Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Warding Off Development: Local Control, Housing Supply, and NIMBYs

Evan Mast*

February 2022

Abstract

Local control of land-use regulation creates a not-in-my-backyard (NIMBY) problem that can suppress housing construction, contributing to rising prices and potentially slowing economic growth. I study how increased local control affects housing production by exploiting a common electoral reform—changing from "at-large" to "ward" elections for town council. These reforms, which are not typically motivated by housing markets, shrink each representative's constituency from the entire town to one ward. Results from a variety of difference-in-differences estimators show that this decentralization decreases housing units permitted by 20%, with similar effects on multi- and single-family permits. Effects are larger in whiter and higher-income towns.

JEL Codes: R31, R38, H77

Keywords: Housing supply, land-use regulation, NIMBYism

*emast@nd.edu, W.E. Upjohn Institute for Employment Research, 300 S. Westnedge Avenue, Kalamazoo, MI 49007. Thanks to Shane Reed and Nathan Sotherland for excellent research assistance and to Brian Asquith, Ray Kluender, Davin Reed, and Amy Schwartz for helpful comments. The Center for Growth and Opportunity provided financial support.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Introduction

Housing construction in the United States has not kept pace with demand. Freddie Mac (2018) estimates that 2017 production fell 20% short of the level required to accommodate population growth and replace dilapidated structures. This not only affects housing markets, where costs continue to rise faster than incomes (Albouy et al. 2016), but may also slow aggregate economic growth (Hsieh and Moretti 2019). Accordingly, ways to address this shortfall have come to the forefront of the policy debate.

The highly decentralized control of land-use regulation and development approval may create a classic NIMBY (not-in-my-backyard) problem that contributes to the shortfall. Because new housing has diffuse benefits and concentrated costs—such as congestion, lost green space, and construction noise—people may prefer less construction near their residence than would be optimal for the town or region. Moreover, citizens can use their outsized influence over proposals in their area to restrict production to their preferred level. At a larger geography, towns within a region can use locally-controlled land-use regulations to behave similarly. These forces are anecdotally strong, leading to recent reforms that reduce local control of land-use regulations in an attempt to increase housing supply.[1] However, there is little empirical evidence on how local control of regulation affects housing supply, likely because of the difficulty of isolating relevant exogenous variation.

In this paper, I obtain causal estimates of the effect of increased local control on housing production by exploiting a common reform to town council elections. The reform is a change from "at-large" elections, in which citizens vote for candidates to represent the town as a whole, to "ward" elections, in which the town is divided into wards and each citizen votes for a single candidate to represent their area. This change decentralizes town governance, as at-large representatives should be responsive to the town's median voter, while ward representatives should respond to the median in their district. Moreover, informal practices

[1]Elmendorf (2019) describes California's requirements for municipal housing production and Oregon's recent prohibition of single-family zoning in most municipalities.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

like log-rolling and aldermanic privilege often provide ward representatives with significant control over issues in their district (Schleicher 2013).

I use a difference-in-differences (DiD) approach to evaluate the effect of switching from at-large to ward voting. I use all initially at-large towns as a control group and report results from both a two-way fixed effect (TWFE) model and the estimator developed in Callaway and Sant'Anna (2020), which I refer to as the CS estimator.[2] However, causal identification is complicated by the fact that towns choose their election type, making switches endogenous. Historically, this choice has been strongly influenced by racial dynamics. White majorities used at-large systems to limit minority representation, creating equity concerns that led to court cases and legislation in the 1980s that caused many towns to switch to ward voting (Trebbi et al. 2008). Towns that switched are thus larger and less white than other at-large towns. However, housing markets are similar, consistent with evidence that housing concerns rarely motivate switches (Hankinson and Magazinnik 2019), and demographic pre-trends are similar within counties, supporting the parallel trends identification assumption.

Event study results from both estimators suggest that, prior to changing to ward voting, treated and control towns were on very similar housing production trends. However, the number of units permitted each year falls sharply in treated towns immediately upon the reform's approval. This leads to DiD estimates of –1.75 units per 1,000 residents in the TWFE specification and –1.74 in the CS specification. This represents about 20% of average permits in the pre-treatment period. The effect size is similar for single- and multi-family permits, but it is significantly larger in towns with higher income and percent white, consistent with anecdotally stronger development opposition from homeowners and in areas with higher socioeconomic status (Fischel 2001). Lastly, I show that the results from the TWFE and CS estimators are quite similar in part because a large number of never-treated units are included in the sample, minimizing potential issues with negative weighting or "forbidden

[2]The sample contains municipalities of many types, but I refer to all as towns for simplicity.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

comparisons" (De Chaisemartin and d'Haultfoeuille, 2020; Borusyak et al., 2021).

Extensions and patterns within the main results suggest that estimates are not driven by indirect effects of non-housing policy set by ward councils. However, I cannot determine the relative importance of ward representatives blocking projects in their district and ward councils passing more restrictive regulation. More broadly, differences between at-large and ward councils are nuanced—an at-large council could exhibit strong NIMBY behavior over the areas that make up the base of their electorate, but not others.

On the whole, increased local control severely reduces housing production, which has important policy implications for areas trying to increase supply. Towns with ward elections may need to increase the set of structures that can be built without a special approval process or reduce the formal and informal policies giving council members more power over developments in their district. Extrapolating to a parallel setting, regional organizations or state governments may need to set minimum standards for housing production or limits on the restrictiveness of municipal land-use regulations. However, ward voting affects outcomes beyond housing production. The literature has notably shown that it increases minority representation,[3] and it could also affect production and location of locally undesirably land-uses in general (e.g. waste facilities, transit infrastructure, or homeless shelters).

The most directly related literature studies land-use regulation and housing markets. As reviewed in Gyourko and Molloy (2015), many studies have shown a robust relationship between tighter regulation and lower housing construction, as well as higher prices.[4] I first add to this literature by providing a quasi-experimental analysis of the switch from at-large to ward elections. In addition, like Parkhomenko (2018), Duranton and Puga (2019), and Khan (2020), this paper connects endogenous local land-use regulations and housing approval

[3]See Engstrom and McDonald (1981), Trebbi et al. (2008), and Abott and Magazinnik (2020).

[4]See, for example, Ihlanfeldt (2007), Glaeser and Ward (2009), Pollakowski and Wachter (1990), and Quigley and Raphael (2004).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

to the underlying institutions and economic forces that generate them. Similarly, Tricaud (2021) shows that consolidation of urban planning authority across French municipalities increased housing permits. Finally, concurrent work by Hankinson and Magazinnik (2019) is most closely related. They identify a large negative effect of ward voting on multifamily units in California between 2010 and 2018, as well as effects on permit location within towns.[5]

More broadly, this paper contributes to the literature on decentralization and fiscal federalism. Most directly related is the NIMBY literature, which has largely focused on mechanisms to fairly allocate socially necessary but locally undesirable land-uses (Frey et al. 1996, Feinerman et al. 2004, Levinson 1999). This study is among the first to obtain a causal estimate of how changes to local control affect such a land-use, illustrating forces that may apply in other settings and for other outcomes.

1 Background on Electoral Reforms

1.1 At-Large and Ward Elections

Towns in the U.S. generally use one of two methods to aggregate individual votes into seats on their council. In at-large elections, each citizen chooses from the same pool of candidates to elect representatives who represent the town as a whole. In ward elections, the city is divided into smaller wards (or districts) in which each citizen votes for a single candidate to represent their area. About two-thirds of towns use at-large elections for town council, about 15% of towns use a purely ward system, and 20% have some representatives elected at-large and some by ward (Clark and Krebs 2012).

Towns are able to choose between different election types, and a literature in both political science and economics studies the determinants of that choice. In theory, any majority bloc—whether defined by race, ideology, or ethnicity—can suppress the representation of a

[5]In addition, Feiock et al. (2008), Clingermayer (1994), and Lubell et al. (2009) document a negative correlation between ward voting and housing supply.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

minority bloc through at-large voting (Davidson and Korbel 1981). For example, a minority representing 15% of the population may not win a seat in an at-large election, but could constitute a majority within a smaller ward. In recent years in the United States, the majority bloc has typically been defined on racial lines. Trebbi et al. (2008) show that majority-white Southern towns implemented at-large systems after African American voting rights were strengthened by the Voting Rights Act, and a number of studies have found that ward voting indeed increases the representation of racial minorities[6]

Accordingly, changes between the two electoral rules have often been motivated by racial equity concerns. A key 1982 Supreme Court case held that at-large elections in Burke County, Georgia violated the Fourteenth Amendment rights of African-Americans, sparking a wave of switches to ward voting in the 1980s and 1990s.[7] While there are other reasons towns may switch, in the 1991 Form of Government Survey from the International City/County Management Association, 31% of the 77 switching towns cite either a court or state mandate. An additional 20% cite a government initiative, 29% a referendum, and 21% another (unlisted) reason. Importantly, housing markets are rarely cited as motivation for reform, as Hankinson and Magazinnik (2019) document in meeting minutes of several example towns.

1.2 Differences in Representative Incentives

Ward systems are more decentralized than at-large, since representation is tied to smaller groups of people. This leads ward and at-large representatives to face very different incentives. In classic models, at-large representatives are responsive to the town's median voter, while ward representatives should respond to the median voter in their district.[8] This difference is likely important for housing approvals. Within the ward containing a proposed

[6]Examples include Engstrom and McDonald (1981), Leal et al. (2004), Trebbi et al. (2008), and Abott and Magazinnik (2020).

[7]Rogers v. Lodge, 458 U.S. 613.

[8]See Trounstine (2010), Tausanovitch and Warshaw (2014), and Schleicher (2013).

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

development, a higher percentage of people will be affected by the project's concentrated costs—such as construction noise, lost green space or views, and congestion—than in the town as a whole. The ward's median voter will thus have a more negative opinion of a proposal than the town's median voter, making ward representatives less likely to support housing developments.

Differences between the two systems are reinforced by common practices and norms within ward councils. For example, in ward systems, the representative for the district containing a proposed development often has a disproportionate influence over its approval due to logrolling and "aldermanic privilege," informal practices in which council members defer to the home representative and, in return, receive the same treatment when an issue arises in their ward. This dynamic not only affects the approval chances of a particular project, but could also lead ward councils to pass more restrictive land-use regulation in general, in order to increase the proportion of projects that require a special council approval.

Of course, in the real world, switching to ward voting has a more complicated effect on housing approvals than in a simple median voter model. At-large council members likely cater to the areas that give them their majority, which should make it easier for such areas to deter new housing than other places in the town. This means that when a town switches to ward voting, it could have a larger effect in areas that previously were not part of the majority, changing the spatial distribution of new housing as well the total quantity. Hankinson and Magazinnik (2019) provide some evidence of this in a small sample of California cities.

Finally, this particular policy change speaks to a larger political economy story about how changes to decentralization affect incentives. Ward voting could influence other locally undesirable land-uses, and similar forces apply at other geographies. For example, allowing towns to set land-use regulation instead of counties or regional bodies could decrease housing supply.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

2 Data and Sample Construction

2.1 Data Sources

a. ICMA Data: Electoral reforms are identified using the International City/County Management Association (ICMA) Municipal Form of Government survey. The ICMA surveys towns (identified by Census Place codes) on their election type and form of government every five years and receives approximately 3,500-5,000 responses. Trebbi et al. (2008) also use this data to study changes in electoral rules and show that the sample is representative of the set of U.S. towns with population over 2,500.

To identify changes from at-large to ward voting, I use the survey question that asks whether a town made this change in the past five years, and, if so, when the reform was approved.[9] Across seven waves of the survey (from 1981 to 2011), this identifies nearly 300 municipalities that switched from at-large to purely ward elections at some point between 1970 and 2011. As shown in Appendix Figure 1, these switches occur across all decades, with 112 and 103 in the 1980s and 1990s and 21 and 30 in the 1970s and 2000s. Because the survey has an imperfect response rate, this approach will miss some changes in electoral rules. However, given the relative infrequency of reforms among survey respondents (<1%), this should be a small proportion of the control group.

b. Census Building Permit Data: Census Place-level data on new housing permits comes from the 1980 to 2018 waves of the Census Bureau Building Permit Survey. Each year, the Census Bureau surveys about 20,000 permit-granting local governments that represent 98% of U.S. housing production. The Census's follow-up surveys show that almost all permitted units are eventually built, making the permits a reasonable measure of housing construction. The data includes both the total number of units permitted and the number permitted in single-family and multifamily structures.[10] Appendix Figure 2 shows the time series of units

[9]The initial 1981 survey asks about reforms between 1970 and 1981.

[10]The Census imputes values in the event that a town does not respond to the survey, but

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

permitted, which is highly correlated with the business cycle.

c. LIHTC Data: Data on buildings receiving Low-Income Housing Tax Credit (LIHTC) financing is provided by the Department of Housing and Urban Development and covers all LIHTC buildings permitted between 1986 and 2015.

2.2 Sample Construction

I merge these three data sources to construct the analysis sample. After removing unincorporated areas from the building permit data, roughly 16,000 towns can be linked across all years of the survey. Of these, I keep only the 5,830 that can be matched to at least one ICMA survey wave. Next, about 6.5% of town-years and 60% of towns are matched to at least one LIHTC project, and I assume that town-years in the time frame covered by the LIHTC data that are not matched to the LIHTC data saw no such units permitted.

I then apply several restrictions to complete the final analysis sample. First, I drop towns with population below 2,500 in 1980, which are poorly represented in the ICMA data. Second, in order to improve the comparability of the treatment and control groups, I keep only towns that either report switching from at-large to ward or always report at-large voting. Third, for simplicity, I drop the additional 46 towns that switch from at-large to a mixed system with some ward and some at-large representatives.[11] The final sample contains approximately 2,500 towns in the years 1980-2018 and includes 238 at-large to ward transitions.

I drop these observations from the sample.

[11]Because the 1986 survey does not ask for the year the reform occurred, I classify all towns that reported switching in this time period as doing so in 1984 for the main analysis.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

3 Empirical Strategy

3.1 Specification

I study the effect of switching to ward voting using a differences-in-differences (DiD) strategy. The outcome variables are measures y of regular and low-income housing permits, and the treated group is towns that made the switch. In order to illustrate the rapidly evolving literature on DiD research designs with staggered treatments, I show estimates from both traditional two-way fixed-effect designs and the estimator developed in Callaway and Sant'Anna (2020).

Outcome variables: Because the effect of reforms is likely proportional to the flow of proposals, the primary dependent variable is total housing units permitted divided by town population in 1980 (in 1000s). This adjusts for town size without over-controlling for the effect of reforms, which may affect population. However, even with this normalization, the distribution (shown in Appendix Figure 3) is highly skewed. To prevent extreme values from driving estimates, I drop the five percent of towns with the highest average permitting per capita.[12] For low-income units, the primary dependent variable is an indicator for whether a town permitted any LIHTC units in a year.

Estimation samples: As described above, the full analysis sample includes treated towns and all towns that are always at-large. In addition, I construct a matched sample containing treated towns and a set of similar always at-large towns. I match each switching town to its ten nearest-neighbors among at-large towns in the same state, based on population, percent white, percent homeownership, median household income, and housing permits per capita,

[12]The median dropped town grew by about 460 percent from 1980 to 2010, versus 14 percent in the rest of the sample. In addition to the possibility that these extreme values may skew results, the housing approval process may be somewhat different in such a high growth setting.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

all measured in 1980.[13] I do this with replacement, allowing control towns to be matched to multiple treatment towns. I then include all towns that are matched to any treatment town in the control group.

TWFE estimator: I first estimate traditional two-way fixed effects models. I include state-year fixed effects to nonparametrically control for local trends and cluster errors at the state level. In order to investigate pre-trends and effect dynamics, I begin by estimating the following event study specification for town i in state s in year t:

$$y_{it} = \gamma_i + \alpha_{st} + \sum_{k \in C} \beta_k d_{it}^k + \epsilon_{it}. \tag{1}$$

I define C and d_{it}^k such that the d variables represent time-to-treatment dummies with a two-year window. I include only years from 1980 to 2003 (recall that the treatment variable is observed from 1970 to 2011) to ensure that all d_{it}^k are correctly defined, following Schmidheiny and Siegloch (2019).

Following the event study, I estimate an average effect of treatment by replacing the time-to-treatment variables with a single after-treatment dummy:

$$y_{it} = \gamma_i + \alpha_{st} + \beta \mathbb{1}(t > t_i^*) + \epsilon_{it}. \tag{2}$$

For this specification, the necessary identification assumption is that, in the absence of an electoral reform, housing permits would have changed in parallel for the treatment group and the control group of towns that were always at-large and towns that have not yet switched to ward in the same state. However, recent literature has highlighted that TWFE estimators may also make implicit comparisons that use already-treated units as a control group for units that are treated later. If treatment effects change over time, this may introduce bias or lead the weighted average of treatment effects returned by the estimator to contain negative

[13] To ensure that these are pre-reform characteristics, I drop the small number of reforms prior to 1980.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

weights (De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021). I assess the severity of this problem using tests suggested by the literature, as well as by comparing results to the CS estimator, which does not have this problem.

CS estimator: Second, I implement the estimator developed in Callaway and Sant'Anna (2020). It avoids the potential issues with heterogeneous treatment effects and negative weighting that arise in TWFE models by separately estimating treatment effects in each period t for all units that were treated in a given period g and explicitly excluding already-treated units from the control group. These comparisons, termed "group-time average treatment effects," are then aggregated into an average with well-defined positive weights. They can also be combined to produce a plot similar to a traditional event study. In addition, the estimator allows the researcher to condition on baseline covariates to relax the necessary identification assumption.

When implementing the CS estimator, I use the matched sample and use only never-treated units as controls. In addition, I include a town's state and 1980 values of population, percent white, owner-occupancy, and median household income as covariates, incorporating them using the doubly robust estimator developed in Sant'Anna and Zhao (2020). The identification assumption is then that treatment towns and never-treated towns must be on parallel counterfactual trends conditional on the included covariates.

3.2 Identification

The primary threats to the identification assumptions introduced above arise because towns in some cases choose whether to switch to ward voting. This may lead the treatment group to be different on both observable and unobservable characteristics. In the TWFE model, I attempt to minimize this issue by including state-specific time effects and by repeating the estimation with a matched sample, both of which make treatment and control groups more similar. In the CS model, I use the matched sample and include baseline covariates, which directly controls for observable differences between treatment and control. Before

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

turning to pre-trends in the event studies, I take a simpler approach to assess identification in Table 1.

[TABLE 1 ABOUT HERE]

In Panel A, I regress several characteristics on an indicator for whether a town ever switches, and, depending on the specification, either state-year or year fixed effects.[14] I include only the decennial census years in which these characteristics are observed and drop post-reform observations in treated towns. Fitting with the history discussed in Section 1.2, towns that switch are significantly larger and less white. However, the difference in housing units permitted per capita is statistically insignificant when including state-year effects, and the difference in the rate of homeownership is only 3.9% (5.9% of the mean). This similarity in housing markets fits with anecdotal and survey evidence that reforms are typically not driven by housing concerns (Hankinson and Magazinnik, 2019).

Panel B of Table 1 assesses differences in trends by repeating the exercise in Panel A using ten-year changes in characteristics as the dependent variable. With state-year fixed effects, the treatment indicator is statistically insignificant for all outcomes except percent white, suggesting that treated towns were evolving similarly to towns in the same state. Finally, in Panel C, I examine if reform timing is correlated with large changes in other characteristics in the preceding decade. I restrict to the sample of treated towns and regress changes in characteristics on a dummy for the decade prior to a reform and a set of town fixed effects, as well as either year or state-year fixed effects. The coefficients of interest are all small and statistically insignificant.

These three exercises support the necessary identification assumptions. Treatment and control towns appear to have similar housing markets and demographic trends but different baseline size and racial composition. The CS model is able to control for size and racial composition directly and rely on parallel trends conditional on covariates. While the TWFE

[14]Appendix Table 1 shows raw characteristics for treatment and control towns in the pre-period.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

model does not control for covariates directly, including state-year fixed effects greatly shrinks treatment-control differences, particularly when considering trends rather than levels.

4 Results

4.1 Main Results

Event study results for total housing units permitted in the TWFE and CS specifications are shown in Figures 1 and 2, respectively. While the confidence intervals are too large to precisely identify an effect in each time period, point estimates suggest minimal pre-trends, supporting the identification assumptions. However, permits fall sharply in treated towns following the reform's approval. The effect appears to reach a stable level by two years after the reform, where it remains until the end of the sample.[15] This lag may occur because the treatment is defined as the approval year, not the year of first ward election, and because reforms do not necessarily occur at the beginning of a time bin.

[FIGURE 1 ABOUT HERE]

[FIGURE 2 ABOUT HERE]

Although the point estimates support the parallel trends assumption, they are relatively noisy. To consider how potential violations of parallel trends could change results, I conduct the sensitivity test developed in Rambachan and Roth (2021). Results, shown in Appendix Figure 6, suggest that including reasonable deviations from parallel trends are unlikely to have a large effect on results.[16]

[15]Appendix Figures 4 and 5 repeat the exercises separately by unit type for the TWFE and CS estimators, respectively, and show similar patterns.

[16]Specifically, I follow Rambachan and Roth (2021) to estimate confidence intervals under a set of possible deviations from parallel trends. I allow deviations from the observed linear pre-trend of up to 0.15 (the maximum observed deviation) and compute the corresponding confidence intervals for the final event time coefficient in both specifications. Including these

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Table 2 shows the average effect of switching to ward voting, beginning with the TWFE estimates in Panel A. For total units, the after-treatment coefficient is -1.75 (S.E. = 0.491, p<0.01). This implies that the reform reduces permits by 1.75 per 1,000 town residents, corresponding to 20% of average permitting in the treatment group in the pre-reform period. The estimates for multi- and single-family units are -0.61 (S.E. = 0.261, p=0.024) and -1.14 (S.E. = 0.358, p<0.01). While the coefficient is larger for single-family than multi-family, both represent about 20% of average pre-reform production. Finally, Column 4 shows a small and statistically insignificant effect on the probability that any LIHTC units are permitted. For all outcomes, TWFE results using the matched sample (shown in Panel C) are extremely similar.[17]

Next, Panel B of Table 2 shows results from the CS estimator. The estimate for total units is quite close to the baseline TWFE specification: -1.74 (S.E. = 0.749, p=0.02). For multi-family and single-family units, the estimates are -1.01 (S.E. = 0.456, p = 0.03) and -0.92, (S.E. = 0.451, p=0.04). Finally, the LIHTC effect is larger than the TWFE but still statistically insignificant.[18]

[TABLE 2 ABOUT HERE]

---

pre-trends has very small effects on the center of the confidence intervals, reflecting the lack of trends in the point estimates. Of course, allowing larger deviations mechanically increases the width of each interval. Given that the event studies are relatively noisy to begin with, this leads the interval to cross zero in both cases when allowing deviations of 0.05 or greater. This highlights that the event study is somewhat under-powered and is best viewed as a diagnostic rather than a main result.

[17]TWFE event study estimates with the matched sample are shown in Appendix Figures 7 and 8.

[18]While using the matched sample significantly reduces the computational burden of the CS estimator, Abadie and Imbens (2006) show that inference following matching can be problematic. In Appendix Table 2, I run the baseline CS specification with the full sample and find very similar results.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Examining why the two approaches return very similar estimates illustrates some themes of the new DiD literature. First, the weighted average returned by TWFE estimate contains very few negative weights—only 0.6% of the total according to the methodology developed in De Chaisemartin and d'Haultfoeuille (2020). Similarly, in the decomposition developed in Goodman-Bacon (2021), "forbidden comparisons" that use already-treated units as the control group account for only 2.3% of the total. An important factor in both of these results is that many never-treated units are included in all specifications, ensuring that the lion's share of implicit comparisons use these units as controls.

In addition, note that the post-period effects in the event studies are relatively stable over time. This means that even in the event that an already-treated unit is implicitly used as a control, it does not provide a rapidly changing benchmark. This leads the average DD estimate from these forbidden comparisons to be -1.18, within the confidence interval of the primary estimate. Finally, to provide an estimate from a third approach, I follow Borusyak et al. (2021), which relies on an imputation method that is quite different from Callaway and Sant'Anna (2020). The point estimate for total units is -1.42 (S.E. = 0.48, $p < 0.01$).[19]

4.2 Heterogeneity

Differences across towns in how decentralization affects housing markets may be important for how metropolitan areas grow, as well as the impact of reforms to land-use or local control. A wealth of research, as well as anecdotal evidence, suggests that NIMBYism is strongest in high owner-occupancy areas, which also tend to be whiter and higher-income. Homeowners may be particularly opposed to development because of their direct stake in

[19] I consider alternative treatment definitions in the first two panels of Appendix Table 3. The effect shrinks but remains statistically significant when expanding the definition to include towns that report an at-large council in one survey wave and ward in the next but do not explicitly report a reform. This may reflect increased measurement error. Expanding the definition to include at-large to mixed switches slightly shrinks estimates.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

property values and because they are likely to remain in the area for longer than renters (Fischel 2001). Homeowners are also more likely to attend community meetings on new developments (Einstein et al. 2019) and less likely to support developments when the site is closer to their residence (Hankinson 2018).

However, because a bundle of characteristics are highly correlated with homeownership, it is difficult to study in isolation. I instead repeat the TWFE specification three times, allowing for different treatment effects in towns that were above-median on owner-occupancy rate, percent white, and median household income in 1980.[20] To allow for different trends among towns with different characteristics, I replace the year × state fixed effects with fixed effects for year × above median on the characteristic of interest in the specification. This yields the following estimating equation:

$$y_{it} = \gamma_i + \kappa_{ht} + \beta_1 \mathbb{1}(t > t_i^*) + \beta_2 \mathbb{1}(t > t_i^*) \times \mathbb{1}(h = 1) + \epsilon_{it}, \tag{3}$$

where h is an indicator for above median on the characteristic of interest. The identification assumption is that, in the absence of a reform, housing permits would have changed in parallel for treatment and control towns in the same category of the characteristic of interest. While this assumption is slightly different than the baseline specification, it prevents potentially heterogeneous trends across categories (which may have averaged out to zero in the baseline) from driving estimates of heterogeneity in the treatment effect.

[20] I focus on the TWFE estimator because the CS estimator does not allow the disaggregation of towns that were treated in the same year. To assess heterogeneity with the CS estimator, I run the baseline specification on subsamples that include only subsets of treated towns. For example, I report estimates in a subsample that includes only treated towns with income above the sample median and a subsample that includes only below-median treated towns. Results, shown in Appendix Table 4, show the same patterns as the TWFE estimates. In addition, Appendix Figure 9 shows the full distribution of estimated CS effects across towns that were treated in different years.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Results are shown in Table 3, with each panel displaying the estimates for a different town characteristic. For all characteristics, the effect on total units is larger in towns that rank above the median, though for owner-occupancy the difference is only significant at the 10% level. The differences are large, suggesting that towns with high socioeconomic status drive the average effects. For example, the effect in low-income towns is –0.81, while the interaction term for high-income towns is –3.9 (S.E.=1.12, $p < 0.01$).[21] Point estimates of the interaction term are also negative for both multi- and single-family units in all cases, but they are not always statistically significant for multifamily. This effect may be limited because above-median income towns, for example, produced only half as many multifamily units as other towns even before the treatment. Finally, Column 4 of Table 3 shows no evidence of heterogeneity for LIHTC permits.

[TABLE 3 ABOUT HERE]

4.3 Mechanisms

Even given correct causal estimates, the mechanism linking electoral reforms to housing supply may not be related to decentralization and NIMBYism. One possibility is that the administrative burden of changing electoral rules could slow the permitting process. However, it seems unlikely that this would persist for very long after the reform, and Panel C of Appendix Table 3 shows that towns that switch from ward to at-large voting do not see decreased permits.

Alternatively, over a longer time period, a ward-elected council could set non-housing

[21] Examining heterogeneity across three covariates introduces a multiple hypothesis testing problem. Because the analysis is exploratory and all variables measure a similar construct of socioeconomic status, I assess the false discovery rate (the probability that a rejected null hypothesis is actually true) by computing the sharpened-Q values developed by Benjamini et al. (2006) using the algorithm provided by Anderson (2008). I find q values of 0.004 for income and 0.011 for percent white.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

policies that make a town less attractive to developers. Prior evidence has robustly found that ward voting increases minority representation, but literature on other effects is limited and mixed. Several pieces of evidence suggest that non-housing effects of ward voting are unlikely to play a major role in my results. First, note that the effect of the reform on permits occurs immediately, meaning that other policy changes would have to be quickly anticipated and passed through to home prices and construction. Second, Appendix Table 4 repeats the main differences-in-differences specifications using other town characteristics as the dependent variable. Across the three specifications and four dependent variables, no estimates are significant at the 10% level. Finally, perhaps the strongest support for NIMBYism and decentralization as the underlying mechanism is provided by the role of town governments in the United States. They have much more extensive control over land-use regulation than other types of policy, so it makes sense that their effect on housing markets occurs primarily through this direct channel.

5 Conclusion

The shortfall in housing construction has attracted increased attention as U.S. housing costs rise. A contributing factor may be that land-use regulation and development approvals are controlled at a very local level, and policymakers aiming to spur construction have begun to target this feature of housing policy. For example, Oregon has limited how restrictive municipal zoning regulation can be, and Minneapolis has eliminated single-family zoning. Despite the increased attention, there is little empirical evidence on the importance of local control of housing regulation. In this paper, I exploit changes in electoral rules that provide rare exogenous variation in local control. My findings support conventional wisdom and anecdotal evidence of NIMBYism, as the reform reduces local housing production by 20%. However, I emphasize two caveats.

First, although I use electoral reforms for variation, my goal is not to estimate their total

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

welfare effect. Ward voting increases minority representation on town councils, and there may also be benefits from local representatives who are attuned to an area's needs. It may also affect other locally undesirable land-uses, changing the location or quantity of things like homeless shelters, heavy industry, and waste processing facilities. Second, while the forces underlying my proposed mechanism are quite general, I study a specific policy, and results may not generalize to every situation. Most proposed reforms consider decreasing local control, and applying my findings to this setting requires a symmetry assumption. Similarly, towns within a region are a parallel but perhaps different setting.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

References

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *econometrica*, 74(1):235–267.

Abott, C. and Magazinnik, A. (2020). At-large elections and minority representation in local government. *American Journal of Political Science*.

Albouy, D., Ehrlich, G., and Liu, Y. (2016). Housing demand, cost-of-living inequality, and the affordability crisis. NBER Working Paper #22816.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Clark, A. and Krebs, T. (2012). Elections and policy responsiveness. In *The Oxford Handbook of Urban Politics*, The Oxford Handbook of Urban Politics. Cambridge University Press.

Clingermayer, J. (1994). Electoral representation, zoning politics, and the exclusion of group homes. *Political Research Quarterly*, 47(4):969–984.

Davidson, C. and Korbel, G. (1981). At-large elections and minority-group representation: A re-examination of historical and contemporary evidence. *The Journal of Politics*, 43(4):982–1005.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Duranton, G. and Puga, D. (2019). Urban growth and its aggregate implications. Technical report, National Bureau of Economic Research.

Einstein, K., Palmer, M., and Glick, D. (2019). Who participates in local government? Evidence from meeting minutes. *Perspectives on Politics*, 17(1):28–46.

Elmendorf, C. S. (2019). Beyond the double veto: Housing plans as preemptive intergovernmental compacts. *Hastings Law Journal*, 71(1).

Engstrom, R. L. and McDonald, M. D. (1981). The election of blacks to city councils: Clarifying the impact of electoral arrangements on the seats/population relationship. *American Political Science Review*, 75(2):344–354.

Feinerman, E., Finkelshtain, I., and Kan, I. (2004). On a political solution to the NIMBY conflict. *American Economic Review*, 94(6):369–381.

Feiock, R., Tavares, A., and Lubell, M. (2008). Policy instrument choices for growth management and land use regulation. *Policy Studies Journal*, 36(3):461–480.

Fischel, W. (2001). *The Homevoter Hypothesis*. Harvard University Press.

Freddie Mac (2018). The major challenge of inadequate U.S. housing supply. Technical report.

Frey, B., Oberholzer-Gee, F., and Eichenberger, R. (1996). The old lady visits your backyard: A tale of morals and markets. *Journal of Political Economy*, 104(6):1297–1313.

Freyaldenhoven, S., Hansen, C., and Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–38.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Glaeser, E. and Ward, B. (2009). The causes and consequences of land use regulation: Evidence from Greater Boston. *Journal of Urban Economics*, 65(3):265–278.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Gyourko, J. and Molloy, R. (2015). Regulation and housing supply. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 1289–1337. Elsevier.

Hankinson, M. (2018). When do renters behave like homeowners? High rent, price anxiety, and NIMBYism. *American Political Science Review*, 112(3):473–493.

Hankinson, M. and Magazinnik, A. (2019). How electoral institutions shape the efficiency and equity of distributive policy. Working paper.

Hsieh, C.-T. and Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39.

Ihlanfeldt, K. (2007). The effect of land use regulation on housing and land prices. *Journal of Urban Economics*, 61(3):420–435.

Khan, A. R. (2020). Decentralized land-use regulation with agglomeration spillovers: Evidence from aldermanic privilege in chicago.

Leal, D. L., Martinez-Ebers, V., and Meier, K. J. (2004). The politics of latino education: The biases of at-large elections. *The Journal of Politics*, 66(4):1224–1244.

Levinson, A. (1999). NIMBY taxes matter: the case of state hazardous waste disposal taxes. *Journal of Public Economics*, 74(1):31–51.

Lubell, M., Feiock, R., and Ramirez De La Cruz, R. (2009). Local institutions and the politics of urban growth. *American Journal of Political Science*, 53(3):649–665.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Parkhomenko, A. (2018). The rise of housing supply regulation in the U.S.: Local causes and aggregate implications. (275).

Pollakowski, H. O. and Wachter, S. M. (1990). The effects of land-use constraints on housing prices. *Land Economics*, 66(3):315–324.

Quigley, J. M. and Raphael, S. (2004). Is housing unaffordable? Why isn't it more affordable? *Journal of Economic Perspectives*, 18(1):191–214.

Rambachan, A. and Roth, J. (2021). An honest approach to parallel trends. *Unpublished manuscript, Harvard University.[99].*

Sant'Anna, P. H. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122.

Schleicher, D. (2013). City unplanning. *Yale Law Journal*, 122.

Schmidheiny, K. and Siegloch, S. (2019). On event study designs and distributed-lag models: Equivalence, generalization and practical implications. CEPR Discussion Paper No. DP13477.

Tausanovitch, C. and Warshaw, C. (2014). Representation in municipal government. *American Political Science Review*, 108(3):605–641.

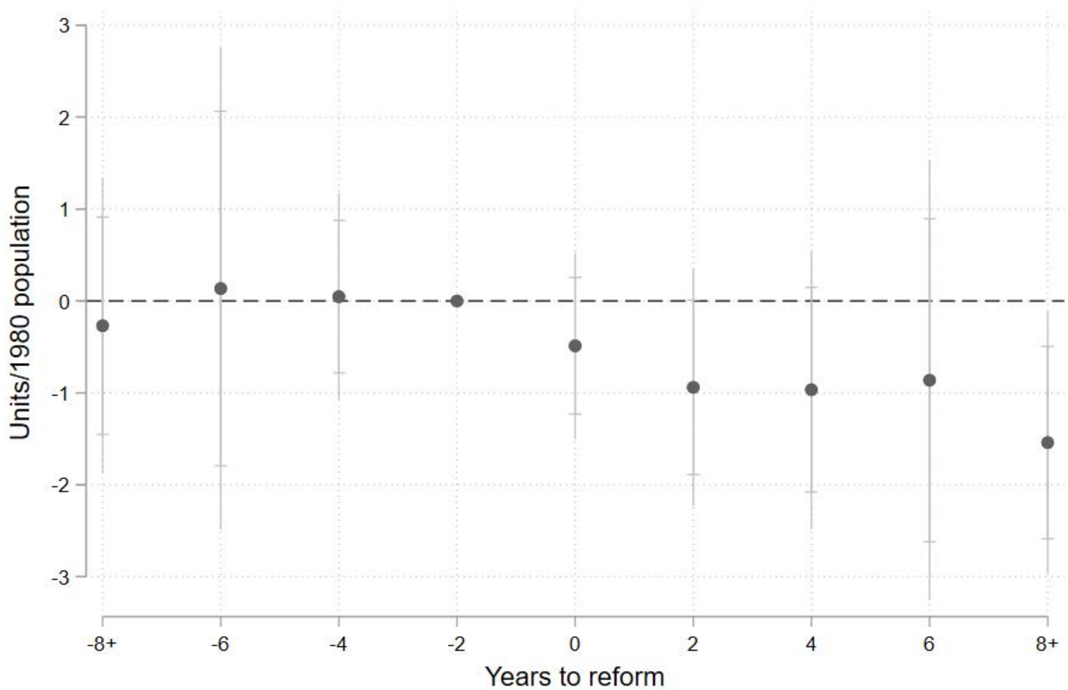
Trebbi, F., Aghion, P., and Alesina, A. (2008). Electoral rules and minority representation in U.S. cities. *Quarterly Journal of Economics*, 123(1):325–357.

Tricaud, C. (2021). Better alone? evidence on the costs of intermunicipal cooperation.

Trounstine, J. (2010). Representation and accountability in cities. *Annual Review of Political Science*, 13:407–423.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

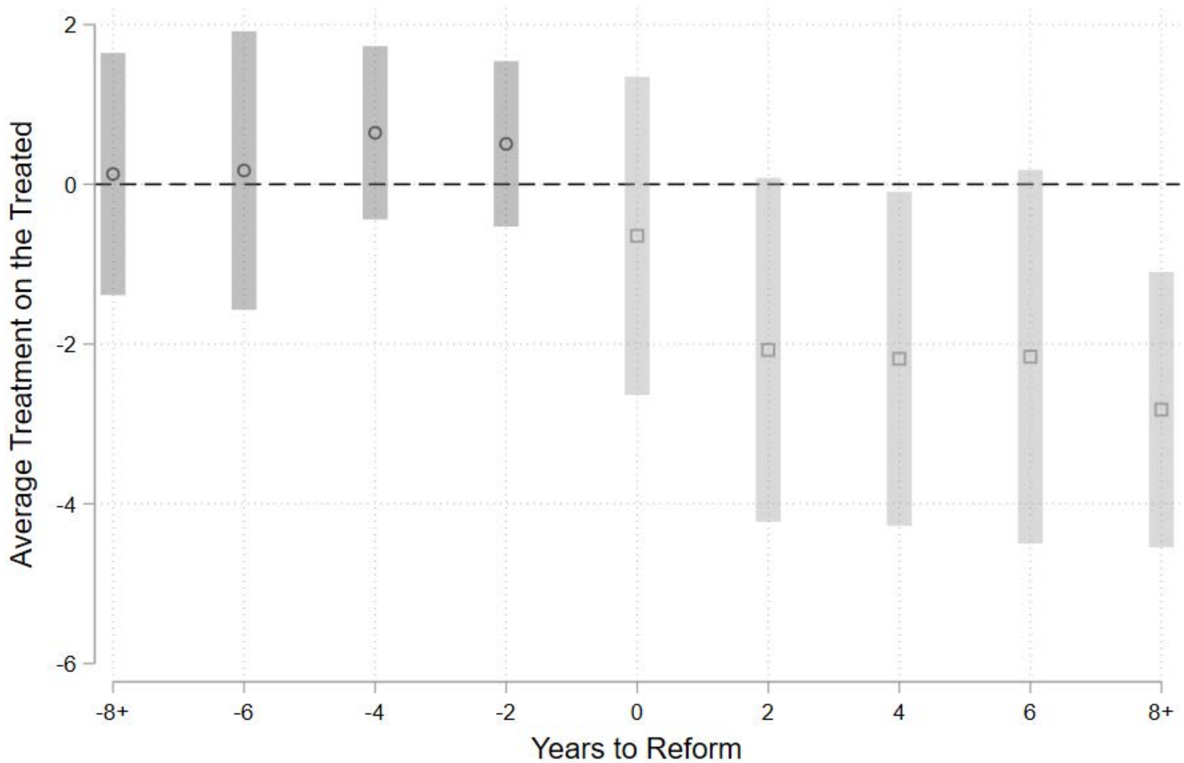
Figure 1: Baseline Event Study for Total Units Permitted



Note: This figure shows event study coefficients from the estimation of Equation 1 with units permitted divided by 1980 population (in 1000s) as the dependent variable. The dummies for years to treatment are bundled into two-year bins, with 0 containing the year that a town approved a switch from at-large to ward elections. The specification includes town fixed effects and state-year fixed effects, and errors are clustered at the state level. The bars represent 95% uniform confidence bands, following Freyaldenhoven et al. (2019), and the hash marks represent 95% pointwise confidence intervals.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Figure 2: CS Event Study for Total Units Permitted



Note: This figure shows treatment effects in two-year bins of event time from the Callaway and Sant'Anna (2020) estimator. The y-axis represents the average treatment effect on the treated. Year 0 is the two-year period in which a town approved a switch from at-large to ward elections, and the dependent variable is units permitted divided by 1980 population (in 1000s). The specification includes controls for a town's state and 1980 levels of population, income, owner-occupancy, and percent white. Errors are clustered at the state level, and bars represent 95% uniform confidence bands. The comparison group is never treated towns, and the sample is the matched sample described in Section 3.1.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Table 1: Comparison of Treatment and Control Towns

	Population		Percent White		New Units per 1000		Owner-Occupancy	
Panel A: Characteristics levels								
Ever switch	19337	24530	-0.102	-0.065	1.513	0.872	-0.033	-0.039
(S.E.)	(6441)	(6369)	(0.019)	(0.012)	(0.623)	(0.528)	(0.013)	(0.010)
Constant	16005	18621	0.90	0.84	5.40	3.95	0.67	0.66
(S.E.)	(2234)	(205)	(0.015)	(0.000)	(0.684)	(0.017)	(0.013)	(0.000)
N	9,733	9,733	9,733	9,733	9,733	9,733	9,733	9,733
Time FE	Year	State-Year	Year	State-Year	Year	State-Year	Year	State-Year
Panel B: Decadal percent change in characteristics								
Ever switch	0.054	0.036	-0.016	-0.017	-0.166	-0.085	0.011	0.005
(S.E.)	(0.049)	(0.036)	(0.010)	(0.007)	(0.065)	(0.046)	(0.005)	(0.004)
Constant	0.18	0.12	-0.05	-0.05	-0.08	-0.14	-0.01	-0.02
(S.E.)	(0.046)	(0.001)	(0.014)	(0.000)	(0.078)	(0.001)	(0.005)	(0.000)
N	9,072	9,072	9,072	9,072	6,567	6,567	9,320	9,320
Time FE	Year	State-Year	Year	State-Year	Year	State-Year	Year	State-Year
Panel C: Decadal percent change in characteristics (only treated observations)								
Decade before switch	-0.056	-0.045	0.005	0.008	-0.074	0.117	0.010	0.004
(S.E.)	(0.041)	(0.054)	(0.009)	(0.011)	(0.078)	(0.067)	(0.008)	(0.009)
Constant	0.14	0.14	-0.07	-0.07	-0.19	-0.23	-0.02	-0.02
(S.E.)	(0.009)	(0.012)	(0.002)	(0.002)	(0.011)	(0.009)	(0.002)	(0.002)
N	852	815	852	815	618	593	865	831
FE	Town, Year	Town, State-Year	Town, Year	Town, State-Year	Town, Year	Town, State-Year	Town, Year	Town, State-Year

Note: This table compares characteristics of treatment and control towns. Panel A depicts regressions of the variable in the column heading on an indicator for if a town ever switches from at-large to ward voting and a vector of either year or state-year fixed effects. Years after reform are dropped for treated towns. Panel B repeats the exercise using decadal changes as the dependent variable. Panel C restricts to treated towns and regresses changes in the dependent variable on an indicator for the decade before the reform. In all specifications, only decennial census years are included and errors are clustered at the state level. Observation counts in Panels B and C differ across outcomes because owner-occupancy is better reported than other outcomes in the 1970 Census and permit data is not available before 1980.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Table 2: Difference-in-Differences Results

	All Units	Multifamily Units	Single-Family Units	1(LIHTC)
Panel A: Baseline, TWFE Estimator				
After*treated	-1.752	-0.611	-1.141	0.022
(S.E.)	(0.491)	(0.261)	(0.358)	(0.018)
Constant	5.89	1.83	4.06	0.08
(S.E.)	(0.030)	(0.016)	(0.022)	(0.001)
Observations	85,727	85,727	85,727	65,278
Control group	All	All	All	All
Time FE	State-year	State-year	State-year	State-year
Panel B: Matched, CS Estimator				
After*treated	-1.740	-1.01	-0.921	0.037
(S.E.)	(0.749)	(0.456)	(0.451)	(0.023)
Observations	37,563	37,563	37,563	28,611
Control group	Matched, never treated	Matched, never treated	Matched, never treated	Matched, never treated
Panel C: Matched, TWFE Estimator				
After*treated	-1.924	-0.677	-1.246	0.021
(S.E.)	(0.475)	(0.210)	(0.383)	(0.017)
Constant	6.62	2.15	4.47	0.11
(S.E.)	(0.059)	(0.026)	(0.047)	(0.002)
Observations	37,563	37,563	37,563	28,611
Control group	Matched	Matched	Matched	Matched
Time FE	State-year	State-year	State-year	State-year

Note: This table shows results from two-way fixed effect and Callaway and Sant'Anna (2020) estimators. The column heading denotes the dependent variable, which is normalized by 1980 population (in 1000s) for all but the LIHTC outcome. Panel A represents the TWFE estimator with the full analysis sample. Panel B shows results from the CS estimator and the matched sample discussed in Section 3.1. Panel C shows results from the TWFE estimator and the matched sample. The TWFE specifications include town fixed effects and state-year time effects. The CS specification uses never-treated observations as the control group and includes includes controls for a town's state and 1980 levels of population, income, owner-occupancy, and percent white. Errors are clustered at the state level.

Review of Economics and Statistics Just Accepted MS.
https://doi.org/10.1162/rest_a_01192

Table 3: Heterogeneity by Town Characteristics (TWFE Estimator)

	All Units	Multifamily Units	Single-Family Units	1(LIHTC)
Panel A: Owner-Occupancy				
After*treated*above-median	-2.087	-0.271	-1.816	-0.039
	(1.070)	(0.438)	(0.813)	(0.032)
After*treated	-0.931	-0.630	-0.300	0.030
	(0.499)	(0.256)	(0.365)	(0.028)
Observations	85,166	85,166	85,166	63,082
Control group	All	All	All	All
Time FE	Above-median × year	Above-median × year	Above-median × year	Above-median × year
Panel B: Median Income				
After*treated*above-median	-3.903	-1.452	-2.451	0.022
	(1.127)	(0.784)	(0.954)	(0.032)
After*treated	-0.814	-0.369	-0.445	0.010
	(0.440)	(0.233)	(0.315)	(0.021)
Observations	85,166	85,166	85,166	63,082
Control group	All	All	All	All
Time FE	Above-median × year	Above-median × year	Above-median × year	Above-median × year
Panel C: Percent White				
After*treated*above-median	-3.575	-0.650	-2.925	0.027
	(1.340)	(0.591)	(1.160)	(0.038)
After*treated	-0.707	-0.502	-0.205	0.000
	(0.675)	(0.284)	(0.452)	(0.027)
Observations	85,166	85,166	85,166	63,082
Control group	All	All	All	All
Time FE	Above-median × year	Above-median × year	Above-median × year	Above-median × year

Note: This table shows how the effect of ward voting varies in towns with different owner-occupancy rates, median household incomes, and percent white. Each panel estimates heterogeneity for a different characteristic. All use the TWFE specification shown in Equation 3, which is similar to the primary exercise, but also includes the after × treated × above-median dummy and a vector of year × above-median fixed effects. The column heading denotes the dependent variable, which is normalized by 1980 population (in 1000s) for all but the LIHTC outcome, and errors are clustered at the state level.