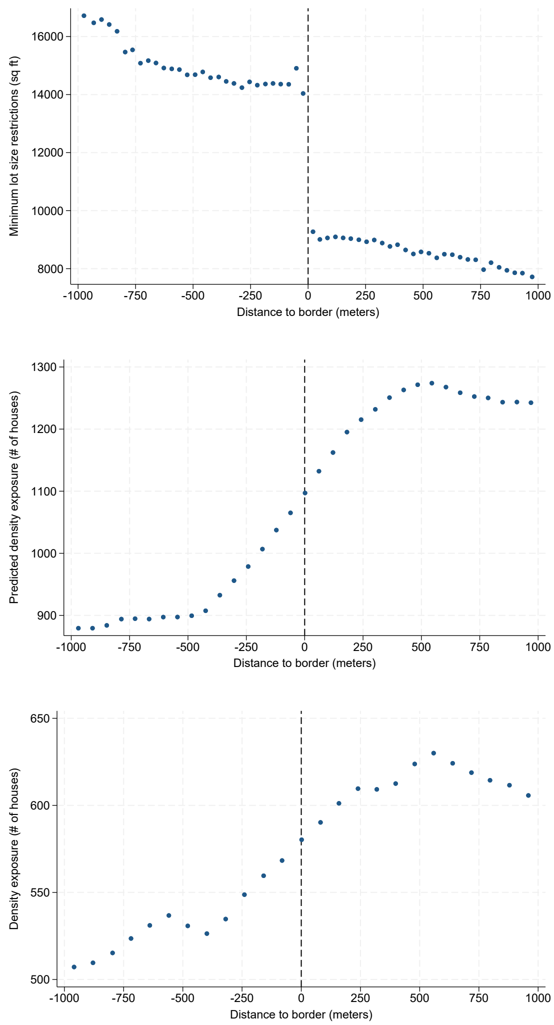
A *Graphical evidence*

Before proceeding to the main results, we provide graphical evidence that lot size and density exposure near borders with minimum lot size discontinuities follow patterns consistent with the border model in Section 2. We then look at how house prices vary around these borders, working toward the hedonic IV estimates presented in the next subsection.

The top graph in Figure 8 plots the minimum lot sizes measured following Song (2025) on the signed distance to the border x where within each border segment we assign a house in the larger minimum lot size area to have the negative of its radial distance to the border as in Figure 1. On average, minimums decrease discontinuously by 4,000-5,000 ft^2 at the border.

[11]For each house, we draw the shortest line to the municipal border. We then form two orthogonal vectors each with length 200 meters at the intersection of the first line and the border. If the endpoints of the two orthogonal vectors are within 15 meters of the border, then we say a house is near a straight border.

Figure 8: Min. lot sizes and density exposure on distance to the border

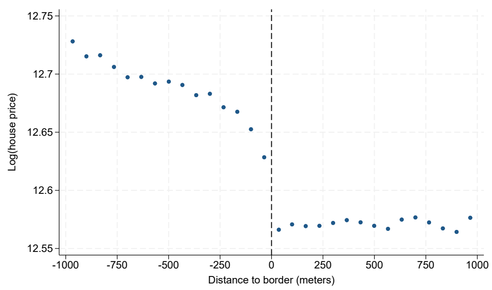


Note: This figure presents binscatter plots of the minimum lot sizes, formula density exposure, and actual density exposure on the signed distance to the municipal boundary pooling together border areas. The top row plots the minimum lot size restrictions measured following Song (2025). The middle row plots the density exposure predicted by Equation (5). The third row is the actual density exposure. For each block group segment, parcels on the more regulated side of the border are assigned negative distances. The plotted relationships adjust for border area fixed effects and select the optimal number of bins following Cattaneo et al. (2024) clustering on border area.

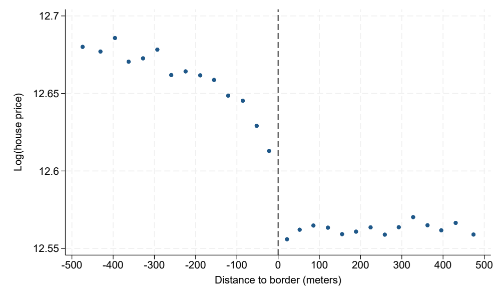
The middle and bottom graphs of Figure 8 plot the density exposure predicted by Equation (5) and the actual density exposure on distance to the border. On average, the density exposure follows a functional form similar to the prediction of the formula in Equation (5). Going left to right, the density exposure is kinked at -500 meters where we set the spillover bandwidth R; it then increases continuously through the municipal border; lastly, it flattens out at +500 meters. Taken together, the plots in Figure 8 show that minimum lot sizes induce variation in density exposure that we can predict using the model in Section 2.

Figure 9: Log(house price) on distance to the border

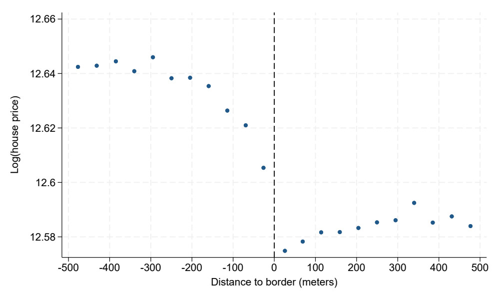
(a) <1,000 meters



(b) <500 meters



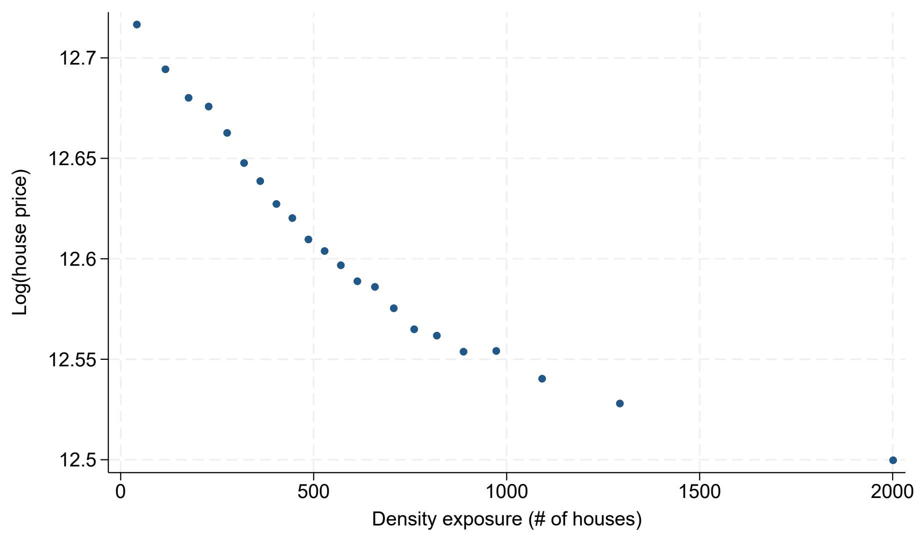
(c) <500 meters, quality-adjusted



Note: This figure presents bin scatter plots of log(house price) on distance to the municipal boundary pooling together the border area. For each block group segment, parcels on the more regulated side of the border are assigned negative distances. The plotted relationships adjust for border area fixed effects and select the optimal number of bins following Cattaneo et al. (2024) clustering on border area. Panel (c) additionally residualizes individual housing traits including a house's lot size, living area, and age.

Our outcome of interest is house price. Figure 9 shows various binscatter plots of log(house price) on distance to the border. Panel (a) of Figure 9 shows the average price gradient within 1 kilometer of a border. Going left to right, house prices decline continously on the more regulated side of the border. They then drop discontinuously at the boundary along with the minimum lot size discontinuity. Prices are essentially flat on the right entering the interior of the less regulated side. We may expect houses farther away from each other to be more unlike in their unobservable characteristics. Thus, we may prefer estimates based on a smaller bandwidth around the borders. Panel (b) of Figure 9 is similar to (a) but now zooms into the smaller geographic area within 500 meters of a border. Within this area, the slope of the price gradient is similar to (a). However, we still may be concerned that other individual housing characteristics such as lot size are correlated with the distance to the border. Panel (c) of Figure 9 again plots log(house price) on distance to the border but now residualizes a house's lot size, living area, and age. Looking within each side of the border, while the gradient on the right side of the border might be modestly positive, the gradient on the more regulated side of the border is negative and steeper. Taken together, Figures 8 and 9 suggest an IV regression using the (signed) distance to the border as an instrument will measure a negative price effect from density exposure.

Figure 10: Log(house price) on density exposure



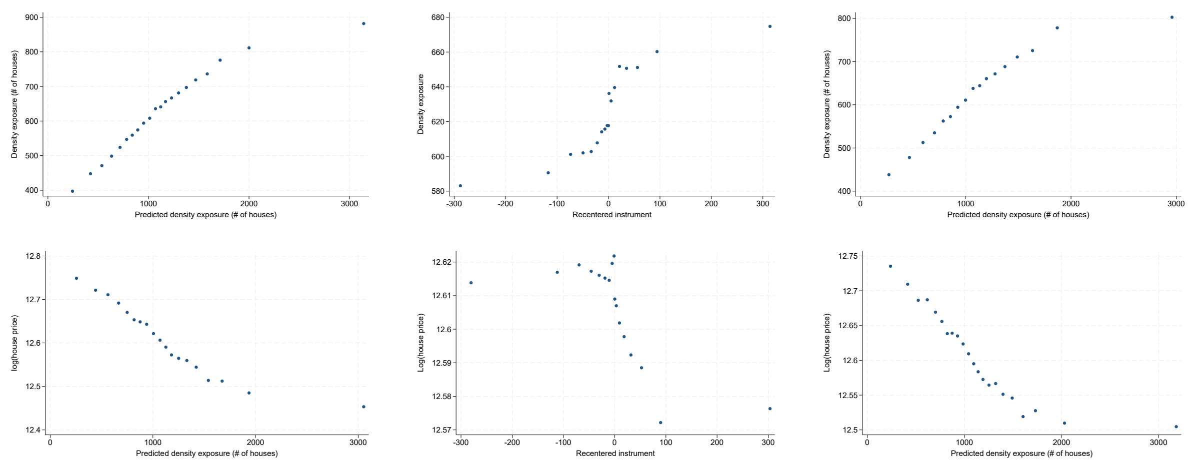
Note: This figure presents a bin scatter plot of log(house price) on the density exposure. The plotted relationship adjusts for border side specific fixed effects following Cattaneo et al. (2024). Note this implies comparisons are within the same town within a border area. We selected the optimal number of bins following Cattaneo et al. (2024) and cluster by border area. This figure additionally residualizes a house's lot size, living area, and age, and the minimum lot size restriction in a house's own zoning district.

The relationship of interest is the effect of density exposure on price. We can alternatively view the problem as a hedonic directly regressing price on density exposure. Figure 10 plots log(house price) on the density exposure. This figure now adjusts for border *side* specific fixed effects, so comparisons are within the same town within a small geographic area. As in the previous figure, we also residualize a house's lot size, living area, and age. House prices are lower for higher density exposure levels. Similar to the binscatter, the OLS estimate for the analogous linear regression is -1%. However, as discussed above, omitted variable bias in a regression of price on density exposure is still plausible, even including observable covariates.

Figure 11 shows binscatter plots analogous to the first stage and reduced form of an IV regression, treating the density exposure predicted by Equation (5) as an instrument. This projects house prices onto only the variation that is a function of the minimum lot sizes and distance to the border. The left column shows plots for the predicted density exposure without recentering. The middle column shows plots for the recentered instrument. The right column again plots the unadjusted density exposure prediction but now treats the expected instrument as a control and residualizes the expected instrument following Cattaneo et al. (2024). The top row of graphs is the first stage for each instrument. The bottom row is the reduced form.

Figure 11: Formula IVs using the predicted and recentered density exposures

(a)	(b)	(c)
z	$\tilde{z}$	z, control for μ



Note: This figure shows binscatter plots analogous to the first stage and reduced form of an IV regression using the formula instruments. The left column uses the density exposure predicted by the formula in (5). The middle column uses the recentered version of the prediction. The third column controls for μ instead of directly recentering the instrument. The top row of graphs plots the actual density exposure on the instruments and the bottom row plots log(house price) on the instruments. The plotted relationship adjusts for border side specific fixed effects following Cattaneo et al. (2024). We selected the optimal number of bins following Cattaneo et al. (2024) and clustered by border area. This figure additionally residualizes a house's lot size, living area, and age, and the minimum lot size in a house's own zoning district.

Starting with the top left plot, the formula density exposure is a strong predictor of the actual density exposure with a positive first stage. In the bottom left plot, house price decreases with the formula density exposure, implying a negative reduced form. The middle column shows analogous plots for the recentered version of the instrument. While significantly noisier, these graphs suggest similar relationships. The recentered instrument is positively associated with the actual density exposure and negatively associated with house price. As noted in Borusyak and Hull (2023), controlling for the expected instrument is another related solution to omitted variable bias, which implicitly recenters the instrument, and may be more efficient in large samples. Doing so yields the relationship plotted in the right-most column. The instrument is positively associated with density exposure and negatively associated with house price. Taken together, we expect an IV regression to indicate density exposure lowers house price.

B *Price effect estimates*

Table 1 presents IV regressions of log(house price) on density exposure, where each column uses a different formula instrument construction defined in Section 3. The first column uses the signed distance to the border, i.e. the x-axis in Figures 8 and 9. The second column uses a formula analogous to the specification of the external effect in Turner et al. (2014). The third column uses the density exposure predicted by Equation (5). The fourth column uses the recentered instrument, which subtracts the expected density exposure from the prediction of Equation (5) using the method illustrated in Section 4. The fifth column, instead of recentering, controls for the expected density exposure.

The first three rows (panels a, b and c) report estimates of the first stage, the reduced form, and of the effect of adding 100 housing units within 500 meters on log(house price). Estimates range from approximately -3 to -7%. Recall that the corresponding OLS estimate was -1% as suggested by Figure 10, so these estimates are significantly larger in magnitude—including the estimate using the recentered instrument at -3.5%.

Figure 12 provides more context on the range of variation underlying these estimates. We partition our sample into 2,126 geographically-adjacent neighborhoods and estimate a separate TSLS coefficient within each neighborhood. The figure plots the distribution of these estimates using the recentered instrument. The mean (-1.7%) and median (-1.4%) are smaller in magnitude than the conventional TSLS estimate. The mean corresponds to the LATE-REWEIGHT estimand described in Section 3 and reported in column (4) of Table 1. Figure 12 shows summary price effect estimates conceal a large degree of heterogeneity: the interquartile range is -7% to +4%.[12]

Panels d and e of Table 1 instead target the sample ATE using the LATE-REWEIGHT and IVCRC approaches, respectively. These approaches require the additional assumptions described in Section 3. The range of estimates is narrower than those for the conventional TSLS, at around -1 to -2%, which is closer to the OLS estimate of -1%. Furthermore, given the confidence intervals, we can not reject that these estimates are the same.

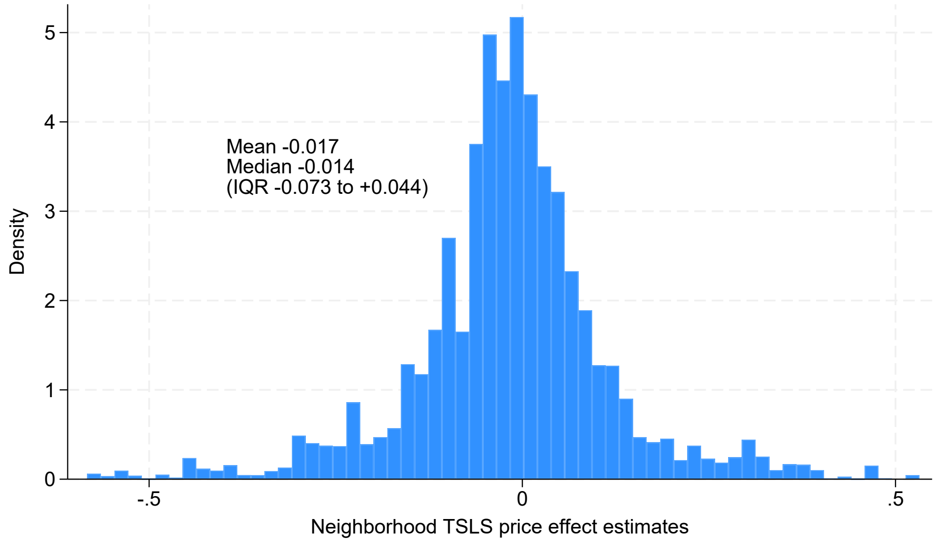
[12] Although it should be caveated that individual neighborhoods estimates are imprecise, accounting for the excess variance in $\hat{\beta}_n$ due to sampling error appears to only modestly reduces dispersion in the distribution. Appendix Figure A1 applies the empirical Bayes deconvolution estimator of Efron (2016) to the estimated price effects. The interquartile range for the estimated mixing distribution after accounting for excess noise is comparable to the range reported in Figure 12.

Table 1: Effect of Density Exposure on Log(House Price), IV estimates

	(1) Distance	(2) THV	(3) Prediction	(4) Recentered	(5) Control μ
Panel a. First stage					
Instrument	0.0527***	-0.0015***	0.1225***	0.1038***	0.1179***
	(0.00784)	(0.00026)	(0.01815)	(0.02690)	(0.02440)
Panel b. Reduced form					
Instrument	-0.00357***	0.00005***	-0.00830***	-0.00362**	-0.00464***
	(0.000709)	(0.000015)	(0.001209)	(0.001688)	(0.001500)
Panel c. TSLS					
100s of houses	-0.068***	-0.036***	-0.068***	-0.035***	-0.039***
	(0.0147)	(0.0078)	(0.0093)	(0.0125)	(0.0110)
Panel d. LATE-RWGT					
100s of houses	-0.010***	-0.023***	-0.020***	-0.017***	-0.017***
	(0.0037)	(0.0043)	(0.0043)	(0.0042)	(0.0056)
Panel e. IVCRC					
100s of houses	-0.013***	-0.013***	-0.019***	-0.013***	-0.011***
	(0.0013)	(0.0014)	(0.0020)	(0.0013)	(0.0015)
Border$\times$Muni FE	Yes	Yes	Yes	Yes	Yes
House controls	Yes	Yes	Yes	Yes	Yes
Own side MLR	Yes	Yes	Yes	Yes	Yes
Effective F	45.1	33.2	45.6	14.9	23.3
Number of obs	1,424,989	1,424,989	1,424,989	1,424,989	1,424,989
Number of Borders	2,432	2,432	2,432	2,432	2,432

Note: Each column reports estimates for an IV regression of log(house price) on density exposure using a different instrument construction as defined in Section 3. The first column uses the signed distance to the border. The second column uses a formula analogous to the specification of the external effect in Turner et al. (2014). The third column uses the density exposure predicted by Formula (5). The fourth column uses the recentered instrument, which subtracts the expected density exposure from the predicted density exposure using the method described in Sections 3 and 4. The fifth column, instead of recentering, controls for the expected density exposure. Panel a. reports first stage estimates. Panel b. reports reduced form regressions. Panel c. reports conventional TSLS *estimates for the effect of density exposure in 100s of houses within 500 meters on log(house price). Panels d. and e. instead use the* LATE-REWEIGHT *and* IVCRC *approaches described in Section 3. All models include geographic fixed effects specific to a side of a border area, house-level controls for lot size, living area, and age, and a control for the minimum lot size restriction in a house's own block group. Panels a. to c. report standard errors clustered by border area, Panel d. reports the standard error of the neighborhood* TSLS *estimates' mean, and Panel e. reports bootstrap standard errors based on 500 draws resampling border areas. Units are divided by 100 for legibility.*

Figure 12: Distribution of neighborhood TSLS estimates



Note: This figure plots the distribution of 2,126 neighborhood-specific TSLS *estimates using the recentered instrument weighted by the number of observations within each neighborhood. Measures of central tendency are reported in the text box.*

C *Falsification tests and robustness checks*

One potential concern is that exposure based on distance to the border is correlated with other geographic factors. Table 2 reports estimates including additional observable controls for local geography. The first three columns use the signed distance to the border as the instrument. The next three use the recentered instrument. Within a set of three, each column successively layers on additional geographic controls. The first column adds a control for distance to the city center. The next column controls for the local topography. The third column controls for a house's distance to a body of water, park, school, and highway. Panel a. reports conventional TSLS/IV estimates, while Panels b. and c. report estimates using the LATE-REWEIGHT and IVCRC approaches.

The estimates generally are not sensitive to inclusion of additional geographic controls with similar patterns as seen above in Table 1. However, the conventional TSLS/IV estimates are again sensitive to whether the signed distance to the border or the recentered formula prediction is used as an instrument (top panel), while the LATE-REWEIGHT and IVCRC are comparable in that they range from -1 to -2%. While Table 2 is reassuring, it is at least plausible that other unobserved geographic factors could violate instrument orthogonality, and so we may prefer to rely on estimates using a recentered instrument.

Table 2: Effect of Density Exposure on House Price, Additional Geographic Controls

	Distance			Recentered		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel a. TSLS						
100s of houses	-0.068***	-0.067***	-0.067***	-0.035***	-0.035***	-0.035***
	(0.0145)	(0.0147)	(0.0147)	(0.0123)	(0.0123)	(0.0121)
Panel b. LATE-REWEIGHT						
100s of houses	-0.020***	-0.020***	-0.010***	-0.023***	-0.021***	-0.012***
	(0.0041)	(0.0041)	(0.0039)	(0.0050)	(0.0073)	(0.0041)
Panel c. IVCRC						
100s of houses	-0.013***	-0.013***	-0.012***	-0.013***	-0.013***	-0.012***
	(0.0013)	(0.0013)	(0.0014)	(0.0013)	(0.0014)	(0.0015)
Border×Muni FE	Yes	Yes	Yes	Yes	Yes	Yes
House controls	Yes	Yes	Yes	Yes	Yes	Yes
Own side MLR	Yes	Yes	Yes	Yes	Yes	Yes
Dist. CBD	Yes	Yes	Yes	Yes	Yes	Yes
Elevation & Slope	No	Yes	Yes	No	Yes	Yes
Dist. to amenities	No	No	Yes	No	No	Yes
Effective F	46.6	45.3	44.3	15.3	15.4	14.5
Number of obs	1,424,989	1,424,989	1,424,989	1,424,989	1,424,989	1,424,989
Number of Borders	2,432	2,432	2,432	2,432	2,432	2,432

Note: Each column reports estimates for an IV *regression of log(house price) on density exposure. The first three columns use the signed distance to the border as the instrument. The next three columns use the recentered instrument. Within each set of three, the columns layer on additional geographic controls. Panel a. reports conventional* TSLS *estimates. Panels b. and c. instead use the* LATE-REWEIGHT *and* IVCRC *methods. Dist. to amenities includes a house's distances to a body of water, a park, a school, and a highway. Panel a. reports standard errors clustered by border area, Panel b. reports the standard error of the neighborhood* TSLS *estimates' mean, and Panel c. reports bootstrap standard errors based on 500 draws resampling border areas.*

However, constructing the recentered instrument required specifying counterfactual minimum lot sizes. We next present the falsification tests suggested in Borusyak and Hull (2023) to test our specification of counterfactual minimums. Table 3 reports results from regressing the actual density exposure and formula instruments on several predetermined geographic variables we may expect to be associated with housing density. Consistent with the classic prediction of the Alonso-Muth-Mills model, Column (1) shows there is a statistically significant negative association between a parcel's actual density exposure and its distance to the city center even after conditioning on border side fixed effects. Similarly, because it is harder to build high density in steep terrain (Saiz, 2010), we find elevation and slope are associated with lower density exposure. Finally, we find distance away from water is associated with higher housing density. Column (2) shows that the formula prediction of the density exposure, while weakening these associations, still has similar associations with these geographic variables.

However, columns (3)-(5) validate the specification of counterfactual minimum lot sizes. The recentered density exposure is regressed on the same predetermined variables as in Columns (1) and (2). Significant associations between these variables would suggest a failure of Assumption 2. The regression coefficients and R^2 values fall significantly relative to Column (1) or (2), and the recentered instrument does not have a statistically significant association with any of these variables. Columns (4) and (5) further show that the recentered instrument is not associated with the expected instrument it is supposed to remove. In summary, these results are consistent with a valid specification of counterfactual minimum lot sizes.

A number of other robustness checks are considered in the appendix. Construction of the formula instruments required several arbitrary bandwidth choices including the density exposure buffer radius R, the buffer size around the border, and the number of nearest neighbors k to draw as counterfactual minimum lot sizes. Table B2 shows estimates analogous to Table 1 after setting k to every value between 2 and 15. Table B3 shows estimates where we narrow the spillover bandwidth and/or the buffer around the border. We also assess sensitivity to functional form. Table B4 reports results for two alternative specifications: a linear-linear specification and a log-log specification. These additional tests show our main findings are robust to a number of different assumptions from those in Table 2. We still see the same patterns: (a) average price effects are negative and economically meaningful; and (b) estimates from conventional TSLS/IV are larger in magnitude and more variable than those from estimators that target the sample ATE.

Table 3: Regression of density exposure and instruments on predetermined geography

	Dens. exp.	Predicted	Recentered		
	(1)	(2)	(3)	(4)	(5)
Dist. CBD (km)	-22.68***	-10.74*	0.335		0.174
	(4.314)	(5.918)	(1.263)		(1.305)
Elevation (m)	-0.925**	-1.077*	-0.0308		-0.0460
	(0.409)	(0.620)	(0.156)		(0.159)
Slope (%)	-3.916***	-1.849***	-0.162		-0.187
	(0.380)	(0.538)	(0.139)		(0.142)
Dist. water (km)	70.89***	-4.335	1.800		1.711
	(12.33)	(14.39)	(3.933)		(3.983)
μ_j				-0.0145	-0.0145
				(0.0109)	(0.0109)
Border X Muni FE	Yes	Yes	Yes	Yes	Yes
Observations	1,424,989	1,424,989	1,424,989	1,424,989	1,424,989
Number of Borders	2,432	2,432	2,432	2,432	2,432
Within R-sq.	0.0344	0.0039	0.0001	0.0014	0.0014
Joint F	49.11	4.63	0.59	1.75	0.85
Joint p-value	0.00	0.00	0.62	0.19	0.49

Note: This table reports coefficients from regressing the density exposure, the exposure predicted by formula (5), and the recentered density exposure prediction on geographic controls and expected instrument values.

6 Discussion

Before we discuss how our results relate to homeowner WTP, we first benchmark our reduced form estimates against three other articles that study the impact of new residential buildings on local housing prices (Diamond and McQuade, 2019, Asquith et al., 2023, Blanco and Sportiche, 2026). Comparing estimates across studies requires adjusting for the radius of spillovers and the average number of new units under consideration because price effects are typically stronger closer to a new development. Hence, estimates constructed at different radii are not directly comparable, and the average number of units in a new building varies across studies. In Appendix G.A., we standardize to a common spillover radius and unit count under simplifying assumptions and find that estimates across studies are similar in magnitude after adjustment. Most estimates range from -1 to -2%, with our TSLS estimate of -3.5% and IVCRC estimate of -1.3% bounding the price effects of the other three studies. As we argue below, we believe the larger negative price effect in the TSLS estimate is due at least in part to sorting by preference for density.

We now move to a discussion of what an average price effect of -1% implies about how much a homeowner is willing to pay to avoid density under a variety of different assumptions. First, the literature cited above has noted a potential problem with interpreting the price decrease from an increase in unit quantity as a disamenity effect if there is significant housing market segmentation within hyper-localized areas. In that case, the price decrease could at least partially reflect the change from the increase in housing supply. However, a NIMBY is still a homeowner. So, even under the conservative assumption that our estimate is due entirely to a supply effect, a homeowner would be willing to pay the property value loss associated with exposure to an additional 100 units. We estimate the average property value loss at approximately $4,000 (a roughly -1% price effect times the $400,000 typical house value in our sample).[13]

Heterogeneity in price effects across neighborhoods is more consistent with a hedonic model featuring sorting on preferences, implying the average homeowner will sustain a utility loss in addition to the property value loss. Under this interpretation, a -1% average effect yields an average homeowner MWTP of $4,000 per 100 units. While this identifies the average marginal willingness to pay for a small increase in density exposure, a welfare analysis for the nonmarginal 100-unit increase actually requires characterizing the entire MWTP function.

One possible approach is to impose assumptions on the shape of these functions. A convenient, and likely conservative, assumption is to assume that homeowners' MWTP functions are linear in density exposure. In this case, the combined asset and amenity value loss simply doubles to -$8,000. While this clearly is a back-of-the-envelope calculation, in Appendix G.C. we show that under weak assumptions on demand such a linear extrapolation yields a conservative lower bound in this setting.[14] Given that a 100-unit density increase is not a huge change in density in our context, we consider $8,000 a reasonable first-order approximation to the amount the average suburban homeowner would be willing to pay to avoid a 100-unit increase in density exposure.

[13]Under treatment effect heterogeneity, converting a price change measured in logs to dollars by multiplying by the average house price, while common in applications, requires an additional assumption. In Appendix G.B., we discuss the implicit assumption required to do this. We also report estimates from a linear-linear specification in B4. The fully linear specification does not require this additional assumption and yields estimates comparable in magnitude.

[14]Bishop and Timmins (2019) make a similar point about the curvature of MWTP functions in evaluating the MWTP to avoid an increase in crime. Appendix G.C. also considers point identification under a parametric assumption following Bajari and Benkard (2005).

A natural follow up question is whether this amenity value loss is large enough to justify existing density levels in U.S. suburbs. A standard result in urban economics is that maximizing total land value also maximizes social welfare (Brueckner, 1983). Another back-of-the-envelope calculation in Appendix G.D. concludes that in the typical U.S. suburb in our sample, an increase in the existing density level would increase total land value. This is consistent with earlier results in Glaeser and Ward (2009) pertaining to Boston area suburbs, and suggests density is inefficiently low not just in coastal markets such as San Francisco (which is the subject of another calculation in the appendix) where house prices are very high compared to construction costs. That said, the extensive heterogeneity in price effects qualifies this conclusion: for certain neighborhoods where homeowners have strong aversions to density and land values are not high relative to construction costs, existing density is not always 'too low' in an efficiency sense. This further analysis in Appendix G.D. validates recent calls for more evidence on how higher density development affects neighborhood amenity values (Baum-Snow, 2023, Baum-Snow and Duranton, 2025). Future research should try to incorporate this heterogeneity in losses into a broader conceptual framework that allows for comparison to the large benefits of increased density.

This has practical implications for thinking about how to address the expanding housing affordability problem in the U.S. There likely are communities with distastes for density that are many multiples of our ATE estimate (see Figure 12 above for our distribution of neighborhood TSLS estimates), making it very unlikely for densification to occur in those places without expensive compensation for the local homeowners. Hence, future research might also focus on where any available compensation would be most likely to increase density.

A related question is whether homebuilders themselves might be able to compensate the existing residents with a distaste for density. While this question should be studied as part of the larger issue of how the local political economy of permitting is constraining housing development, a very simple, back-of-the-envelope calculation suggests that the answer is 'no'. The typical house in our sample sells for about $400,000. Presuming a 10% profit for the builder yields $40,000. Our density exposure measure from Appendix Table B1 indicates there are 625 homes within 500 meters of the typical owner, so compensation is presumed to be paid to 625 households whose distaste for density averages $8,000 for an additional 100 houses. On average, that implies $80 in distaste per added home. Multiplying that number by the 625 owners who have to be compensated works out to $50,000, a figure more than 20 percent larger than the hypothetical profit. There are a number of assumptions in that very simple calculation that might prove inaccurate, but it certainly looks as if the builders will not be able to compensate the existing residents. Most, if not all, of that compensation will have to come from another source.

Given that the average suburban homeowner dislikes an increase in density, another natural follow up question is why they do not like it. Density could affect neighborhood character through a variety of channels. For example, there are many historical examples where zoning regulations were adopted intentionally to induce sorting on certain demographic characteristics such as race. We could include a possible mediating factor as a control in the hedonic regression specification and then reestimate the coefficient on density exposure. However, this would only identify the price effect explained by this factor if the mediator is as good as randomly assigned conditional on the instrument and observable controls. Hence, without additional assumptions, our econometric identification strategy requires us to remain agnostic about what precisely underlies our preference estimates.

We do provide suggestive evidence for two possible channels: demographics and crime.[15] Figure 13 plots how the share of the local population that is white and the number of gun violence incidents over the past decade vary as one moves closer to the border, going from the more regulated community towards and through the less regulated community as usual. The y-axis in each plot is an analogous exposure variable for each of these factors, i.e., the share white within 500 meters and the number of gun incidents within 500 meters. The x-axis assigns distance to the border as in Figure 8. On average, white share decreases and gun violence increases approaching an area with a smaller minimum lot size.

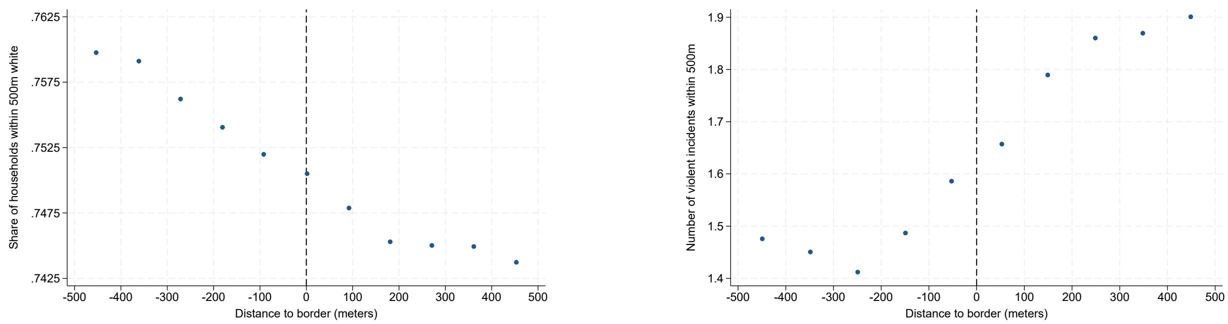
These patterns continue to hold if we estimate the effect of density exposure on these potential mediators by TSLS using the recentered instrument. Specifically, we reestimated the conventional TSLS specification from Table 1 using the white share and gun violence incidents, respectively, as the outcome variable. Doing so yields a negative and statistically significant coefficient on white share of -0.014, which implies that each additional 100 units of density exposure decreases the share white by 1.4 percentage points. The coefficient on gun violence incidents is positive and statistically significant at 0.72. This implies an additional 100 units in density exposure leads to an increase in gun violence of about 3/4 of an incident per decade on average.[16]

[15]There are many other candidates such as traffic congestion and noise pollution that well could be associated with more density. Our examples are just that and are not intended to be an exhaustive list.

[16]Complete regression output for these two specifications is in Table B6, along with estimates using the signed distance to the border as the instrument.

Figure 13: Potential mechanisms on distance to the border

(a) Share white (b) Gun violence incidents



Note: This figure presents binscatter plots of exposure measures for race and gun violence on signed distance to the border analogous to Figure 1. As before, outcomes are residualized on border fixed effects.

While statistically significant, the change in white share is not economically large. The standard deviation in our sample is 0.23 about a mean of 0.75, with an interquartile range running from 0.64 to 0.92. One possible interpretation of this is that race may not play a vital role in mediating the distaste for density because density exposure does not strongly shift race in our sample. However, we caution that our research design does not rule out the possibility that people care a lot about small changes in racial composition.

The impact on gun violence incidents is economically significant. The typical suburban neighborhood in our sample has no gun violence incidents over the past decade and the interquartile range runs from 0 to 1. Hence, it is possible that homeowners would care a lot about such an underlying effect if it represents a change from no gun violence to some gun violence. If so, policy makers interested in increasing housing supply might bundle new construction with some type of subsidy for more public safety (e.g., more police on the street) if densification is allowed. We reiterate that this speculation is not the result of a complete mediation analysis, as that would require a separate paper. However, it does indicate the potential payoff from subsequent research that deepens our understanding of the current permitting environment and how we might structure incentives to change it.

7 Conclusion

This paper derived a formula instrument using proximity to discontinuities in minimum lot sizes to identify the impact of local exposure to housing density on prices. The results suggests that, for a U.S. suburban house, an increase of 100 units within 500 meters lowers its value by at least 1% on average. However, treatment effects were heterogeneous across neighborhoods, consistent with a model where individuals sort according to their preference for density. Estimates for some subpopulations were as large as -7% in magnitude. Interpreting an average effect of -1% in the hedonic model, we estimate the typical suburban household would pay at least $8,000 to avoid this density increase. This -1% price externality suggests the typical U.S. suburb has an inefficiently low density in the sense that local land value is not being maximized.

Considering heterogeneity in this context is essential for future research. Some neighborhoods would be relatively more harmed from densification than others and understanding this dispersion in tastes has implications for the political economy of the building permit approval process. Another direction for future research is to develop a deeper understanding of why increased density is viewed as a disamenity from the perspective of existing owners. Insights into the mechanisms underlying this preference could suggest where policy interventions would lower the perceived costs of higher density.

References

Angrist, J. D. and Fernández-Val, I. (2013). Extrapolate-ing: External validity and overidentification in the late framework. In Acemoglu, D., Arellano, M., and Dekel, E., editors, *Advances in Economics and Econometrics: Tenth World Congress*, Econometric Society Monographs, pages 401–434. Cambridge University Press.

Arkhangelsky, D. and Imbens, G. W. (2024). Fixed effects and the generalized mundlak estimator. *The Review of Economic Studies*, 91(5):2545–2571.

Asquith, B., Mast, E., and Reed, D. (2023). Local effects of large new apartment building in low-income areas. *Review of Economics and Statistics*, 105(2):359–375.

Bajari, P. and Benkard, C. L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy*, 113(6):1239–1276.

Baum-Snow, N. (2023). Constraints on city and neighborhood growth: The central role of housing supply. *Journal of Economic Perspectives*, 37(2):53–74.

Baum-Snow, N. and Duranton, G. (2025). Housing supply and housing affordability. In Donaldson, D. and Redding, S., editors, *Handbook of Regional and Urban Economics*, volume 6. Elsevier, Amsterdam.

Benson, D., Masten, M., and Torgovitsky, A. (2022). ivcrc: An instrumental variables estimator for the correlated random coefficients model. *Stata Journal*, 22(3):469–495.

Bishop, K. and Timmins, C. (2019). Estimating the marginal willingness to pay function without instrumental variables. *Journal of Urban Economics*, 109:66–83.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2):577–599.

Blanco, H. and Sportiche, N. (2026). There goes the neighborhood? the local impacts of state policies that override municipal zoning. *American Economic Journal: Applied Economics*.

Blandhol, C., Bonney, J., Mogstad, M., and Torgovitsky, A. (2025). When is tsls actually late? *Working paper*.

Borusyak, K. and Hull, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica*, 91(6):2155–2185.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Brueckner, J. K. (1983). Property value maximization and public-sector efficiency. *Journal of Urban Economics*, 14:1–15.

Cattaneo, M., Crump, R., Farrell, M., and Feng, Y. (2024). On binscatter. *American Economic Review*, 114(5):1488–1514.

Chay, K. Y. and Greenstone, M. (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy*, 113(2):376–424.

Chen, J. and Roth, J. (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Cui, T. (2024). Did race fence off the american city? the great migration and the evolution of exclusionary zoning. *Working paper*.

Cui, T. and Been, V. (2025). Are zoning district borders color lines? how lot size requirements alter neighborhood racial composition. *Manuscript*.

Davidoff, T., Pavlov, A., and Somerville, T. (2022). Not in my neighbour's back yard? laneway homes and neighbours' property values. *Journal of Urban Economics*, 128:103405.

Davis, M., Glaeser, E., Gyourko, J., and Hankinson, M. (2026). Can side payments solve NIMBYism? experimental evidence. *Mimeo*.

Diamond, R. and McQuade, T. (2019). Who wants affordable housing in their backyard? An equilibrium analysis of low-income property development. *Journal of Political Economy*, 127(3):1063–1117.

Duranton, G. and Puga, D. (2023). Urban growth and its aggregate implications. *Econometrica*, 91(6):2219–2259.

Efron, B. (2016). Empirical Bayes deconvolution estimates. *Biometrika*, 103(1):1–20.

Glaeser, E. and Gyourko, J. (2018). The economic implications of housing supply. *Journal of Economic Perspectives*, 32(1):3–30.

Glaeser, E. L. and Ward, B. A. (2009). The causes and consequences of land use regulation: Evidence from greater boston. *Journal of Urban Economics*, 65:265–278.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gyourko, J. and Krimmel, J. (2022). The impact of local residential land use restrictions on land values across and within single family housing markets. *Journal of Urban Economics*, 126.

He, X., Pan, X., Tan, K. M., and Zhou, W.-X. (2023). Smoothed quantile regression with large-scale inference. *Journal of Econometrics*, 232:367–388.

Herkenhoff, K. F., Ohanian, L. E., and Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the u.s. economic slowdown. *Journal of Monetary Economics*, 93:89–109.

Huber, M. (2019). A review of causal mediation analysis for assessing direct and indirect treatment effects. *Technical report*.

Imbens, G. W. and Newey, W. K. (2009). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica*, 77(5):1481–1512.

Kashner, Z. and Ross, B. (2025). Measuring and mitigating traffic externalities. *manuscript*.

Kline, P., Rose, E. K., and Walters, C. R. (2022). Systemic discrimination among large u.s. employers. *The Quarterly Journal of Economics*, 137(4):1963–2036.

Kulka, A., Sood, A., and Chiumenti, N. (2026). Under the (neighbor)hood: Understanding interactions among zoning regulations. *Review of Economics and Statistics*, Forthcoming.

Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2):391–413.

Macek, J. (2024). Housing regulation and neighborhood sorting across the united states. *Working paper*.

Masten, M. A. and Torgovitsky, A. (2016). Identification of instrumental variables correlated random coefficients models. *The Review of Economics and Statistics*, 98(5):1001–1005.

McCulloch, S. (2026). Valuing neighborhood amenities with continuous spatial difference-in-differences. *Manuscript*.

Miguel, E. and Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.

Mogstad, M. and Torgovitsky, A. (2024). Instrumental variables with unobserved heterogeneity in treatment effects. In *Handbook of Labor Economics*, volume 5, pages 1–114. Elsevier.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46(1):69–85.

Narasimhan, B. and Efron, B. (2020). deconvolveR: A G-modeling program for deconvolution and empirical Bayes estimation. *Journal of Statistical Software*, 94(11):1–20.

Pennington, K. (2021). Does building new housing cause displacement? The supply and demand effects of construction in San Francisco. *Manuscript*.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55.

Roth, J. and Kwon, S. (2026). Testing mechanisms. *Review of Economic Studies*. Forthcoming.

Rothstein, R. (2017). *The Color of Law: A Forgotten History of How Our Government Segregated America*. Liveright Publishing, New York.

Saiz, A. (2010). Geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Severen, C. and Plantinga, A. J. (2018). Land use regulations, property values, and rents: Decomposing the effects of the california coastal act. *Journal of Urban Economics*, 107:65–78.

Song, J. (2025). The effects of residential zoning in U.S. housing markets. *Journal of Urban Economics*.

Turner, M., Haughwout, A., and van der Klaauw, W. (2014). Land use regulation and welfare. *Econometrica*, 82(4):1341–1403.

Walters, C. (2023). Empirical bayes and large-scale inference. `https://github.com/Mixtape-Sessions/Empirical-Bayes/tree/main`.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2nd edition.