

Online Appendix

for “Why zoning is too restrictive”

Jack Favilukis* Jaehee Song†

1 Supplementary Figures and Tables

Table A.1 — Additional data sources

Variable	Source
Land area by uses (total residential land area, % residential, % agricultural, % commercial, % industrial)	CoreLogic property tax records in 2018
% residential built after 1850, 1910, 1970	CoreLogic property tax records in 2018
1969 demographics (log population, % white)	U.S. County Population Data (SEER)
1940 demographic and economic characteristics (log population, % white, % Black, % Hispanic, % homeownership, log avg home value, log avg rent, % with nonwage income, log avg income, 90/10 income ratio)	IPUMS complete-count Census
Presidential election returns (% dem. votes from 2000 to 2020)	MIT Election Data + Science Lab
Weather condition (avg temperature and precipitation yearly and in Jan)	PRISM Weather data
Land developability index from 2021	National Land Cover Database
2020 industry shares (measured by # of establishments at the 2-digit NAICS level)	County Business Patterns

*Sauder School of Business, University of British Columbia. Email: jack.favilukis@sauder.ubc.ca

†Leeds School of Business, University of Colorado. Email: jaehee.song@colorado.edu

Table A.2 — Summary statistics

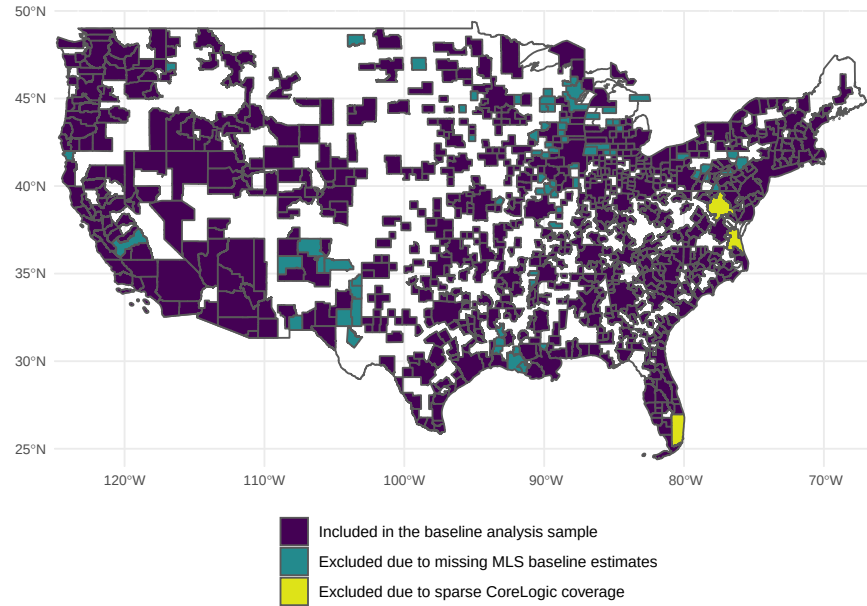
This table reports summary statistics, including the mean, standard deviation, 10th percentile, 25th percentile, 50th percentile (median), 75th percentile, and 90th percentile for key variables at the CBSA level. The reported variables include various measures of zoning stringency, such as the average of municipality-level median minimum lot sizes (baseline), the exponential of the average log minimum lot sizes, the first quartile of municipality-level medians, the third quartile of municipality-level medians, and the weighted average of the Wharton Density Restriction Index (DRI), which serve as outcome variables. Additionally, the table provides summary statistics for zoning jurisdiction HHI (baseline) and the number of municipalities, which serve as treatment variables. Finally, the table presents summary statistics for instrumental variables, including HHI as of 1875 (baseline), the number of municipalities as of 1875, and the number of rivers crossing CBSAs, which serve as instruments.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	SD	10th	25th	50th	75th	90th
<i>Zoning stringency</i>							
avg. median MLS (in SF)	39,570	30,481	12,416	19,075	30,544	50,177	78,233
exp(avg. log MLS) (in SF)	26,539	15,197	11,519	16,140	23,694	32,899	44,249
Q1 median MLS (in SF)	16,116	11,970	7,405	9,210	12,632	17,890	27,225
Q3 median MLS (in SF)	48,919	52,104	11,107	18,491	42,307	46,282	88,209
Wharton DRI ($N = 574$)	1.78	1.09	0.59	1	1.58	2.53	3.40
<i>Municipal fragmentation</i>							
zoning jurisdiction HHI	4,639	3,298	523	1,255	4,474	7,806	9,189
# municipality	23.3	41.1	4	6	12	24	50
<i>Instruments</i>							
HHI as of 1875	5,583	3,928	559	1,217	5,654	10,000	10,000
# municipality as of 1875	13.15	24.77	0	1	4	15.75	31
# rivers	34.7	37.4	9	14	22	41	74

Figure A.1 — Sample CBSA maps

Panel (a) maps the CBSA sample used in the baseline analysis. CBSAs are classified as included (baseline sample), excluded due to missing baseline minimum-lot-size (MLS) estimates, or excluded due to sparse CoreLogic parcel coverage. Panel (b) restricts to CBSAs included in the baseline sample and indicates whether any municipalities existed as of 1875 in our historical municipal dataset.

(a) Baseline analysis sample coverage



(b) Municipality identification as of 1875 (within baseline sample)

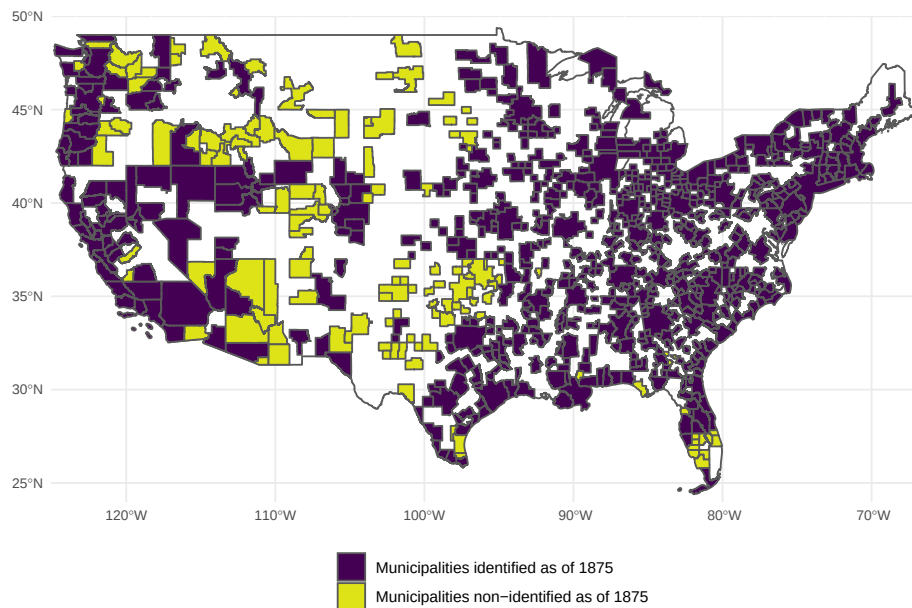


Figure A.2 — Correlation between zoning jurisdiction HHI and MLS

The figure depicts the scatter plots (black dots) and their fitted lines (blue solid lines) where the x-axis is zoning jurisdiction HHI, and the y-axis is minimum lot size, aggregated at the CBSA level by taking the mean of median municipality-level MLS. Both axes are on a logarithmic scale. It also depicts binned data points with 20 bins (red crosses) and shows the coefficient estimate and R^2 from the univariate regression of log minimum lot size on log HHI at the CBSA level ($N = 832$).

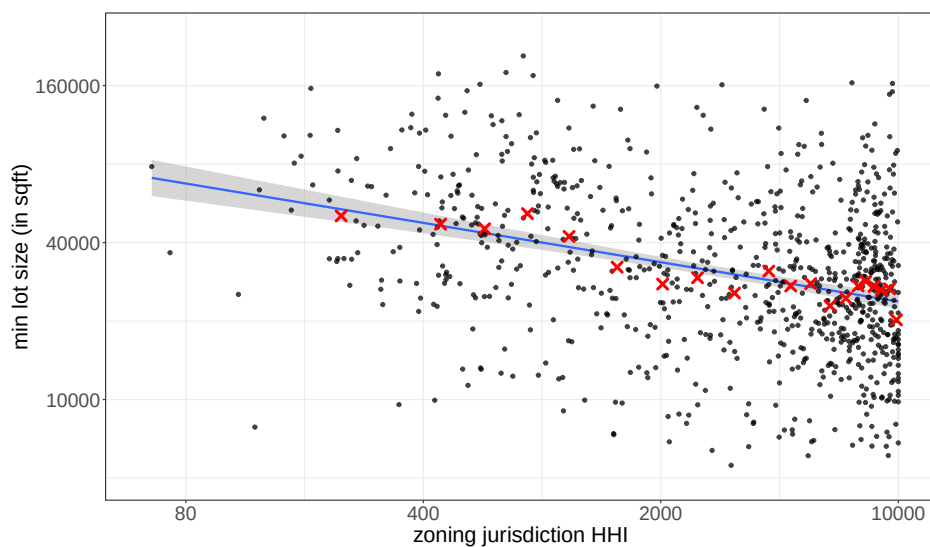


Figure A.3 — Correlation between key variables and log HHI

The figure depicts the coefficient estimates and 95% confidence intervals from univariate OLS regressions where the regressor is log HHI and the dependent variables (y-axis) are key control or outcome variables used in the main empirical analysis. All coefficients are standardized by the standard deviation of the respective dependent variables.

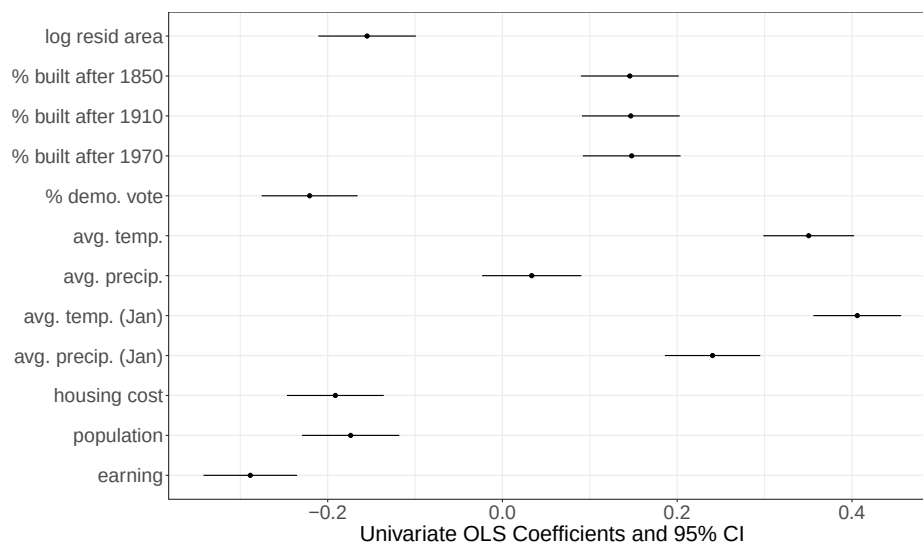


Table A.3 — Estimated effects on other CBSA-level outcomes (log specifications)

This table presents IV regression results from the main paper’s specification with alternative outcomes, where the dependent variables, all in logs, are: median monthly housing cost (Column 1), total population in thousands (Column 2), the number of individuals who lived in a different CBSA one year prior (Column 3), and median earnings (Column 4). Columns (1), (2), and (4) use 2015 ACS 5-year estimates, and Column (3) is based on the ACS public-use microdata. Each column reports the baseline CBSA-level IV regressions, incorporating the full set of control variables described in Table A.1. Due to space constraints, the table omits the coefficients for the control variables.

	(1) median cost (\$)	(2) population (k)	(3) # in-migrants (k)	(4) median earnings (\$)
$\log \widehat{MLS}$	0.1217*** (0.0346)	0.0785 (0.0529)	-0.3493 (0.2638)	0.1080*** (0.0365)
Observations	832	832	261	832
Adjusted R ²	0.70602	0.97395	0.71307	0.25833

Significance: ***: 0.01, **: 0.05, *: 0.1

Table A.4 — Estimated effects on other CBSA-level outcomes (alternative specification)

This table reports regression results from a specification analogous to the main paper’s baseline IV regression, where the treatment variable is $\log HHI$ (instrumented with $\log HHI$ as of 1875) and the dependent variable varies by column. The dependent variables are: median monthly housing cost (Column 1), total population in thousands (Column 2), the number of individuals who lived in a different CBSA one year prior (Column 3), and median earnings (Column 4). Columns (1), (2), and (4) use 2015 ACS 5-year estimates, and Column (3) is based on the ACS public-use microdata. Each column reports the baseline CBSA-level IV regressions, incorporating the full set of control variables described in Table A.1. Due to space constraints, the table omits the coefficients for the control variables.

	(1)	(2)	(3)	(4)
	median cost (\$)	population (k)	# in-migrants (k)	median earnings (\$)
$\log \widehat{HHI}$	-30.88*** (6.769)	121.1** (50.97)	7.116** (2.886)	-532.2** (207.9)
Observations	906	906	272	906
Adjusted R ²	0.74478	0.44173	0.65623	0.43919

Significance: ***: 0.01, **: 0.05, *: 0.1

Figure A.4 — Model results: Migration across metros (extended range)

This figure presents results from the *Migration across metros* model. It is identical to the corresponding model figure in the main paper but extends the x-axis to a larger number of municipalities. The x-axis represents the number of municipalities N , while the y-axis plots metro-wide housing quantity h (equivalently, the inverse of zoning restrictions), house prices p , congestion ϕ , number of migrants m , average wage w , and utility u in each panel.

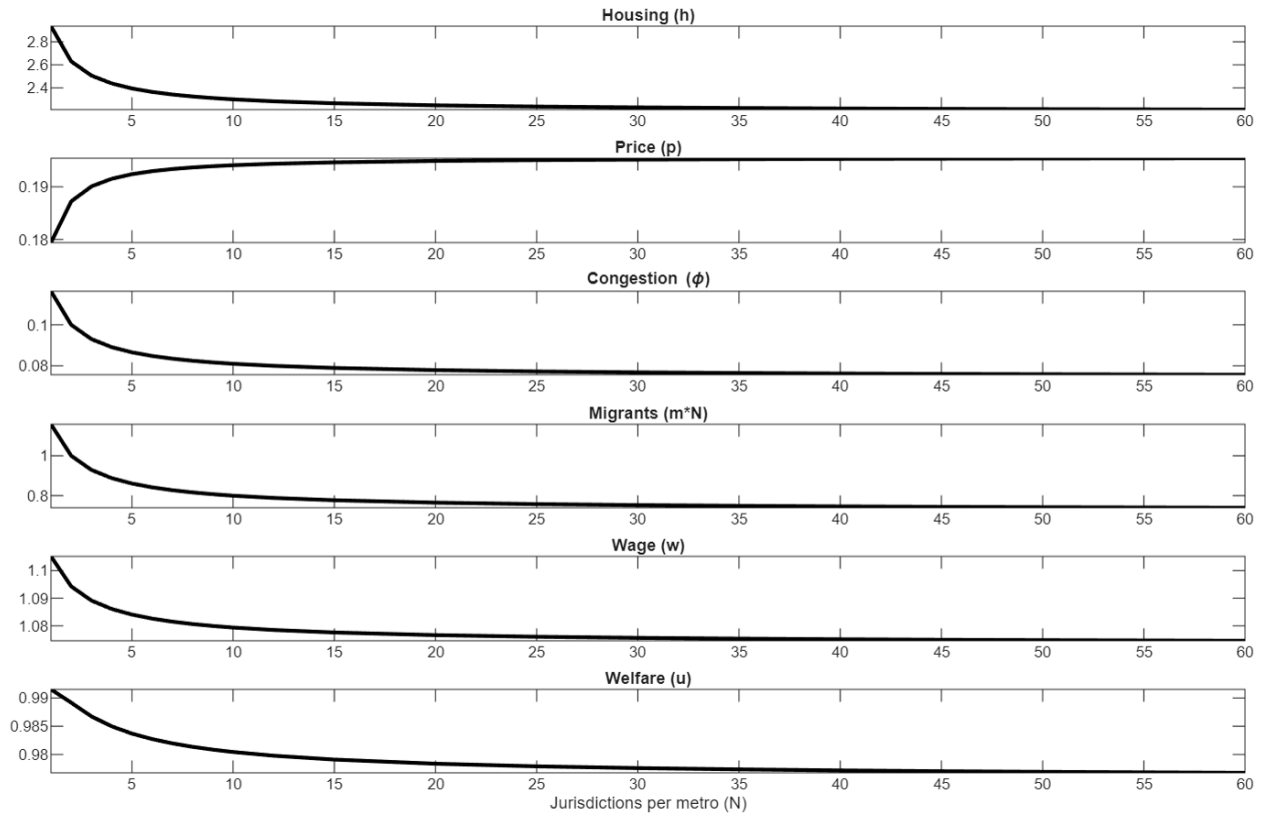


Figure A.5 — Model results: Migration within metro

This figure presents results from the *Migration within metro* model. The x-axis represents the number of municipalities N , while the y-axis plots metro-wide housing quantity h (equivalently, the inverse of zoning restrictions), house prices p , congestion ϕ , number of migrants m , average wage w , and utility u in each panel.

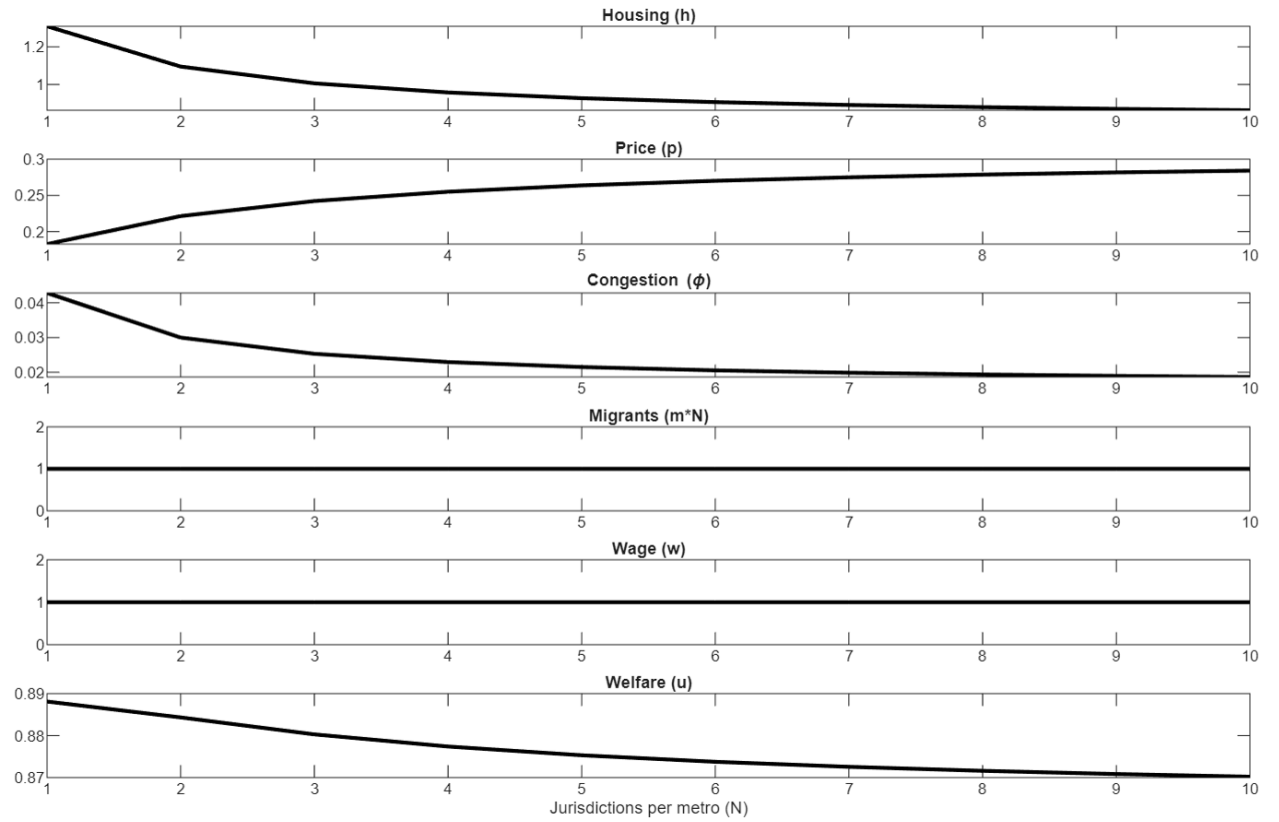


Figure A.6 — Model results: Migration across metros with poorer migrants

This figure presents results from the *Migration across metros* model with migrants who are poorer than incumbents. The parameters remain the same as in the main paper's calibration, except ψ is reduced from 1.0 to 0.7 and $\underline{u}_m$ from 0.835 to 0.574. The x-axis represents the number of municipalities N , while the y-axis plots metro-wide housing quantity h (equivalently, the inverse of zoning restrictions), house prices p , congestion ϕ , number of migrants m , average wage w , and utility u in each panel.

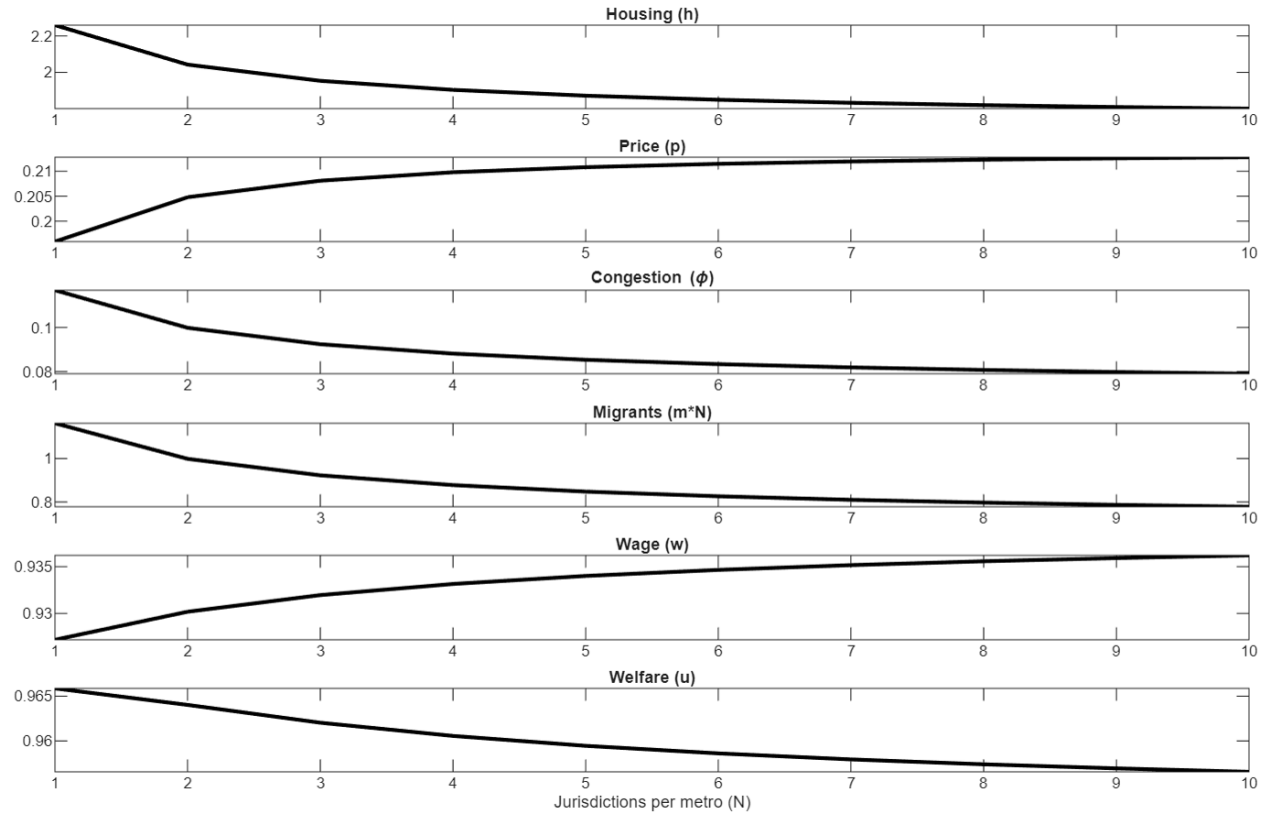


Figure A.7 — Model results: Migration across metros with public goods, taxes, and heterogeneity in income

This figure presents results from the *Migration across metros* model extended to include public goods, taxes, and income heterogeneity, as described in the main paper. The x-axis represents the number of municipalities N , while the y-axis plots metro-wide housing quantity h (equivalently, the inverse of zoning restrictions), house prices p , congestion ϕ , number of migrants m , average wage w , and utility u in each panel. The parameters are identical to the baseline *Migration across metros* model, additionally $\gamma = 0.04$, $\theta = 0.05$, and $\underline{u}_m$ is lowered to 0.71.

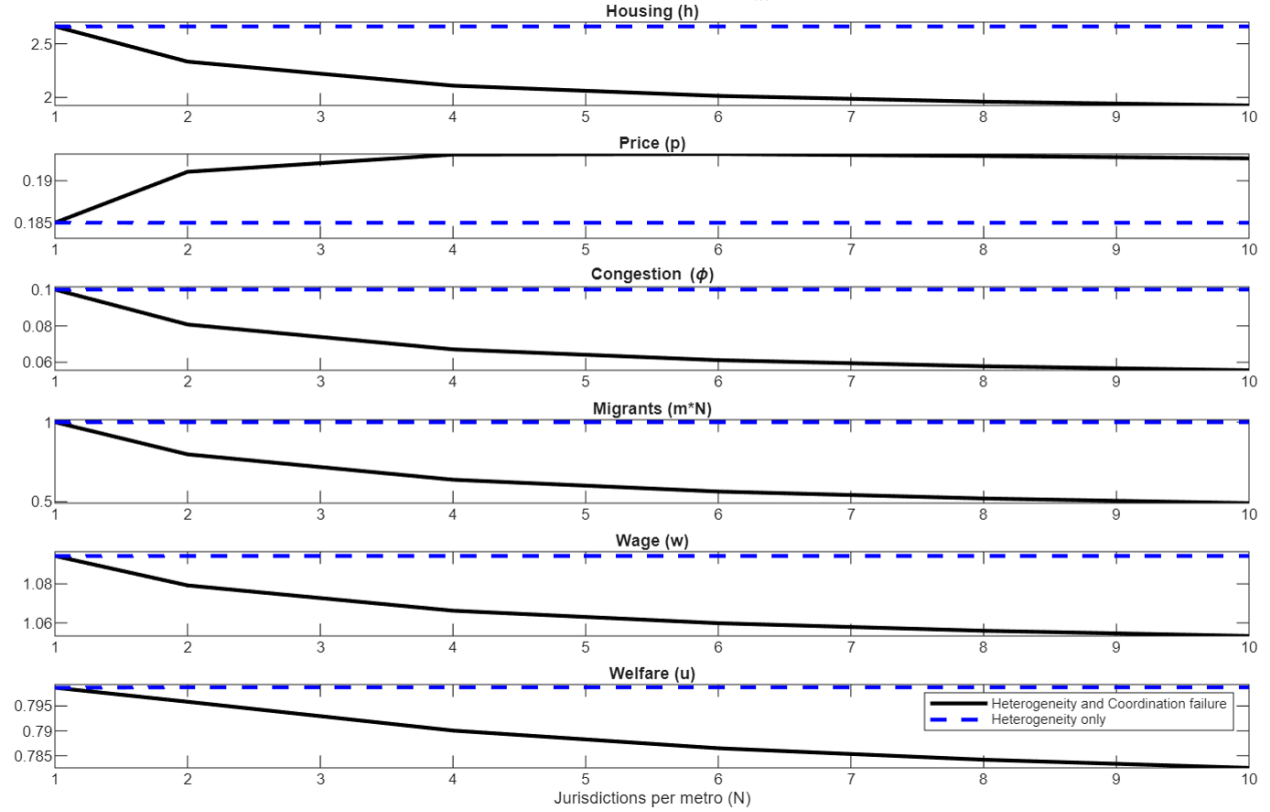
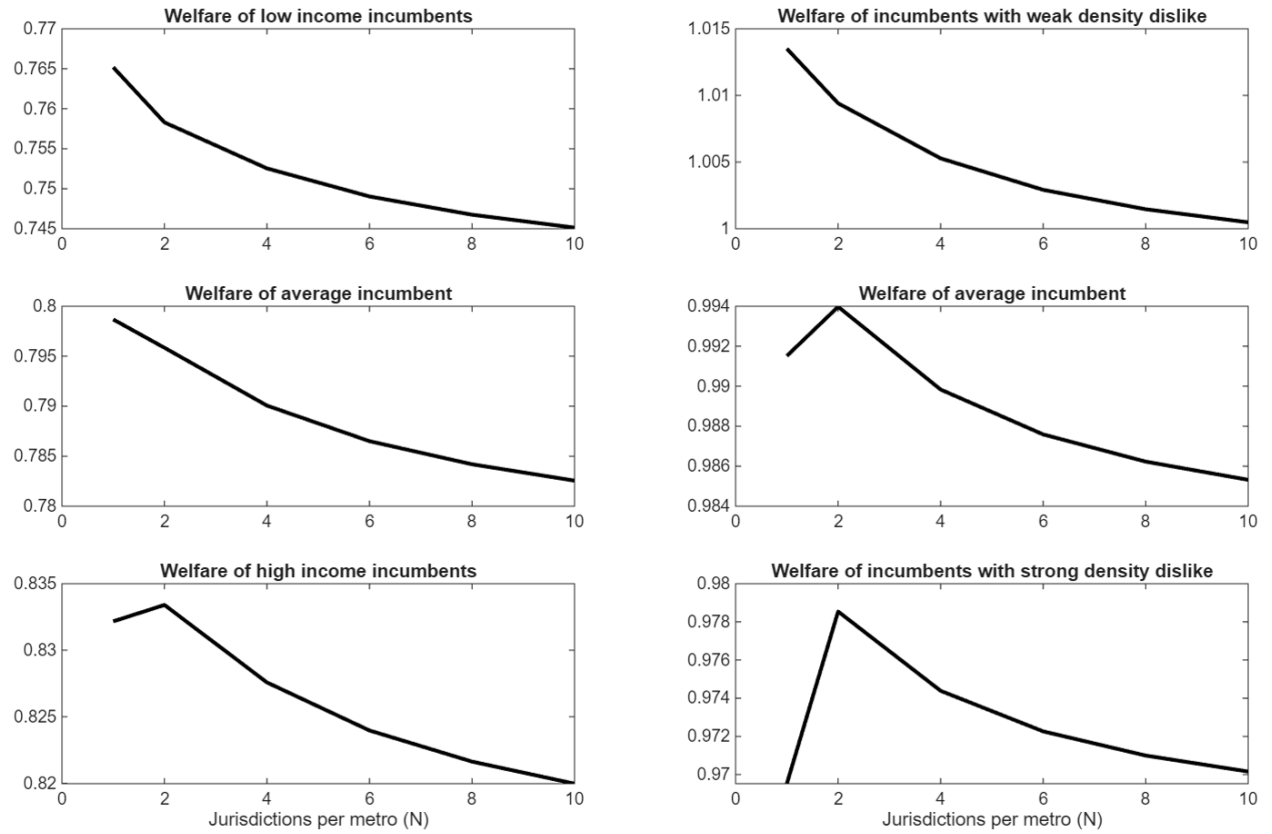


Figure A.8 — Resistance to amalgamation

This figure presents results from the *Migration across metros* model with public goods, taxes, and income heterogeneity in the left three panels and from the *Heterogeneity in preference for density* model in the right three panels. The x-axis represents the number of municipalities N . The y-axis plots the utility of incumbent land owners with low income (left) or low dislike of density (right) in the top two panels, average incumbents in the middle two panels, and incumbents with high income (left) or high dislike of density (right) in the bottom two panels.



2 Model extensions

2.1 Migration within metro model

This model closely follows the *Migration across metros* framework presented in the main paper. To avoid redundancy, we summarize the key differences here rather than presenting the full model. Unlike the *Migration across metros* model, this version assumes no migration from other metros, meaning that the wage is constant and normalized to one. The congestion externality function is still local but now depends explicitly on the zoning rule $h_{0,i}$, $\phi(h_{0,i}) = \phi_0 h_{0,i}^{\phi_1}$, where stricter zoning (small $h_{0,i}$) leads to lower congestion. Additionally, we set the construction cost $\lambda = 0$. The results remain qualitatively similar with positive costs, but the model works even with a zero cost.¹

A key distinction in this model is that between $t = 0$ and $t = 1$, landowners remain in their initial municipality with probability $1 - q$ but receive a random moving shock with probability q . Conditional on receiving a moving shock, they have equal chances of moving to any municipality, including their own. Therefore, the probability of moving to any different municipality is $\hat{q} = q \frac{N-1}{N}$, and the probability of moving from municipality i to municipality k is $q_{ik} = \frac{\hat{q}}{N-1}$ if $k \neq i$. A household that moves from municipality i to municipality k at $t = 1$ solves:

$$\begin{aligned} u_{1,j}(p_i h_{0,i}) &= \max_{c_{ik}, h_{ik}} c_{ik}^{\alpha_c} h_{ik}^{\alpha_h} \\ \text{s.t. } p_k h_{ik} + c_{ik} &= p_i h_{0,i} + w(1 - \phi(h_{0,k})) \end{aligned} \quad (\text{A.1})$$

This equation mirrors the time-1 landowner's problem in the main paper, but we emphasize that the household's net worth depends on the value of housing in their initial municipality i , while its housing expenditure is determined by the price in municipality k . This problem's analytical solution is identical to the analytical solution in the main paper, and the ex-ante expected utility of a household who owns in neighborhood i is:

$$\begin{aligned} u_{0,i} &= (1 - \hat{q})u_{1,ii} + \sum_{i \neq k} q_{ik} u_{1,ik} \\ &= \bar{\alpha} \left((1 - \hat{q}) p_i^{-\alpha_h} (h_{0,i} p_i + (1 - \phi_i)w) + \hat{q} p_k^{-\alpha_h} (h_{0,i} p_i + (1 - \phi_k)w) \right) \end{aligned} \quad (\text{A.2})$$

As in the *Migration across metros* model in the main paper, we search for a symmetric Nash equilibrium to solve this problem, which leads to six equilibrium conditions with six unknowns ($h_{0,i}$,

1. Because of the congestion externality, there is an interior solution for h_0 even when $\lambda = 0$. The *Migration across metros* model in the main paper has a different specification for the congestion externality. As a result, it does require a construction cost to get an interior solution for h if the utility is Cobb-Douglas because, without construction costs, optimal structures h are infinite. However, construction costs are not important for the mechanism. For example, if the utility function is CES with stronger complementarity than Cobb-Douglas, then the model has an interior solution for h even with zero construction costs. In this case, our key results on the relationship between the number of voting municipalities m and the restrictiveness of zoning h are the same as in the *Migration across metros* model.

$h_{0,k}, p_i, p_k, \phi_i, \phi_k$:

$$\begin{aligned}
h_{0,j} &= \frac{\alpha_h}{p_j} ((1 - \hat{q})(h_{0,j}p_j + (1 - \phi_j)w) + \hat{q}(h_{0,k}p_k + (1 - \phi_i)w)) & j \in (i, k) \\
\phi_j &= \phi_0 h_{0,j}^{\phi_1} & j \in (i, k) \\
h_{0,i} &= \bar{\alpha} \left((1 - \hat{q})p_i^{-\alpha_h}(h_{0,i}p_i + (1 - \phi_i)w) + \hat{q}p_k^{-\alpha_h}(h_{0,i}p_i + (1 - \phi_k)w) \right) \\
h_{0,i} &= h_{0,k}
\end{aligned} \tag{A.3}$$

The first line equates the aggregate supply of housing in municipality i or k with the aggregate demand, which includes demand from current residents as well as residents moving from other neighborhoods. Since this problem is solved from the point of view of residents of municipality i , taking as given the zoning of all other municipalities, they internalize the fact that the demand in every other municipality will be affected by q_{ik} residents from municipality i moving there. The second line defines the congestion externality. The third line represents the Nash equilibrium condition, ensuring that, conditional on their beliefs about all other endogenous variables, the owners in municipality i choose housing supply to maximize their expected utility from Equation A.2. Finally, the fourth equation specifies that the equilibrium is symmetric.

These conditions can be reduced to two equations and two unknowns. It is also easy to check that if there is a single municipality ($N = 1$) or if there is no moving ($q = 0$) then the solution is equivalent to one chosen by a social planner.

2.2 Heterogeneity in municipal size

We now revisit the baseline *Migration across metros* model presented in the main paper but allow municipalities to have different sizes. Specifically, we solve the model for metros with $N \in [1, 10]$ municipalities. For each $N > 1$, we randomly assign each municipality a relative size within the metro and solve the model for several different random realizations. This exercise demonstrates that HHI is a more relevant measure of fragmentation than the simple inverse of the number of municipalities ($1/N$). The top panel of Figure A.9 shows the relationship between HHI and its average housing size (the inverse of zoning restrictiveness) where each dot represents a metro. The relationship is nearly linear. In contrast, the bottom panel of Figure A.9 shows the relationship between a metro's $1/N$ and its average housing size; this relationship is much noisier.

2.3 Calibration

The model is static and we think of it as reflecting long-term zoning, housing, prices, and migration. In all models, we set $\alpha_h = 0.2$ as this is approximately the housing expenditure share in the BLS; we assume that congestion is quadratic in population, implying $\phi_1 = 2.0$; we set the severity of congestion cost (strength of anti-density preferences) $\phi_0 = 0.025$ to roughly match the slope of MLS with respect to municipal HHI; we normalize the wage level parameter $\omega_0 = 1$ and set the wage agglomeration parameter to $\omega_1 = 0.13$ following estimates in Glaeser and Gottlieb (2009).

The remaining parameters differ across models. In the *Migration across metros* model we set the migrant’s outside option $\underline{u}_m = 0.835$ so that when $N = 2$, migrants are 50% of the population. We choose this target because $N = 2$ corresponds to the median HHI in our sample and because [Dávila and Zlate \(2022\)](#) estimate that up to 40% of U.S.-born residents live outside of their state of birth. Moreover, the foreign-born population is around 15%. For the *Migration within metro* model, we set $\underline{u}_m = \infty$ as there are no migrants from outside the metro, but set the probability of moving within the metro $q = 0.5$. For the *Heterogeneity in preference for density*, *Heterogeneity in income among incumbents*, and *Heterogeneity in income between immigrants and incumbents* models we continue to target a 50% migrant ratio when $N = 2$, leading to $\underline{u}_m$ taking the values, respectively, of 0.835, 0.710, and 0.574.

For the *Heterogeneity in preference for density* model, the strength of anti-density preferences ϕ_0 are 0.02 and 0.03 for the high and low types.

For the *Heterogeneity in income among incumbents* model, we set $\theta = 0.05$, implying that wage is 5% above (below) average for the high (low) type. Additionally, we set the public good share in the utility function to $\gamma = 0.04$. We choose this value because [Barro and Sala-I-Martin \(1992\)](#) show that when chosen optimally, public goods expenditure share should be equal to γ . While exact expenditure share on municipality-provided public goods is difficult to estimate, calculations by the [Tax Policy Center \(2021\)](#) suggest that 4% is reasonable.²

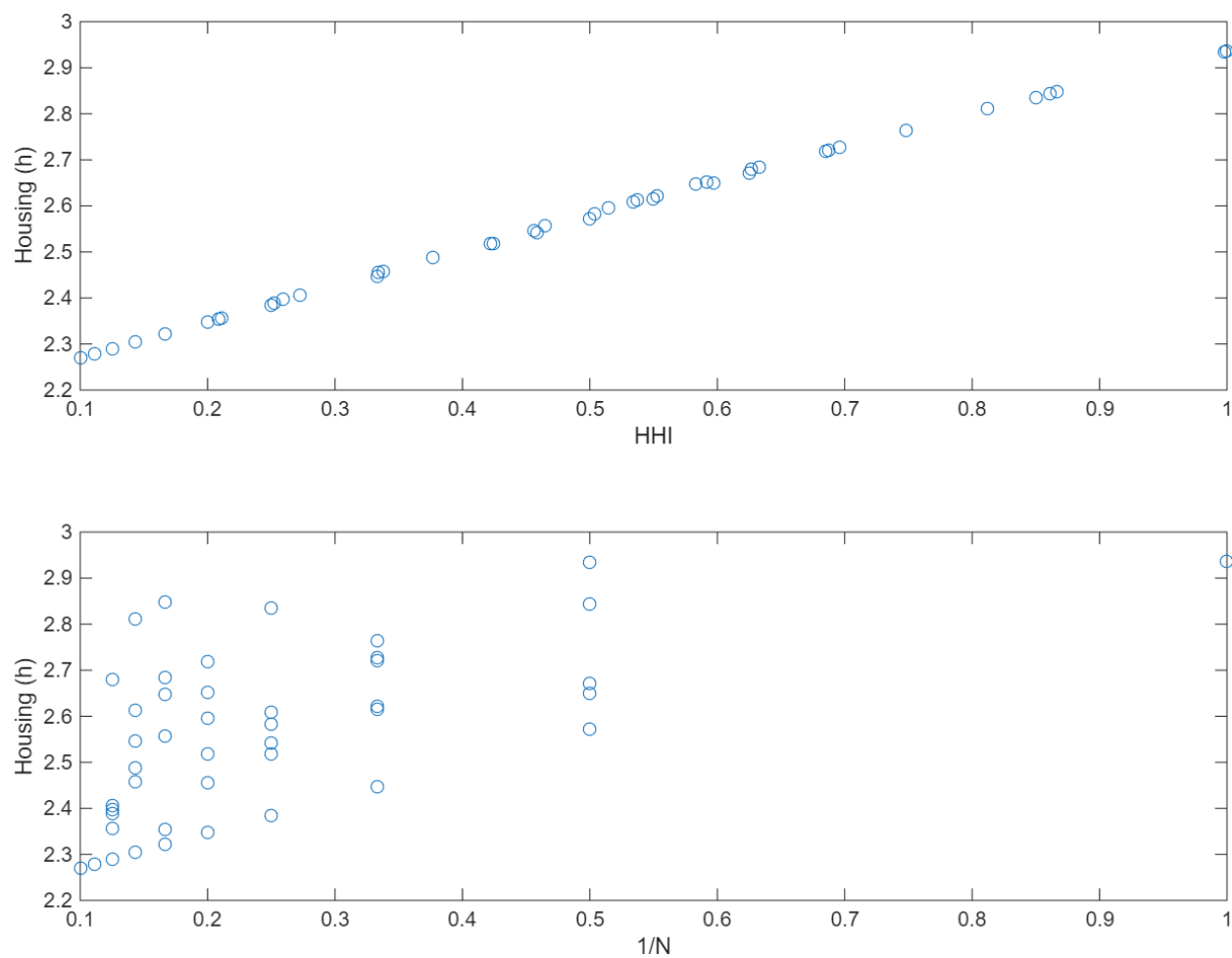
For the *Heterogeneity in income between immigrants and incumbents* model, we set the ratio of immigrant to local income $\psi = 0.7$ to approximately match the slope of average wage with respect to municipal HHI.

Finally, we must choose the construction cost parameter λ . This parameter affects the level of housing stock and the level of house prices. However, we found that numerically, the level of congestion, the migrant share, and the wage are invariant to this parameter; additionally, the slope of the logarithm of price, quantity, and utility with respect to the logarithm of N are also invariant to this parameter. We set $\lambda = 0.1$ and confirm that our results are consistent for alternative values.

2. For example, in 2021 total municipal property plus income taxes were 2.8% of GDP, total municipal taxes were 3.7% of GDP, and total municipal revenues (excluding transfers from the state) were 5.7% of GDP.

Figure A.9 — Model results: Migration across metros with heterogeneity in municipal size

This figure presents results from the *Migration across metros* model with heterogeneity in metro size. We vary the number of municipalities $N = 1, 2, \dots, 10$. For each number of municipalities, we solve the model several times, each time randomly choosing each municipality's proportion of total metro population. Each point in the scatter-plot represents the model's solution for a metro with a specific N and municipality size distribution. The top panel presents the relationship between a metro's HHI and housing quantity (equivalently, the inverse of zoning restrictions); the bottom panel presents the relationship between a metro's $1/N$ and housing quantity.



3 Empirical Analysis and Robustness Check Details

3.1 Construction of the Instrument

The instrument is constructed using the same methodology as the current HHI described in the main paper, but it is restricted to jurisdictions that existed in 1875. To identify municipalities that existed in 1875, we compile their year of establishment, defined as the year of incorporation for incorporated places and the year of creation for minor civil divisions and counties, using multiple data sources.

First, we rely on the Municipal Incorporation Data compiled by [Goodman \(2023\)](#), which provides near-complete coverage of incorporation years for incorporated places across most states, except Nebraska (NE), Oklahoma (OK), South Dakota (SD), and Utah (UT). To supplement this dataset, we use the 1850 1%, 1860 1%, and 1870 1% IPUMS samples to determine whether incorporated places appeared in the Census. The 1% samples include the INCORP variable, which identifies the incorporated place where a household resides. We match state and municipality names with the IPUMS dictionary for the INCORP variable to verify whether a given municipality existed before 1875.³

To identify minor civil divisions (MCDs) that existed before 1875, we use the MCD variable in the IPUMS Census data, which records the minor civil division where a household resides. Since the IPUMS data is based on sample surveys, we further digitize historical maps drawn before 1875 to capture the names of minor civil divisions. We classify an MCD as having been established before 1875 if it appears in either the IPUMS Census data or historical maps.

Through this process, we find that 56% of municipalities in our sample existed as of 1875. In comparison, we find that 29% of municipalities existed as of 1850 and 72% as of 1900. Using this subset of municipalities, we calculate HHI as of 1875 to measure the historical concentration of local governments and use it as the baseline instrument. Note that for CBSAs with no municipalities identified to have existed as of 1875, we assign an HHI value of 10,000, equivalent to a fully consolidated jurisdiction.

3.2 Sample Summary Statistics

Table [A.2](#) presents summary statistics for key variables in the sample of 906 CBSAs. At the CBSA level, the average of municipality-level median MLS has a mean of 39,570 square feet (0.91 acres) and a standard deviation of 30,481 square feet (0.70 acres). The 10th percentile CBSA has an MLS of 12,416 square feet (0.29 acres), while the median (50th percentile) is 30,544 square feet (0.70 acres), and the 90th percentile reaches 78,233 square feet (1.80 acres).

Alternative MLS aggregation methods also show substantial variation. The exponential of the average log MLS, which places less weight on large minimum lot sizes, has a mean of 26,539 square feet (0.61 acres). The first quartile (Q1) of municipality-level median MLS, representing the less

3. The INCORP variable is not available in the full-count Census, meaning that relying solely on IPUMS data to determine incorporation status before 1875 may result in false negatives. Similarly, the MCD variable described below is also absent from the full-count Census and may suffer from the same limitation.

restrictive end of zoning regulations within a CBSA, has a mean of 16,116 square feet (0.37 acres). In contrast, the third quartile (Q3), reflecting the more restrictive end, has a mean of 48,919 square feet (1.12 acres). The Wharton Density Restriction Index (DRI), which reflects the largest MLS regulation within CBSAs, is available for 574 CBSAs and exhibits notable variation as well, with a mean of 1.78 and a standard deviation of 1.09.⁴

The zoning jurisdiction HHI exhibits substantial variation, with a mean of 4,639 and a standard deviation of 3,298, ranging from 523 at the 10th percentile to 9,189 at the 90th percentile.⁵ The number of municipalities averages 23.3, with values ranging from 4 at the 10th percentile to 50 at the 90th percentile. In 1875, the average number of municipalities was 13.15. Consequently, HHIs as of 1875 are higher than the current HHI, with a mean of 5,583 and a median of 5,654, reflecting the formation of new municipalities over time and an overall trend of increasing fragmentation, though significant variation persists. Lastly, the number of rivers crossing CBSAs has a mean of 34.7 and a standard deviation of 37.4.

Figure A.3 shows the correlation between the HHI and several key control and outcome variables used in the main empirical analysis. Of note is that less fragmented places tend to be warmer (0.4 correlation of HHI with average January temperature) and lower earning (−0.29 correlation of HHI with earnings). Less fragmented places also tend to be smaller, though the correlation is only −0.16 and there is significant variation in HHI within each size quartile.

3.2.1 Identifying Historical Municipal Mergers and Annexations

To identify historical municipal mergers and annexations cases, we begin with the IPUMS dictionary of incorporated places and minor civil divisions used in the 1850–1930 samples. We restrict the list to incorporated places in states where incorporated places typically function as local governments and to minor civil divisions in states where MCDs typically function as local governments. We then construct name keys by dropping common suffixes and prefixes, such as “town,” “township,” and “city of,” and retain unique name keys within each state, yielding 26,300 potential municipalities observed between 1850 and 1930. We merge this list to the 2019 Census place and county-subdivision files using state-by-name-key matches and identify 4,201 non-matches. Any amalgamations must be a subset of these 4,201 non-matches because the matched names still exist as independent Census places in 2019. We manually review these non-match cases and classify 1,651 as not municipalities (e.g., unidentifiable entries such as “91 r43” or “t10 r8”), 238 as neighborhoods that were never municipalities, 177 as extinct, 150 as merged or annexed, and 181 as unknown; the remainder still exist but our algorithm failed in matching them to the 2019 files due to spelling differences. Excluding non-municipal

4. The Wharton DRI takes a value of 0 if there is no MLS regulation, 1 if the largest MLS is no larger than 0.5 acres, 2 if it falls within 0.5–1.0 acres, 3 if it falls within 1.0–2.0 acres, and 4 if it exceeds 2.0 acres.

5. Figure A.3 shows the correlation between the HHI and key control and outcome variables used in the main empirical analysis. For example, one potential concern is that larger metros may be more fragmented (i.e., have lower HHI), which could confound the results. However, the correlation between log HHI and log residential land area is modest at −0.155. Importantly, our main IV regressions include log residential land area as a control variable, and we further confirm robustness across CBSA size quartiles in Table A.9. We conduct similar robustness checks for other potential confounders as well.

entries, neighborhoods that were never municipalities, and unknown cases, mergers/annexations account for only 0.6% of municipalities in existence before 1930 (150 of 24,230), and these 0.6% occurred over the course of more than a century (the denominator includes municipalities in existence before 1930, but some of them merged well before 1930). Even if all unknown cases are classified as mergers/annexations, mergers/annexations account for only 1.4% of municipalities. Of the 112 amalgamations for which we could identify dates, 25% occurred 1850–1900, 38% 1901–1950, 30% 1951–2000, and 8% post-2000.

3.3 Robustness Checks

3.3.1 Alternative instruments: HHI as of 1900/1850 and river instruments

We evaluate the robustness of our findings by using alternative instruments. First, we construct historical HHIs as instruments for different years, rather than the baseline year of 1875. Specifically, we compute HHIs as of 1900 and 1850 to test the consistency of our results over different historical periods. Second, we use the number of rivers crossing each CBSA as an alternative instrument, following the approach of [Hoxby \(2000\)](#). Using the USA Rivers and Streams database, we count the total number of streams crossing CBSAs.⁶

Table [A.5](#) presents the results using alternative instruments to assess the robustness of our findings. Columns (1) and (2) report the IV estimates using historical HHIs as instruments, while Column (3) reports results using the river-based instrument. The first-stage regressions corresponding to each specification are presented in Columns (4) through (6).

The historical HHI instruments, based on 1900 and 1850 data, exhibit strong first-stage relationships, with F-statistics of 3,500.6 and 167.2 in Columns (4) and (5), respectively. The IV estimates in Columns (1) and (2) remain negative and statistically significant, with p-values less than 0.0001, aligning with our baseline results. Column (3) explores an alternative instrument based on the number of rivers crossing each CBSA. While the IV estimates continue to indicate a negative and statistically significant relationship at the 5% level, the first-stage regression in Column (6) shows weaker instrument strength, with F-statistics of 5.80 (p-value = 0.017). This suggests that while the river-based instrument captures some variation in municipality fragmentation, specifically, a greater number of rivers leads to greater fragmentation, they are less predictive than the historical HHI-based instruments. Accordingly, the IV coefficient in Column (3) is noisier, yielding a coefficient estimate of -1.036 with a standard error of 0.5040. Nonetheless, the consistently negative estimates across all specifications reinforce our main hypothesis that more fragmented zoning authority leads

6. We exclude artificial paths, intermittent streams, canals, aqueducts, and intracoastal waterways from the dataset to ensure that only natural, continuously flowing rivers and streams are considered in our instrument construction. Additionally, we count the number of larger streams extending beyond 3.5 miles, consistent with the definition of large streams in [Hoxby \(2000\)](#), and those extending beyond 5.75 miles, which represents the median stream length in the dataset, and include them as additional instruments as an alternative specification. The numbers of large streams do not exhibit statistical significance, with p-values of 0.226 and 0.612, when included alongside the total number of streams in the first stage. Moreover, the main IV coefficient is statistically indistinguishable compared to using the total number of streams alone as an instrument, suggesting that the additional distinction between large and small streams does not provide additional explanatory power for municipal fragmentation in this context.

to stricter zoning regulations.

3.3.2 Geographic Controls and Exclusions

Table A.6 presents CBSA-level IV regressions that examine whether the estimated effect varies across broad regions of the United States. Column (1) augments the baseline specification with Census-region fixed effects (Midwest, Northeast, South, and West). Column (5) then allows for regional heterogeneity by interacting both the treatment variable (log HHI) and the instrument (log HHI as of 1875) with Census-region indicators.⁷ The estimated effects are broadly similar across regions. The point estimate is largest and most precisely estimated in the Midwest (coefficient estimate of -0.2729 , p -value < 0.0001) and smallest in the West (coefficient estimate of -0.1564 , p -value $= 0.2340$), with the latter also having the largest standard error.⁸ The coefficient differences across regions are not statistically significant at conventional levels, and the estimates remain consistently negative across all regions. Overall, these results reinforce the robustness of the main finding that more fragmented metropolitan governance (lower HHI) is associated with more restrictive zoning.

Columns (2), (3), and (6) provide additional robustness checks based on alternative region definitions. Column (2) replaces Census-region fixed effects with fixed effects defined by territorial acquisition categories, while Column (3) replaces them with fixed effects defined by the decade of state admission. Column (6) goes one step further by allowing the treatment and instrument to vary flexibly across territorial-acquisition groups through interactions. Across these alternatives, the estimated effect remains negative and similar in magnitude, indicating that the main results are not driven by broad regional differences tied to territorial history or to earlier-versus-later incorporation eras.

Lastly, Column (4) excludes the top destination CBSAs during the Second Great Migration. We define destinations using the number of Black in-migrants during 1935–1940 from [Derenoncourt \(2022\)](#)’s county-level migration data, and exclude Southern CBSAs when constructing the ranking, so that the exercise focuses on major receiving metros.⁹ This subsample analysis addresses the possibility that large migration shocks affected municipal fragmentation and zoning stringency in ways correlated with our historical instrument ([Boustan, 2010](#); [Sahn, 2021](#); [Cui, 2024](#)). The estimated effect remains statistically indistinguishable from the baseline.

7. Eleven CBSAs span multiple Census regions (e.g., Cincinnati, OH–KY–IN falls within both the Midwest and South). For these CBSAs, the fixed-effect category in Column (1) is defined by the combination of regions (e.g., “MW/S”), while in Column (5) the region indicators for all constituent regions are set to one.

8. The first-stage F-statistic for the West ($N = 168$) is 213.0, compared with 298.0 for the Midwest ($N = 246$), 7,184 for the Northeast ($N = 92$), and 248.1 for the South ($N = 358$).

9. Column (4) excludes the top 50 destination CBSAs, which corresponds to 49 CBSAs in the baseline analysis sample. We also consider alternative cutoffs excluding the top 10 and top 100 destinations; the estimates remain similar, with coefficient estimates of -0.275 (standard error $= 0.040$) and -0.246 (standard error $= 0.045$), respectively.

3.3.3 Alternative measurements

We assess the robustness of our main findings by exploring alternative measurements for both our outcome and treatment variables and present the results in Table A.7. First, since the baseline MLS estimates from Song (2025) are missing in some CBSAs due to gaps in the construction year data within CoreLogic, we conduct an alternative specification to verify the validity of our results. Panel A uses an alternative MLS estimate from Song (2025), which infers neighborhood-level minimum lot sizes based on all single-family construction, rather than restricting the analysis to post-1940 construction, thereby achieving complete coverage.

Second, Panels B to D retain the baseline MLS estimates but apply different aggregation schemes at the CBSA level. Panel B takes the average of log MLS, which downweights large MLS estimates, while Panels C and D capture different portions of the MLS distribution within municipalities. Panel C reports the 1st quartile (Q1) of municipality-level median MLS, representing less restrictive zoning environments, while Panel D reports the 3rd quartile (Q3) of municipality-level median MLS, reflecting more restrictive zoning municipalities within CBSAs. Beyond MLS-based zoning stringency measures, Panel E uses the Density Restriction Index (DRI) from the 2018 Wharton Land Use Survey, a widely recognized measure of zoning restrictiveness, as the outcome variable. The survey covers 2,450 municipalities and categorizes each jurisdiction based on its most restrictive minimum lot size requirement. Specifically, the DRI distinguishes whether a municipality imposes an MLS and, if so, whether the largest MLS falls into one of four categories: less than 0.5 acres, 0.5 to 1 acre, 1 to 2 acres, or more than 2 acres. We aggregate this measure at the CBSA level by computing a weighted average of the DRI across municipalities.

Finally, Panel F replaces the zoning jurisdiction HHI with the number of municipalities as an alternative measure of zoning authority granularity. In this case, we instrument the number of municipalities using historical municipality counts from 1900 to address potential endogeneity concerns.

Across all specifications, we find a consistent and statistically significant relationship between municipal fragmentation and zoning restrictiveness, reinforcing the central prediction of our theoretical model: greater fragmentation of zoning authority, reflected in lower HHI or a higher number of municipalities, leads to more stringent zoning regulations, as measured by higher MLS or DRI. In Panel E, the IV estimate in Column (4) is smaller in magnitude and less precisely estimated, primarily due to the smaller sample size and the coarser nature of the DRI measure. Similarly, the results in Panel F are weaker when using the number of municipalities as the treatment variable. This attenuation is expected: as shown by the extended model and by Figure A.9, the relationship between zoning stringency and the number of municipalities (N) becomes noisier when municipality sizes vary. In contrast, our baseline measure of granularity, the HHI, more effectively captures fragmentation by accounting for size heterogeneity among jurisdictions. Nonetheless, the consistency of our findings, regardless of MLS estimation method, aggregation strategy, or alternative zoning restrictiveness measures, provides strong support for the robustness of our empirical results.

3.3.4 Additional Robustness Checks

We report several additional robustness checks. First, given the baseline specification includes 42 control variables for a sample of 909 CBSAs, we implement a LASSO-based variable selection following Chernozhukov et al. (2015). The resulting IV estimate remains stable at -0.1981 (standard error = 0.0313). Second, Table A.8 reports estimates from a flexible specification that allows the effect of fragmentation to vary across quartiles of the HHI distribution. We find consistently negative effects across all quartiles.^{10,11} Third, we allow for heterogeneity by CBSA population and land area quartiles (Table A.9) and again find uniformly negative coefficients. Fourth, as discussed in the heterogeneity model extensions, within-CBSA heterogeneity may affect fragmentation. Empirically, Figure A.10 shows that greater income dispersion and a higher share of white population are associated with more fragmentation (left panel). However, these correlations largely attenuate once we condition on observables (right panel). More importantly, Table A.10 demonstrates that our main IV estimates remain negative and statistically significant across all quartiles of CBSA-level heterogeneity. Fifth, in the baseline specification, when no municipalities are observed in a CBSA as of 1875, we code the 1875 HHI as 10,000 (fully centralized). As an alternative robustness check, we drop these CBSAs and re-estimate the IV regressions. The results remain consistent with the baseline (coefficient = -0.266 , $t = -7.3$).

10. The estimated coefficients are larger in magnitude than the baseline, though standard errors are also higher due to reliance on within-quartile variation. Nonetheless, the results remain robust.

11. Figure A.4 suggests that the marginal effect of HHI on zoning should be lower when HHI is small (the number of municipalities N is large), which would suggest that we should observe almost no effect for very fragmented metros. Table A.8 finds no statistical difference in the size of the effect across HHI quartiles. We conjecture that this is because most of the variation in HHI occurs in a region where the effect is strong throughout. Table A.2 shows that instrumented HHI is 1217 at the 25th percentile and 10,000 at the 75th percentile; this approximately corresponds to 8 and 1 equally sized municipalities. Figure A.4 shows large effects of HHI on zoning in that range.

Table A.5 — Robustness to using alternative instruments

This table presents IV regression results using alternative instrumental variables. Columns (1) to (3) report the main coefficient estimates, while Columns (4) to (6) present the corresponding first-stage regression results. Columns (1) and (4) use the HHI as of 1900 as an instrument, Columns (2) and (5) use the HHI as of 1850 as an instrument, and Columns (3) and (6) use the number of rivers crossing the CBSA as an instrument. All specifications include the full set of control variables described in Table A.1. Due to space constraints, the table omits the coefficients for the control variables.

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: log MLS			First stage		
$\log \widehat{HHI}$	-0.1973*** (0.0334)	-0.4069*** (0.0747)	-1.036** (0.5040)			
log HHI as of 1900				0.9891*** (0.0167)		
log HHI as of 1850					0.4337*** (0.0335)	
# rivers						-0.0022** (0.0009)
Observations	832	832	832	832	832	832
Adjusted R ²	0.35098	0.30070	-0.33121	0.93552	0.70909	0.64947

Significance: ***: 0.01, **: 0.05, *: 0.1

Table A.6 — Estimated effects of HHI on zoning stringency by region

This table presents IV regression results, where the dependent variable is zoning stringency, measured as the log of the minimum lot size aggregated at the CBSA level. Column (1) incorporates Census-region fixed effects into the baseline IV regression. Columns (2) and (3) replace Census-region fixed effects with fixed effects defined by territorial acquisition categories and by the decade of state admission, respectively. Column (4) excludes the top 50 destination CBSAs that experienced the highest Black in-migration during 1935–1940 to test the robustness of the results to historical migration patterns. Columns (5) and (6) extend the baseline specification by interacting both the treatment variable ($\log \widehat{HHI}$) and the instrument ($\log \widehat{HHI}$ as of 1875) with Census-region indicators and territorial-acquisition indicators, respectively, allowing for heterogeneity in the estimated effects across groups. All columns control for the full set of variables described in Table A.1. Due to space constraints, the table omits the coefficients for the control variables.

	(1) Census FE	(2) Territorial FE	(3) Adm.-decade FE	(4) Excl. GM top 50	(5) Census: x/IV int.	(6) Territorial: x/IV int.
$\log \widehat{HHI}$	-0.2771*** (0.0462)	-0.2742*** (0.0407)	-0.2631*** (0.0449)	-0.2591*** (0.0420)		
$\log \widehat{HHI} \times \mathbf{1}(\text{Midwest})$					-0.2729*** (0.0497)	
$\log \widehat{HHI} \times \mathbf{1}(\text{Northeast})$					-0.2649*** (0.0903)	
$\log \widehat{HHI} \times \mathbf{1}(\text{South})$					-0.2463** (0.0954)	
$\log \widehat{HHI} \times \mathbf{1}(\text{West})$					-0.1564 (0.1401)	
$\log \widehat{HHI} \times \mathbf{1}(\text{Internal formation})$						-0.2300*** (0.0866)
$\log \widehat{HHI} \times \mathbf{1}(\text{Louisiana Purchase})$						-0.4529*** (0.0871)
$\log \widehat{HHI} \times \mathbf{1}(\text{Mexican Cession})$						-0.3625** (0.1518)
$\log \widehat{HHI} \times \mathbf{1}(\text{Multiple acquisitions})$						-0.2786*** (0.0657)
$\log \widehat{HHI} \times \mathbf{1}(\text{Northwest Territory})$						-0.1314** (0.0615)
$\log \widehat{HHI} \times \mathbf{1}(\text{Oregon Country})$						-0.3126 (0.3507)
$\log \widehat{HHI} \times \mathbf{1}(\text{Original 13 colonies})$						-0.3068*** (0.0542)
$\log \widehat{HHI} \times \mathbf{1}(\text{Original U.S. territory post-1783})$						-0.3197 (0.2339)
$\log \widehat{HHI} \times \mathbf{1}(\text{Spanish Cession: Florida})$						-0.5417 (0.3346)
$\log \widehat{HHI} \times \mathbf{1}(\text{Texas Annexation})$						-0.4080** (0.1624)
Observations	830	832	832	783	830	832
Adjusted R ²	0.41402	0.42770	0.44003	0.38838	0.41382	0.43197
Within R ²	0.30321	0.28515	0.37380		0.30298	0.29049

Significance: ***, 0.01, **, 0.05, *, 0.1

Table A.7 — Robustness to using alternative measurements

This table presents robustness checks to alternative measurements. Panel A uses alternative MLS estimates from Song (2025) using all single-family construction with complete coverage. Panels B to D aggregate the baseline MLS estimates to the CBSA level using different methods: Panel B uses the average of the log MLS, Panel C uses the 1st quartile of municipality-level medians, and Panel D uses the 3rd quartile of municipality-level medians. Panel E substitutes the zoning stringency measure with the Density Restriction Index (DRI) from the 2018 Wharton Land Use Survey. Panel F maintains the same dependent variable as the baseline regression but replaces HHI with the number of municipalities as the independent variable, instrumented by the number of municipalities as of 1875. For each panel, Columns (1) and (2) present OLS regressions without and with controls, while Columns (3) and (4) report IV regressions without and with controls. The control variables included in Columns (2) and (4) are detailed in Table A.1. Due to space constraints, the table omits the coefficients of the control variables.

	A. Outcome: MLS inferred from all construction				B. Outcome: Mean of log			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	OLS		IV		OLS		IV	
log HHI	-0.2736*** (0.0183)	-0.2584*** (0.0264)	-0.2988*** (0.0203)	-0.2930*** (0.0340)	-0.1176*** (0.0158)	-0.1034*** (0.0214)	-0.1623*** (0.0176)	-0.1693*** (0.0279)
Obs.	906	906	906	906	832	832	832	832
Adj. R^2	0.19657	0.43224	0.19489	0.43111	0.06112	0.42926	0.05210	0.42234
	C. Outcome: Q1 of medians				D. Outcome: Q3 of medians			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	OLS		IV		OLS		IV	
log HHI	-0.0597*** (0.0166)	-0.1088*** (0.0233)	-0.0936*** (0.0184)	-0.1500*** (0.0303)	-0.2364*** (0.0240)	-0.2024*** (0.0364)	-0.3041*** (0.0267)	-0.3306*** (0.0474)
Obs.	832	832	832	832	832	832	832	832
Adj. R^2	0.01414	0.35393	0.00919	0.35136	0.10331	0.31849	0.09475	0.30765
	E. Outcome: DRI in Wharton Survey				F. Treatment: # of municipalities			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	OLS		IV		OLS		IV	
log HHI (N for Panel F)	-0.1564*** (0.0380)	-0.1319** (0.0625)	-0.1468*** (0.0417)	-0.1309* (0.0790)	0.0022*** (0.0006)	0.0003 (0.0007)	0.0037*** (0.0008)	0.0023* (0.0012)
Obs.	574	574	574	574	832	832	832	832
Adj. R^2	0.02706	0.11744	0.02695	0.11744	0.01557	0.32538	0.00850	0.31830
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Significance: ***, 0.01, **, 0.05, *, 0.1

Table A.8 — Flexible specification by HHI quartiles

This table presents regression results where the dependent variable is zoning stringency, measured as the log of the minimum lot size aggregated at the CBSA level. The treatment variable, $\log(\text{HHI})$, is interacted with indicators for HHI quartile groups ($\mathbb{1}(\text{HHI} \in Q_1)$ through $\mathbb{1}(\text{HHI} \in Q_4)$). These interaction terms are instrumented using $\log(\text{HHI})$ as of 1875, also interacted with the corresponding HHI quartile indicators. Column (1) reports a univariate regression, while Column (2) reports a multivariate regression that incorporates the full set of controls. Columns (3) and (4) report CBSA-level IV regression estimates using $\log \text{HHI}$ as of 1875 as an instrument, with Column (4) including the full set of controls. The control variables included in Columns (2) and (4) are detailed in Table A.1. Due to space constraints, the table omits the coefficients for the control variables.

	(1)	(2)	(3)	(4)
	OLS		IV	
$\log \text{HHI} \times \mathbb{1}(\text{HHI} \in Q_1)$	-0.1096 (0.0726)	-0.1247* (0.0691)	-0.6430*** (0.1317)	-0.5473*** (0.1267)
$\log \text{HHI} \times \mathbb{1}(\text{HHI} \in Q_2)$	-0.1489** (0.0589)	-0.1323** (0.0574)	-0.5810*** (0.1066)	-0.4810*** (0.1051)
$\log \text{HHI} \times \mathbb{1}(\text{HHI} \in Q_3)$	-0.1491*** (0.0524)	-0.1383*** (0.0522)	-0.5326*** (0.0948)	-0.4545*** (0.0955)
$\log \text{HHI} \times \mathbb{1}(\text{HHI} \in Q_4)$	-0.1454*** (0.0501)	-0.1409*** (0.0506)	-0.5120*** (0.0905)	-0.4469*** (0.0924)
Observations	832	832	832	832
Controls	No	Yes	No	Yes
Adjusted R ²	0.12438	0.34973	0.06693	0.31846

Significance: ***: 0.01, **: 0.05, *: 0.1

Table A.9 — IV Results by CBSA Size

This table presents regression results where the dependent variable is zoning stringency, measured as the log of the minimum lot size aggregated at the CBSA level. The treatment variable, $\log(\text{HHI})$, is interacted with indicators for CBSA size quartile groups ($\mathbb{1}(\text{CBSA} \in Q_1)$ through $\mathbb{1}(\text{CBSA} \in Q_4)$). These interaction terms are instrumented using $\log(\text{HHI})$ as of 1875, also interacted with the corresponding HHI quartile indicators. Panel A defines CBSA size by total CBSA population, while Panel B uses total residential land area. In each panel, Column (1) reports a univariate regression, while Column (2) reports a multivariate regression that incorporates the full set of controls. Columns (3) and (4) report CBSA-level IV regression estimates using $\log \text{HHI}$ as of 1875 as an instrument, with Column (4) including the full set of controls. The control variables included in Columns (2) and (4) are detailed in Table A.1. Due to space constraints, the table omits the coefficients for the control variables.

Panel A. By CBSA population

	OLS		IV	
	(1)	(2)	(3)	(4)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_1)$	-0.2490*** (0.0202)	-0.1620*** (0.0312)	-0.2928*** (0.0224)	-0.2556*** (0.0401)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_2)$	-0.2561*** (0.0201)	-0.1696*** (0.0303)	-0.2991*** (0.0222)	-0.2617*** (0.0393)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_3)$	-0.2722*** (0.0209)	-0.1710*** (0.0306)	-0.3183*** (0.0232)	-0.2655*** (0.0399)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_4)$	-0.3178*** (0.0224)	-0.1772*** (0.0320)	-0.3673*** (0.0248)	-0.2730*** (0.0413)
Observations	832	832	832	832
Controls	No	Yes	No	Yes
Adjusted R ²	0.19811	0.35021	0.19322	0.34217

Panel B. By CBSA land area

	OLS		IV	
	(1)	(2)	(3)	(4)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_1)$	-0.2128*** (0.0210)	-0.1599*** (0.0311)	-0.2653*** (0.0233)	-0.2480*** (0.0398)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_2)$	-0.2171*** (0.0207)	-0.1704*** (0.0303)	-0.2694*** (0.0229)	-0.2603*** (0.0392)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_3)$	-0.2225*** (0.0215)	-0.1766*** (0.0311)	-0.2775*** (0.0238)	-0.2706*** (0.0404)
$\log \text{HHI} \times \mathbb{1}(\text{CBSA} \in Q_4)$	-0.2177*** (0.0225)	-0.1871*** (0.0323)	-0.2753*** (0.0249)	-0.2822*** (0.0416)
Observations	832	832	832	832
Controls	No	Yes	No	Yes
Adjusted R ²	0.11745	0.35125	0.11015	0.34358

Significance: ***: 0.01, **: 0.05, *: 0.1

Table A.10—IV Results by CBSA Heterogeneity

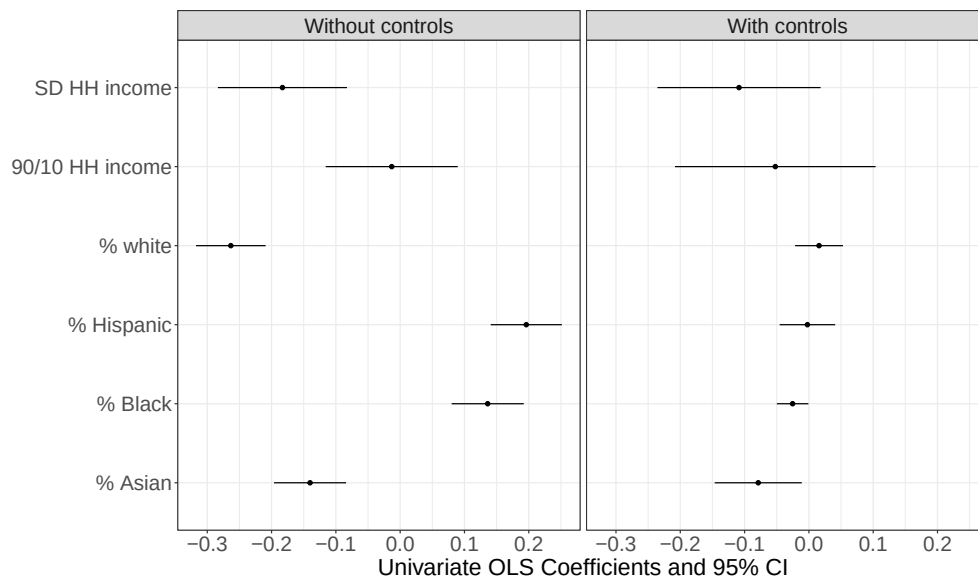
This table presents regression results where the dependent variable is zoning stringency, measured as the log of the minimum lot size aggregated at the CBSA level. The treatment variable, $\log(\text{HHI})$, is interacted with indicators for heterogeneity quartile groups ($\mathbf{1}(\text{CBSA} \in Q_1)$ through $\mathbf{1}(\text{CBSA} \in Q_4)$). Columns (1) and (2) examine income heterogeneity, measured by the standard deviation of household income and the 90–10 income ratio, respectively, constructed from the ACS public-use microdata. Columns (3) to (6) examine racial composition, using the share of the population that is white, Black, Asian, and Hispanic, respectively, constructed from the ACS 5-year estimates. All columns report IV estimates, including the full set of control variables, detailed in Table A.1. Due to space constraints, the table omits the coefficients for the control variables.

	(1)	(2)	(3)	(4)	(5)	(6)
	Income		Racial Composition			
$\log \text{HHI} \times \mathbf{1}(\text{CBSA} \in Q_1)$	-0.3266*** (0.0633)	-0.3406*** (0.0624)	-0.2605*** (0.0413)	-0.2792*** (0.0396)	-0.2626*** (0.0393)	-0.2805*** (0.0393)
$\log \text{HHI} \times \mathbf{1}(\text{CBSA} \in Q_2)$	-0.3322*** (0.0642)	-0.3274*** (0.0637)	-0.2602*** (0.0400)	-0.2645*** (0.0395)	-0.2606*** (0.0397)	-0.2604*** (0.0393)
$\log \text{HHI} \times \mathbf{1}(\text{CBSA} \in Q_3)$	-0.3268*** (0.0655)	-0.3177*** (0.0637)	-0.2521*** (0.0397)	-0.2621*** (0.0397)	-0.2617*** (0.0406)	-0.2659*** (0.0394)
$\log \text{HHI} \times \mathbf{1}(\text{CBSA} \in Q_4)$	-0.3321*** (0.0661)	-0.3034*** (0.0641)	-0.2681*** (0.0393)	-0.2399*** (0.0402)	-0.2639*** (0.0408)	-0.2867*** (0.0403)
Observations	261	261	832	832	832	832
Adjusted R ²	0.50928	0.52185	0.34528	0.34778	0.34149	0.35105

Significance: ***: 0.01, **: 0.05, *: 0.1

Figure A.10 — Correlation between CBSA heterogeneity measures and log HHI

The figure depicts the coefficient estimates and 95% confidence intervals from OLS regressions where the regressor is log HHI and the dependent variables (y-axis) are CBSA-level income heterogeneity and racial composition measures. The left panel shows the results from univariate regressions, while the right panel shows results from OLS regressions with the full set of controls, detailed in Table A.1. All coefficients are standardized by the standard deviation of the respective dependent variables.



4 Toronto Municipality Amalgamation Case Study

In this section, we present a case study of municipality amalgamation in Toronto, Canada. In Canada, provinces have the authority to create and dissolve municipalities. In January 1998, the Government of Ontario dissolved six previously incorporated municipalities—York, North York, East York, Scarborough, Etobicoke, and old Toronto—and merged them into a single municipality, the new City of Toronto, in what became known as the Amalgamation of Toronto. Notably, no changes were made to the other municipalities in the Toronto area. We exploit this municipality boundary change to provide additional empirical evidence for our theory.

According to Toronto’s chief administrative officer, the primary goal of the amalgamation was to reduce administrative and operating costs and increase governmental efficiency (Garrett, 1999). Projections included a 10% reduction of the workforce and annual savings of \$150 million. However, the move was unpopular—a non-binding 1997 referendum saw a large majority of voters opposing the merger. Beyond administrative restructuring, the amalgamation also affected urban planning and development: the city established the new Department of Urban Planning and Development Services and introduced a new Strategic Plan outlining a vision for the city’s future development.

Our theoretical model predicts that when multiple fragmented municipalities are amalgamated into a single entity, the new unified government optimally balances the costs and benefits of density, leading to less restrictive regulations for new development. While we cannot measure regulatory burdens directly in this setting, we find empirical evidence that supports our model prediction: following the amalgamation, housing starts grew faster in Toronto than in neighboring municipalities. Notably, this expansion in housing supply occurred without substantial price and rent responses, suggesting that supply regulations were eased post-amalgamation.

4.1 Empirical strategy: difference-in-difference

We apply a standard difference-in-difference (DiD) approach, using the six former municipalities that now comprise the city of Toronto as the treatment group and the five neighboring municipalities—Mississauga, Brampton, Vaughan, Markham, and Pickering—as the control group. These control municipalities share a border with Toronto but were not affected by the amalgamation. We compare pre- and post-1998 housing starts across the treatment and control groups. Since individual data for the six amalgamated municipalities is unavailable post-1998, we aggregate them into a single treated region and refer to it as “amalgamated Toronto” both before and after 1998. At the time of amalgamation, amalgamated Toronto had a population of approximately 2.4 million, compared to 1.3 million in the five neighboring municipalities. Figure A.11 presents a map of the region.

4.2 Empirical results

We begin by plotting the raw data.¹² The top panel of Figure A.12 illustrates housing starts in amalgamated Toronto, compared to the aggregated starts in the five control municipalities, with both series normalized to their 1997 levels. Prior to 1997, both series show a similar trend, suggesting similar growth trajectories. However, after 1997, housing starts in amalgamated Toronto accelerated significantly, diverging from the trend observed in the control municipalities. By 2010, housing starts in Toronto had more than doubled relative to the level they would have reached if the city followed the same trend as its neighbors.

To statistically test the impact of amalgamation on housing supply, we estimate a difference-in-differences (DiD) regression using the following specification:

$$\log Y_{i,t} = \alpha_t + \text{Treated}_i + \beta \text{Treated}_i \times \text{Post}_t + \gamma X_{i,t} + \epsilon_{i,t} \quad (\text{A.4})$$

where $Y_{i,t}$ represents the outcome variable for municipality i in year t , including housing starts, average rent, or average home price, each scaled by its 1997 level. The independent variables include time fixed effects (α_t), a treatment group dummy (Treated_i), which equals one for amalgamated Toronto, and a post-treatment dummy (Post_t), which equals one for time periods after amalgamation ($t \geq 1998$). We include control variables ($X_{i,t}$), which consist of each municipality’s lagged median wage and density levels as of 1989. Our coefficient of interest, β , measures the impact of amalgamation on the outcome variables.

In Panel A of Table A.11, we present the results of the DiD analysis by comparing amalgamated Toronto to the five neighboring municipalities, with Toronto being the single treated unit and the five municipalities as separate control units. In Panel B, we aggregate the five control municipalities into a single control unit as an alternative specification, comparing one treated unit (amalgamated Toronto) to one combined control. This alternative specification accounts for differences in the relative size of the control municipalities, which are implicitly treated equally in Panel A. The results remain largely consistent across both specifications.

Columns (1) and (2) of Table A.11 indicate that the number of housing permits issued in Toronto was 43% to 78% higher than it would have been had Toronto followed on the same trend as its neighbors in the 22-year period after amalgamation. This finding aligns with our model and the main empirical results in the main paper.¹³ These estimates are statistically significant, with t-statistics between 2.8 and 3.5. These estimates are somewhat larger in magnitude than those reported by Mast (2024) (20% change) and Tricaud (2025) (12.5% change).¹⁴

12. Data on housing starts, rents, and prices of new construction is from the Canada Mortgage and Housing Corporation (CMHC). Data on income, population, and density is from Statistics Canada. All data is publicly available.

13. These estimates are for housing starts and they are not directly comparable to the main paper’s baseline IV estimates, which imply that going from six (equal-sized) jurisdictions to one causes MLS to change by $e^{-0.194 \log(10000/6)} / e^{-0.194 \log(10000)} - 1 = 41.8\%$.

14. One possible explanation is that while Mast (2024) and Tricaud (2025) compare treated and control groups over an 8-year period, our analysis spans 23 years, during which the effect appears to grow over time. Another reason,

There are some limitations to this empirical approach. While both treated and control areas include a mix of urban and suburban neighborhoods, amalgamated Toronto tends to be more urban, denser, and lower earning, and characterized by smaller households and fewer children. These differences introduce the possibility of confounding factors. For example, if a demand shock around 1998 made urban living more attractive or if a productivity shock around 1998 increased earnings in urban areas, then the observed increase in housing development in amalgamated Toronto could be due to rising demand rather than supply-relaxation resulting from the Amalgamation of Toronto.

The standard approach in the literature to address such concerns is to use a discontinuity analysis that focuses on a narrow area along the border of treated and control regions, where housing characteristics and demographics are more comparable.¹⁵ Unfortunately, our data lacks the geographic granularity to implement a boundary discontinuity design. Instead, we examine the response of rents and prices following the amalgamation to rule out many potential confounders. If a demand shock around 1998 was the primary driver of increased housing starts in amalgamated Toronto, we would expect that rents and prices rise concurrently. Given reasonable estimates of supply elasticities, the magnitude of price increases should be comparable to the increase in housing supply.¹⁶ Furthermore, market prices can adjust more quickly than real estate investment decisions, so if demand shocks were responsible, we would expect rents and prices to rise before the increase in housing starts.

The bottom two panels in Figure A.12 provide no clear evidence of a substantial increase in housing prices and only a marginal increase in rents in the treated region relative to the control regions. This pattern is particularly evident in the period leading up to the surge in housing starts—any observed rent growth in the treated region occurs only after 2015, well after the increase in new construction. By contrast, housing starts in the treated region nearly tripled relative to the control trend by 2015. This lack of a substantial and immediate price response suggests that the increase in housing supply was not demand-driven, but rather a result of eased supply regulations following amalgamation.

Columns (3) to (6) in Table A.11 reinforce the conclusion that the increase in housing starts was driven by a supply-side response. Columns (3) and (4) examine the effect of the amalgamation on rents, showing statistically insignificant results in Panel A, but statistically significant results in Panel B, where the control municipalities are aggregated, smoothing out the time-series variation in rent growth across control municipalities. Despite some estimates being statistically significant, the point estimates in both panels remain an order of magnitude smaller than those for housing starts, indicating that the increase in housing starts was largely driven by supply shocks rather than demand shocks. Columns (5) and (6) report the effects on prices of newly built units. None of the

consistent with our theoretical model, is that the magnitude of the effect depends on both the size of the pre-merger municipalities and the scale of the newly amalgamated jurisdiction. Mast and Tricaud average across many mergers, some involving substantial changes and others relatively minor.

15. For example, see [Diamond and McQuade \(2019\)](#).

16. [Saiz \(2010\)](#) estimates the housing supply elasticity to be 1.75 for the average U.S. metro, and around 0.8 for larger metros like Toronto. This implies that a 1% rise in prices or rents (capitalized into prices) should be associated with a 0.8% increase in supply.

estimates are statistically significant, with larger but still insignificant estimates in Panel A, driven by smaller municipalities with greater price volatility. In Panel B, the point estimates are close to zero, further suggesting that it is unlikely that the rise in housing starts in amalgamated Toronto after 1998 was driven by demand shocks.¹⁷

17. If construction costs were low and delays short, then a demand shock would result in a rise in construction but little effect on prices or rents as new supply would offset new demand. However, in the real world, construction delays are significant. Delays, even with low costs, lead to a temporary rise in prices in response to demand shocks. Furthermore, real-world construction costs exhibit convexity. For example, once all buildable land has been used, it becomes increasingly expensive to flatten hills or dry swamps. Similarly, building tall skyscrapers is more expensive per square foot than single-family homes. Such costs lead to permanent increases in prices in response to demand shocks. On the other hand, an increase in housing starts but little change in prices or rents is consistent with a case where individuals are relatively indifferent between living in treated and control regions—leading to little difference in prices or rents—but where relaxed regulation leads to more construction in the treated regions.

Figure A.11 — Toronto map of treated and control areas

The figure depicts a map of the city of Toronto, as well as several neighboring municipalities, all in the Toronto metropolitan area. The municipal borders are drawn by the thicker lines. Mississauga, Brampton, Vaughan, Markham, and Pickering are in the control group, while Toronto is in the treated group. Within Toronto, the former municipal borders (dissolved in 1998) of Etobicoke, North York, York, East York, Scarborough, and old Toronto are outlined by the thinner lines.

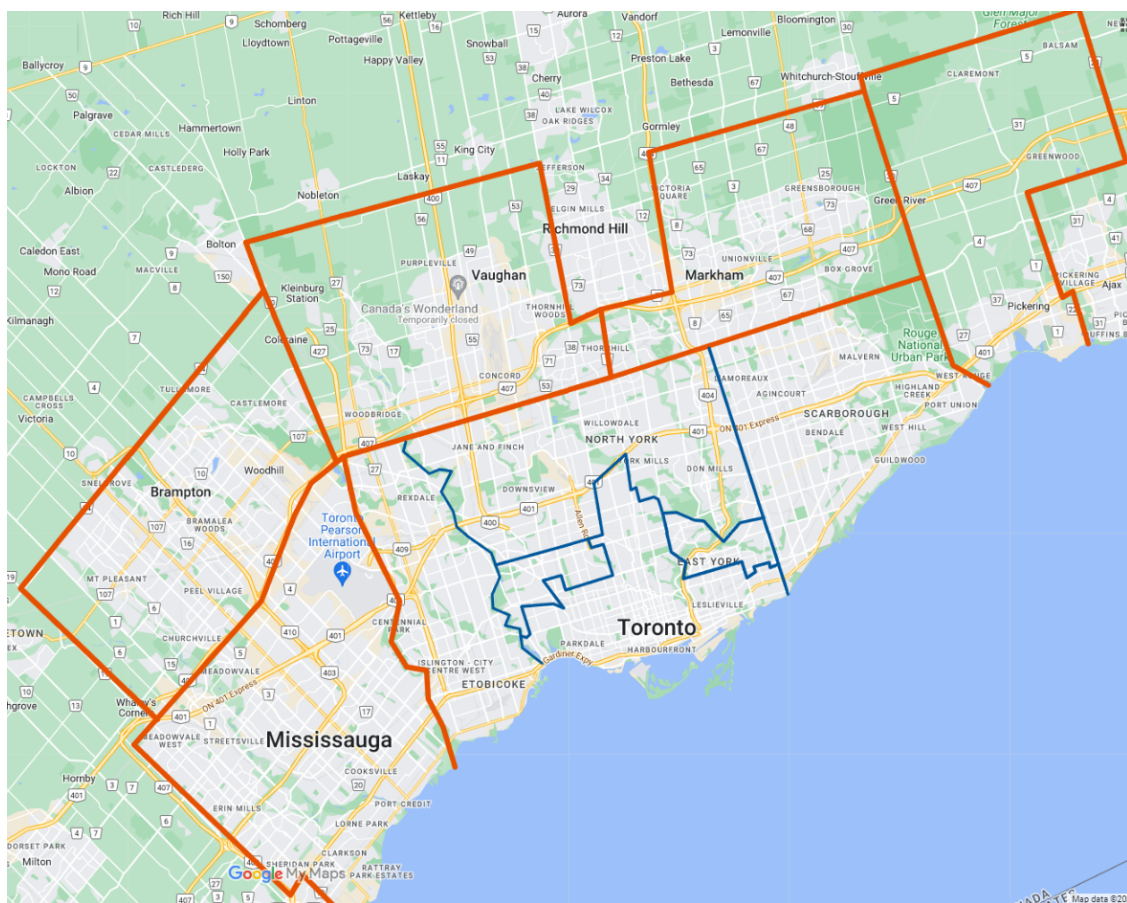


Figure A.12 — Toronto amalgamation, raw data

The figure depicts the 1990–2020 time series of total housing starts, average rent, and average price for newly built properties. The solid line represents the current Toronto municipality, which, before 1998, encompassed the former municipalities of Etobicoke, North York, York, East York, old Toronto, and Scarborough. The dashed line represents the aggregate of five neighboring municipalities—Mississauga, Brampton, Vaughan, Markham, and Pickering—which each share a contiguous boundary with the City of Toronto. When aggregating, average rents and prices were weighted by each city’s number of rental and newly built properties, respectively. The data comes from the Canada Mortgage and Housing Corporation (CMHC).

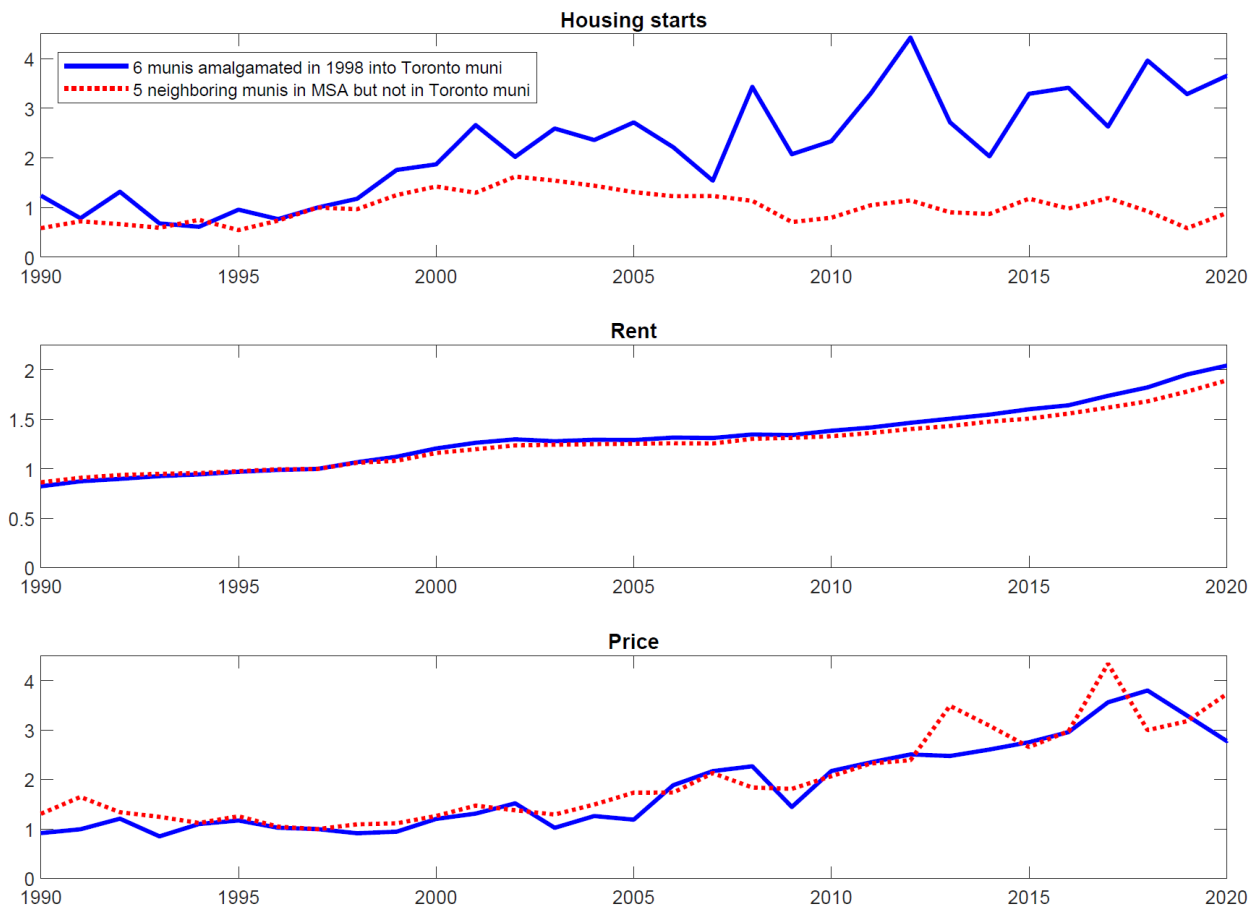


Table A.11 — Estimated effect of Toronto’s amalgamation in the housing market

This table presents the results of the regression in Equation A.4. The dependent variable in Columns (1) and (2) is the number of housing starts, in Columns (3) and (4) is the rent, and in Columns (5) and (6) is the price of newly built units, all in logs. The data is collected from the Canada Mortgage and Housing Corporation (CMHC) for the period 1990 to 2020. All columns control for fixed effects and a treated dummy. Additionally, Columns (2), (4), and (6) include controls for median income in each municipality by year and for population density as of 1989. These controls, obtained from Statistics Canada, are available in 5-year intervals and are interpolated for intermediate years.

Outcome	(1) log # starts	(2) log # starts	(3) log rent	(4) log rent	(5) log price	(6) log price
Panel A: Treated versus individual untreated municipalities						
Treated x Post	0.7453*** (0.2566)	0.7846*** (0.2572)	0.0343 (0.0296)	0.0085 (0.0218)	0.4263 (0.3012)	0.3774 (0.3010)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Income, Density controls	No	Yes	No	Yes	No	Yes
Geographies (N)	6	6	5	5	6	6
Years (T)	31	31	31	31	31	31
Observations (N×T)	186	155	186	186	155	186
Adjusted R^2	0.3213	0.3252	0.9162	0.9554	0.0566	0.0686
Panel B: Treated versus single merged untreated						
Treated x Post	0.6522*** (0.1983)	0.4438*** (0.1161)	0.0660*** (0.0079)	0.0606*** (0.0070)	0.0732 (0.0788)	0.0901 (0.0812)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Income	No	Yes	No	Yes	No	Yes
Geographies (N)	2	2	2	2	2	2
Years (T)	31	31	31	31	31	31
Observations (N×T)	62	62	62	62	62	62
Adjusted R^2	0.6634	0.8907	0.9961	0.9971	0.9114	0.9108

Significance: ***: 0.01, **: 0.05, *: 0.1

References

- Barro, R. J. and Sala-I-Martin, X. (1992). Public finance in models of economic growth. *The Review of Economic Studies*, 59(4):645–661.
- Boustan, L. P. (2010). Was postwar suburbanization “white flight”? evidence from the black migration. *The Quarterly Journal of Economics*, 125(1):417–443.
- Chernozhukov, V., Hansen, C., and Spindler, M. (2015). Post-selection and post-regularization inference in linear models with many controls and instruments. *American Economic Review*, 105(5):486–490.
- Cui, T. (2024). Did race fence off the american city? the great migration and the evolution of exclusionary zoning. Technical report, Working Paper.
- Dávila, E. and Zlate, A. (2022). Monetary policy in times of structural reallocation. Finance and Economics Discussion Series 2022-003, Board of Governors of the Federal Reserve System.
- Derenoncourt, E. (2022). Can you move to opportunity? evidence from the great migration. *American Economic Review*, 112(2):369–408.
- 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.
- Garrett, M. (1999). Building the new city of toronto: Status report on amalgamation. Technical report, City of Toronto, Toronto, Canada.
- Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature*, 47(4):983–1028.
- Goodman, C. B. (2023). Municipal incorporation data, 1789–2020. Software, version 1.0.0. Available at <https://github.com/cbgoodman/muni-incorporation/>.
- Hoxby, C. M. (2000). Does competition among public schools benefit students and taxpayers? *American Economic Review*, 90(5):1209–1238.
- Mast, E. (2024). Warding off development: Local control, housing supply, and nimbys. *Review of Economics and Statistics*, 106(3):671–680.
- Sahn, A. (2021). Racial diversity and exclusionary zoning: Evidence from the great migration. *Working paper*.
- Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.
- Song, J. (2025). The effects of residential zoning in us housing markets. *Journal of Urban Economics*, 149:103784.

Tax Policy Center (2021). What are the sources of revenue for state and local governments? Tax Policy Center Briefing Book.

Tricaud, C. (2025). Better alone? evidence on the costs of intermunicipal cooperation. *American Economic Journal: Applied Economics*, 17(1):160–207.