Exclusionary Zoning and Transportation

Jaehee Song

Yale University

This note is work-in-progress and the results are subject to be updated.

January 1, 2022

Abstract

Historians document that suburban communities historically set stricter zoning laws to prevent the inflow of racial minorities. They argue such practices were especially common in response to transportation developments, which allowed minorities to more easily access suburban communities. In this project, I study the effect of increased potential inflow of racial minorities on the stringency of residential zoning by leveraging variations in transportation networks and out-migration shocks. To do this, I assemble a data set on the railroad and highway networks during the 1900s. I first estimate a model of migration using 1935 to 1940 migration flows. Then, I use the estimated model to predict the inflow of Black households from 1940 to 1970 given the change in transportation networks and economic conditions in their prior locations. My preliminary findings indicate that municipalities that expected more Black inflow imposed stricter minimum lot area regulations.

Historical background of exclusionary zoning

Local governments have been accused of imposing restrictive zoning ordinances

- for the purpose of excluding racial minorities and low-income households
- e.g. banning multi-family homes, imposing stringent dimensional requirements
- requiring residential development to take expensive forms (i.e. large single-family homes)
- hence preventing racial minorities and low-income households from moving to their communities

Such exclusionary zoning practices were especially popular during post-World War II suburbanization

- whites moved to suburbs ("white flight") in response to the influx of Black population into cities in the North during the Second Great Migration (Boustan, 2010)
- these suburban communities enacted restrictive zoning policies (Nicolaides and Wiese, 2017)
- central cities also imposed restrictive zoning in response to the Black migration (Sahn, 2020)

This paper

1. **Provide new empirical evidence of exclusionary zoning in both urban and suburban areas**
   - Preliminary findings suggest local communities that faced higher influx of Black population during the Second Great Migration imposed stricter zoning
   - Empirical strategy adopts a Bartik instrument (Boustan, 2010; Derenoncourt, 2021) as well as a new railroad instrument

2. **Use transportation connectivity as a shifter of "exposure" to Black population**
   - Railroad instrument for the Second Great Migration captures the migration costs between southern counties and northern counties
   - Highway connection as a local community-level variation of potential inflow of Black population from nearby central cities (in progress)

3. **Investigate the relationship between exclusionary zoning and postwar suburbanization**
   - Preliminary findings suggest local communities developed since the mid 1900s imposed stricter zoning

Data

Demographics data

Data 1: the complete-count 1940 U.S. Census

- individual-level demographics including race
- geocoded using SmartyStreets

Data 2: City and County Data Books (CCDB) 1944-1977 series from ICSPR

- population demographics including # white, # non-white, and # Black

→ **Measure Black migration flows during 1935-1940 and during 1940-1970**

- Δ Black population/Total population

→ **Identify suburban communities created during the mid 1900s**

- # individuals appeared on 1940 Census

Zoning data

Data: Song (2021), "The Effects of Residential Zoning in U.S. Housing Markets"

- nationwide data set of neighborhood-level minimum lot area estimates

→ **Estimated stringency and restrictiveness of local zoning**

- minimum lot area restriction levels
- % bunching at the minimum lot areas

Railroad data

Data: Jeremy Atack, "Historical GIS database of U.S. Railroads for 1911"

→ **Measure the connectivity between Southern counties and Northern counties**

- whether rail connection exists
- shortest rail distance connecting the two

Highway data

Data 1: National Highway Planning Network 2000

- comprehensive GIS database of major U.S. highway systems: Rural Arterials, Urban Principal Arterials and all National Highway System routes

Data 2: PR-511 data (Baum-Snow, 2007)

- highway segments funded by the 1956 Federal Highways Act, other federally funded segments, locally funded segments, and toll segments for highway miles constructed between 1950 and 1990

→ **Measure the connectivity between suburban communities and their nearby urban cities**

Descriptive Analysis

Suburban communities impose more restrictive zoning

Table: Summary statistics of minimum lot areas (cities vs. suburbs)

	percentile					
Min lot area (in sq.ft.)	10th	25th	50th	75th	90th	mean
Cities	3,750	6,098	8,712	12,192	20,038	11,227
Suburbs	8,882	14,375	30,056	43,821	87,120	37,722

Note. This table reports summary statistics of minimum lot areas for cities and suburbs. Cities and suburbs are identified using the 2019 Census TIGER: cities are regions that intersect Census places, and suburbs are regions in Core-Based Statistical Areas that do not belong to any Census places. The data is restricted to Core-Based Statistical Areas in states where county subdivisions perform as general-purpose local governments (CT, IL, IN, KS, ME, MA, MI, MN, MO, NE, NH, NJ, NY, ND, OH, PA, RI, SD, VT, and WI). Minimum lot area estimates are constructed using data from Song (2021) and restricted to its robust minimum lot area estimates.

Communities developed after 1940 impose more restrictive zoning

Table: Summary statistics of minimum lot areas (developed before 1940 vs. after 1940)

	percentile					
Min lot area (in sq.ft.)	10th	25th	50th	75th	90th	mean
Developed before 1940	3,750	5,316	7,814	11,011	15,246	9,706
Developed after 1940	7,500	10,454	20,038	43,560	80,150	31,148

Note. This table reports summary statistics of minimum lot areas by the development status. Local communities are defined as "Developed after 1940" if no geocoded addresses in the full-count 1940 Census is located in the community. The data is restricted to Core-Based Statistical Areas in states where county subdivisions perform as general-purpose local governments (CT, IL, IN, KS, ME, MA, MI, MN, MO, NE, NH, NJ, NY, ND, OH, PA, RI, SD, VT, and WI).

The Effects of the Great Migration on Zoning Stringency

Empirical framework

Regress zoning stringency on Black migration inflow

Regression equation:

$$y_k = \beta_{GM} GM_k + \beta_X X_k + \varepsilon_k$$

- k: a destination county in the North
- y: stringency of zoning, measured by min lot areas
- GM: Black population change during the Second Great Migration
- X: observable location characteristics

$\Rightarrow \beta_{GM}$ is the parameter of interest where $\beta_{GM} > 0$ suggests exclusionary zoning practices

Empirical framework

Defining the Second Great Migration flow

The Great Migration (GM) is measured by

$$GM_k = \frac{b_{k,1970} - b_{k,1940}}{pop_{k,1940}}$$

- $b_{k,t}$: # Black at county k in year t
- $pop_{k,t}$: total population at county k in year t

$\Rightarrow$ Endogeneity of GM

- endogeneity migration destinations: unobservable characteristics of k affect both GM_k and ε_k
- reverse causality: having strict zoning may deter Black migration

$\Rightarrow$ Often addressed by using the Bartik instrument (Boustan, 2010; Derenoncourt, 2021; Sahn, 2021)

The Bartik instrument

Interacting pre-1940 migration rate with origin push factors

Decompose # Black migrants by their origin county o:

$$\begin{aligned}\tilde{GM}_{k,Bartik} &= \frac{1}{pop_{1940,k}} \sum_o \hat{n}_{o,k,1940to1970} \\ &= \sum_o \underbrace{\frac{1}{pop_{1940,k}} \cdot \frac{n_{o,k,1935to1940}}{N_{o,1935to1940}}}_{\substack{\text{pre-1940 linkage between } o \text{ and } k \\ (\text{“exposure”})}} \cdot \underbrace{\hat{N}_{o,1940to1970}}_{\substack{\text{predicted outmigration at the origin} \\ (\text{“shocks”})}}\end{aligned}$$

- $\hat{n}_{o,k,t}$: predicted # black migrants in county k from county o in period t
- $n_{o,k,t}$: # black migrants in county k from county o in period t
- $N_{o,t}$: Total migration from origin o in period t ($= \sum_k n_{o,k}$)
- $\hat{N}_{o,t}$: Predicted outmigration in period t using “push” factors of southern counties (mostly agriculture/mining industry factors) using LASSO (Derenoncourt, 2021)

Constructing the new railroad instrument

Interacting railroad linkage with origin push factors

$$\tilde{GM}_{k,rail} = \sum_{o} \underbrace{\frac{1}{pop_{1940,k}} \cdot \widehat{\frac{n_{o,k,1935to1940}}{N_{o,1935to1940}}}}_{\substack{\text{pre-1940 linkage between } o \text{ and } k \\ \text{attributed to railroad network} \\ (\text{“exposure”})}} \cdot \underbrace{\hat{N}_{o,1940to1970}}_{\substack{\text{predicted outmigration at the origin} \\ (\text{“shocks”})}}$$

Constructing the new railroad instrument

Railroad networks shift migration—migration less likely with longer travel distances

$$\text{Migration rate}_{o,k} = \frac{1}{pop_{1940,k}} \cdot \frac{n_{o,k,1935to1940}}{N_{o,1935to1940}}$$

	Outcome variable: migration rate (in 10^{-8})					
	(1)	(2)	(3)	(4)	(5)	(6)
log rail distance	−2.22***	−2.66***	−3.51***	−25.64***	−16.60***	−15.90***
× 𝟙(rail connected)	(0.03)	(0.04)	(0.05)	(0.84)	(0.84)	(0.90)
log centroid distance	−1.49***	−1.79***	−1.99***	−24.86***	−13.49***	−9.71***
× 𝟙(rail not connected)	(0.05)	(0.05)	(0.06)	(1.93)	(1.64)	(1.85)
𝟙(rail connected)	11.36***	13.62***	23.74***	24.94	50.24*	103.4***
	(0.82)	(0.89)	(1.04)	(28.09)	(24.16)	(26.29)
Destination fixed effects	No	Yes	Yes	No	Yes	Yes
Origin fixed effects	No	No	Yes	No	No	Yes
Non-zero migration only	No	No	No	Yes	Yes	Yes

Note. Destination counties are restricted to Core-Based Statistical Ares in states where county subdivisions are general-purpose local governments: CT, IL, IN, KS, ME, MA, MI, MN, MO, NE, NH, NJ, NY, ND, OH, PA, RI, SD, VT, and WI. Origin counties are restricted to AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA, and WV. Rail distance is defined as the shortest distance from a location in the origin county to a location in the destination county along the railroad network in 1911. Centroid distance is defined by the straight line distance between the centroid of the origin county and the centroid of the destination county. Specifications #1–#3 use the full sample of all possible origin–destination pairs while Specifications #4–#6 use the subsample of origin–destination pairs with non-zero Black migrants during 1935–1940.

▸ More about migration rate and travel distances

Empirical evidence of exclusionary zoning

Higher GM led to stricter min lot area restrictions

i: municipality, k: northern county

$$y_{ik} = \beta_{GM} GM_{ik} + \beta_X X_{ik} + \varepsilon_{ik}$$

	outcome variable: minimum lot area (in acre)					
	OLS		Bartik IV		Railroad IV	
	(1)	(2)	(3)	(4)	(5)	(6)
GM (Black migration)	−1.036***	−1.114***	0.717*	1.115***	2.997***	3.102***
	(0.133)	(0.126)	(0.378)	(0.358)	(0.422)	(0.400)
$\mathbb{1}$(postwar suburb)		0.593***		0.591***		0.589***
		(0.009)		(0.009)		(0.009)
log agricultural land	0.007***	0.002***	0.006***	0.002***	0.006***	0.002***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
log industrial land	−0.019***	−0.013***	−0.019***	−0.013***	−0.019***	−0.013***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
log total land	0.086***	0.093***	0.086***	0.093***	0.086***	0.093***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
log population in 1940	0.253***	0.865***	−1.362***	−1.189***	−3.456***	−3.021***
	(0.225)	(0.213)	(0.396)	(0.376)	(0.432)	(0.411)
log # white in 1940	−0.362	−0.909***	1.215***	1.097***	3.261***	2.887***
	(0.226)	(0.214)	(0.390)	(0.370)	(0.425)	(0.404)

Note. Destination counties are restricted to Core-Based Statistical Ares in states where county subdivisions are general-purpose local governments: CT, IL, IN, KS, ME, MA, MI, MN, MO, NE, NH, NJ, NY, ND, OH, PA, RI, SD, VT, and WI. Origin counties are restricted to AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA, and WV. "Exposure" term in Railroad IV (as defined in page 16) is constructed using Specification #5 in page 17. "Shocks" term (as defined in page 15 and page 16) in both Bartik IV and Railroad IV is calculated following Derenoncourt (2021) using LASSO. GM is calculated using the full-count 1940 Census and CCDB. $\mathbb{1}$(postwar suburb) indicates whether the municipality appears in the geocoded full-count 1940 Census. Agricultural, industrial, and total (developable) land are calculated using CoreLogic property tax data from 2018. Total population in 1940 and # white in 1940 are taken from the full-count 1940 Census.

Empirical evidence of exclusionary zoning

Results on logarithm of min lot area

i: municipality, k: northern county

$$y_{ik} = \beta_{GM} GM_{ik} + \beta_X X_{ik} + \varepsilon_{ik}$$

	outcome variable: log minimum lot area					
	OLS		Bartik IV		Railroad IV	
	(1)	(2)	(3)	(4)	(5)	(6)
GM (Black migration)	−1.173***	−1.295***	0.978**	1.595***	0.613	0.778**
	(0.136)	(0.119)	(0.387)	(0.339)	(0.427)	(0.374)
$\mathbb{1}$(postwar suburb)		0.919***		0.918***		0.917***
		(0.008)		(0.008)		(0.008)
log agricultural land	0.014***	0.007***	0.014***	0.007***	0.014***	0.007***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
log industrial land	−0.022***	−0.013***	−0.022***	−0.013***	−0.022***	−0.013***
	(0.001)	(0.0005)	(0.001)	(0.0005)	(0.001)	(0.0005)
log total land	0.047***	0.057***	0.047***	0.057***	0.047***	0.057***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)
log population in 1940	−0.505**	0.445**	−2.485***	−2.217***	−2.143***	−1.466***
	(0.230)	(0.201)	(0.405)	(0.355)	(0.438)	(0.383)
log # white in 1940	0.202	−0.567***	2.217***	2.034***	1.882***	1.299***
	(0.231)	(0.202)	(0.399)	(0.350)	(0.431)	(0.377)

Note. Destination counties are restricted to Core-Based Statistical Ares in states where county subdivisions are general-purpose local governments: CT, IL, IN, KS, ME, MA, MI, MN, MO, NE, NH, NJ, NY, ND, OH, PA, RI, SD, VT, and WI. Origin counties are restricted to AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA, and WV. "Exposure" term in Railroad IV (as defined in page 16) is constructed using Specification #5 in page 17. "Shocks" term (as defined in page 15 and page 16) in both Bartik IV and Railroad IV is calculated following Derenoncourt (2021) using LASSO. GM is calculated using the full-count 1940 Census and CCDB. $\mathbb{1}$(postwar suburb) indicates whether the municipality appears in the geocoded full-count 1940 Census. Agricultural, industrial, and total (developable) land are calculated using CoreLogic property tax data from 2018. Total population in 1940 and # white in 1940 are taken from the full-count 1940 Census.

Conclusion

Preliminary findings:

- Post-World War II suburbs tend to have higher min lot area (stricter zoning)
- 1 standard deviation ↑ of Black migration during the Great Migration ⇒ min lot area ↑ by 0.5 acre
- Using the Bartik or railroad IV reverses the relationship between Black migration and zoning stringency

Moving forward:

- Incorporating highway connectivity between central cities and suburbs in the analysis (Baum-Snow, 2007)
  - Did suburbs with highway connections to cities that experienced Black migration (and thus faced higher potential of Black influx) impose stricter zoning?
- Interpreting the railroad IV in the Bartik IV framework (Goldsmith-Pinkham et al., 2018; Adao et al., 2019; Borusyak and Hull, 2021)

Appendix

Migration rate and travel distances

Even after controlling for centroid distances, rail distances are related to migration rate

$$\text{Migration rate}_{o,k} = \frac{1}{pop_{1940,k}} \cdot \frac{n_{o,k,1935to1940}}{N_{o,1935to1940}}$$

	outcome variable								
	migration rate (in 10^{-8})						log migration rate		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log rail distance	-1.543^{***}	-1.814^{***}	-2.512^{***}	-13.545^{***}	-12.168^{***}	-12.523^{***}	-0.266^{***}	-0.100^{***}	-0.040^{***}
	(0.044)	(0.049)	(0.062)	(1.069)	(0.960)	(1.002)	(0.019)	(0.019)	(0.017)
log centroid distance	-0.840^{***}	-1.045^{***}	-1.090^{***}	-23.156^{***}	-10.118^{***}	-9.943^{***}	-1.143^{***}	-0.925^{***}	-1.030^{***}
	(0.036)	(0.040)	(0.051)	(1.269)	(1.156)	(1.846)	(0.023)	(0.022)	(0.032)
$\mathbb{1}$(rail connected)	22.162^{***}	26.187^{***}	37.172^{***}	190.919^{***}	169.131^{***}	185.317^{***}	3.429^{***}	1.124^{***}	0.411^{***}
	(0.638)	(0.699)	(0.904)	(14.867)	(13.473)	(14.570)	(0.270)	(0.262)	(0.251)
Destination fixed effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Origin fixed effects	No	No	Yes	No	No	Yes	No	No	Yes
Non-zero migration only	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Note. Destination counties are restricted the states where county subdivisions are general-purpose local governments: CT, IL, IN, KS, ME, MA, MI, MN, MO, NE, NH, NJ, NY, ND, OH, PA, RI, SD, VT, and WI. Origin counties are restricted to AL, AR, FL, GA, KY, LA, MS, NC, OK, SC, TN, TX, VA, and WV. Rail distance is defined by the shortest distance from a location in the origin county to a location in the destination county along the railroad network in 1911. Centroid distance is defined by the straight line distance between the centroid of the origin county and the centroid of the destination county. Specifications #1–#3 use the full sample of all possible origin–destination pairs while Specifications #4–#9 use the subsample of origin–destination pairs with non-zero Black migrants during 1935–1940.