Milestone 6

Maria Burzillo

3/28/2020

Introduction

This is my pdf document. Please refer to the Github repository of my final project for further information.[1].
You can also access all of the original replication materials made available by Trounstine on Harvard Dataverse
here. I make use of Trounstine (2016), Pencharz and Ball (2003), Xie (2020), Wickham (2019), and Xie
(2015).

Summary of Trounstine (2016)

Trounstine's *Segregation and Inequality in Public Goods* attempts to explain differences in public goods
provision and political polarization through a racial lens by examining the relationships between polarization,
goods provision, and segregation. Trounstine measures segregation with Theil's H index, which measures the
degree to which the diversity of a neighborhood differs from the diversity of the entire city. The main finding
of the paper is that segregation, not simply diversity or political views, is an important determinant of both
political polarization and spending on public goods. In general, segregation leads to the coincidence of racial
and spatial political cleavages, which can make compromise on taxation and public spending difficult and
tends to generally drive down the rate of spending on public goods. Because minorities are much more likely
to live in racially segregated areas than whites, this suggests that public goods are also segregated across
racial lines.

Replication

Table 1 was able to be replicated exactly. The replication for Table 2 was close, but not exact, as were
the replications for Table 3 and 4, which combined replicated Table 3 in the main paper. However, the
implications of the main results are essentially the same for all of these regressions. The IV regression was
able to be replicated exactly and the results presented in Tables 5 and 6 reconstruct the results from Table 5
in the original paper. So far, I have not been able to successfully recreate the marginal effects; however, I
am working to figure out what the problem is. I believe that with a little bit more time I will be able to
successfully recreate all of the results given my success in Table 1 and Tables 5 and 6.

As for the paper's Appendix Tables, I was able to successfully recreate Tables A1 and A2. I have not yet
been able to sucessfully recreate Table A3 in Stata because I am having difficulty in figuring out why I have
fewer observations and also how they calculated some of their statistics, as it seems that they filter for some
conditions only for some rows, which I have yet to figure out how to do in R. I did not attempt to recreate
Table A4 due to time constraints and also because it seemed extraneous.

---

[1] All sources, analysis, and further information are available on my Github repository for this project

Extension Ideas

There are a variety of ways that I could build upon this analysis. Because I have not narrowed it down to one yet, I will use this as an opportunity to propose a few ideas.

How well does this theory apply to more recent elections, and can we use new data to test it? For example, I could try to hunt down some more recent election data and use some sort of prediction funciton to see how well the model predicts the actual results. However, this could be quite difficult if the data is too messy or difficult to obtain (or even non-existent).

Another idea could be to redo the results but using the dissimilarity index, the most common measure of segregation, which Trounstine, perhaps controversially, chooses not to use in her analysis in favor of Theil's entropy score. It could be interesting to see whether her results hold up against this sort of robustness check.

Another interesting thing to do would be to find a city that has recently become less segregated and see if political polarization has decreased and public goods provision has increased.

Finally, another approach to put to practice some more of the skills we've learned in this class would be to try and adpot a Bayesian framework to some of Trounstine's analyses.

Main Analysis

Table 1

Margins from Table 1 Calculations

TO DO

Table 2

Predicted Effects Following Table 2

Table 3

Figure 1

Main Analysis 4

Main Analysis 5

Table 5

Appendix

TABLE A2 Cities Included in Racial Polarization Data
Replication

City Name	Segregation: Mean H Index		Largest Racial Divide, Number of Elections		
	Multigroup	Two-Group	Black/White	Latino/White	Black/Latino
Austin, TX	0.204	0.208	1	0	0
Baltimore, MD	0.510	0.516	3	1	0
Charlotte, NC	0.269	0.287	2	0	0
Chicago, IL	0.572	0.460	7	0	1
Cleveland, OH	0.558	0.531	2	0	0
Columbus, OH	0.316	0.284	3	0	1
Dallas, TX	0.359	0.339	4	0	1
Denver, CO	0.289	0.254	1	2	0
Detroit, MI	0.398	0.255	1	0	1
Houston, TX	0.339	0.308	7	0	2
Indianapolis, IN	0.292	0.293	0	0	1
Jacksonville, FL	0.233	0.222	2	0	0
Los Angeles, CA	0.351	0.366	3	0	5
Memphis, TN	0.470	0.474	2	0	0
Milwaukee, WI	0.423	0.360	3	0	0
New York, NY	0.468	0.474	5	3	1
Oklahoma, OK	0.231	0.165	1	0	0
Philadelphia, PA	0.492	0.487	5	0	0
Phoenix, AZ	0.255	0.270	0	1	0
San Antonio, TX	0.237	0.225	0	4	0
San Diego, CA	0.255	0.266	3	0	1
San Francisco, CA	0.223	0.161	3	0	1
San Jose, CA	0.186	0.198	0	2	1
Tucson, AZ	0.185	0.192	1	0	0
Washington, DC	0.464	0.491	3	0	0

Bibliography

Pencharz, Paul B., and Ronald O. Ball. 2003. "Different Approaches to Define Individual Amino Acid Requirements." *Annual Review of Nutrition* 23. Annual Reviews: 101–16.

Trounstine, Jessica. 2016. "Segregation and Inequality in Public Goods: SEGREGATION AND INEQUALITY IN PUBLIC GOODS." *American Journal of Political Science* 60 (3): 709–25. https://doi.org/10.1111/ajps.12227.

Wickham, Hadley. 2019. *Stringr: Simple, Consistent Wrappers for Common String Operations.* https://CRAN.R-project.org/package=stringr.

Xie, Yihui. 2015. *Dynamic Documents with R and Knitr.* 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. https://yihui.org/knitr/.

———. 2020. *Knitr: A General-Purpose Package for Dynamic Report Generation in R.* https://CRAN.R-project.org/package=knitr.

Table 1: **Racial Polarization in Segregated Cities**

	Dependent variable:		
	biggestsplit		
	(1)	(2)	(3)
H_citytract_multi_i	0.932**		
	(0.394)		
H_citytract_NHW_i		0.756**	0.835***
		(0.297)	(0.296)
diversityinterp	0.385	0.518	0.584*
	(0.362)	(0.323)	(0.323)
pctasianpopinterp	−0.115	0.120	−0.004
	(0.527)	(0.558)	(0.522)
pctblkpopinterp	−0.432	−0.237	−0.133
	(0.269)	(0.216)	(0.212)
pctlatinopopinterp	−0.191	−0.059	0.095
	(0.257)	(0.254)	(0.278)
medincinterp	−0.00000	−0.00001	−0.00000
	(0.00001)	(0.00001)	(0.00001)
pctrentersinterp	−0.580	−0.806*	−0.419
	(0.422)	(0.431)	(0.454)
pctcollegegradinterp	0.328	0.723	0.123
	(0.711)	(0.729)	(0.869)
biracial	0.210***	0.208***	0.192***
	(0.037)	(0.037)	(0.036)
nonpartisan	−0.090	−0.089	−0.034
	(0.066)	(0.066)	(0.065)
primary	−0.092***	−0.090***	−0.071**
	(0.032)	(0.032)	(0.030)
logpop	0.035	0.048	−0.011
	(0.055)	(0.055)	(0.061)
whiteideology_fill2			−0.051
			(0.032)
Constant	−0.242	−0.393	0.236
	(0.569)	(0.563)	(0.605)
Observations	91	91	86
Log Likelihood	62.774	63.190	67.461
Akaike Inf. Crit.	−55.548	−56.381	−64.922
Bayesian Inf. Crit.	32.332	31.499	20.981

4

Table 2: **Effect of Segregation on Overall per Capita City Expenditures**

	Dependent variable:		
	dgepercap_cpi		
	(1)	(2)	(3)
H_citytract_NHW_i	−1.158***	−1.021***	−1.805***
	(0.221)	(0.254)	(0.501)
diversityinterp	0.106		
	(0.134)		
pctblkpopinterp	0.667***	0.736***	−0.027
	(0.167)	(0.161)	(0.555)
pctasianpopinterp	−0.385	−0.853**	−0.562
	(0.302)	(0.348)	(0.810)
pctlatinopopinterp	1.525***	1.561***	1.378***
	(0.186)	(0.204)	(0.407)
chng5pctblk		−1.791***	−3.444***
		(0.646)	(1.105)
chng5pctlatino		−2.060**	−1.427
		(0.824)	(1.106)
chng5pctasian		−0.846	−3.585
		(1.094)	(2.513)
medinc_cpi	0.00000*	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)
pctlocalgovworker_100	0.013	0.005	−0.029
	(0.016)	(0.018)	(0.066)
pctrentersinterp	0.539	0.565	0.224
	(0.333)	(0.386)	(0.810)
pctover65	0.069	0.494	−0.708
	(0.646)	(0.452)	(0.948)
pctcollegegradinterp	5.379***	6.232***	7.288***
	(0.403)	(0.419)	(1.244)
logpop	−0.232***	−0.279***	−0.504***
	(0.044)	(0.068)	(0.112)
ideology_fill			−0.008
			(0.030)
Observations	13,684	11,149	1,759
R^2	0.864	0.897	0.902
Adjusted R^2	0.831	0.865	0.874
Residual Std. Error	0.502 (df = 11050)	0.464 (df = 8513)	0.390 (df = 1370)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3: **Effect of Segregation on Public Goods A**

	Dependent variable:		
	highwayspercapNC_cpi	policepercapNC_cpi	parkspercapNC_cpi
	(1)	(2)	(3)
H_citytract_NHW_i	−0.041***	−0.215***	−0.047***
	(0.016)	(0.023)	(0.017)
diversityinterp	0.003	0.060***	−0.002
	(0.010)	(0.013)	(0.012)
pctblkpopinterp	0.052***	0.138***	0.034**
	(0.013)	(0.018)	(0.016)
pctasianpopinterp	−0.036	−0.054	−0.061***
	(0.026)	(0.035)	(0.023)
pctlatinopopinterp	0.025*	0.330***	0.049***
	(0.014)	(0.019)	(0.014)
medinc_cpi	0.00000**	0.000	−0.00000
	(0.00000)	(0.00000)	(0.00000)
pctlocalgovworker_100	−0.0004	−0.001	0.001
	(0.001)	(0.002)	(0.001)
pctrentersinterp	0.007	0.082***	0.030
	(0.022)	(0.028)	(0.020)
pctover65	0.148***	0.140***	0.110***
	(0.031)	(0.045)	(0.039)
pctcollegegradinterp	0.209***	0.769***	0.421***
	(0.026)	(0.039)	(0.036)
logpop	−0.014***	−0.049***	−0.003
	(0.004)	(0.004)	(0.003)
Observations	13,684	13,684	13,684
R^2	0.570	0.835	0.749
Adjusted R^2	0.468	0.796	0.689
Residual Std. Error (df = 11050)	0.039	0.043	0.034

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: **Effect of Segregation on Public Goods B**

	Dependent variable:		
	sewerspercapNC_cpi	welfhoushealthNC_cpi	genrevownpercap_cpi
	(1)	(2)	(3)
H_citytract_NHW_i	−0.159***	−0.130***	−0.777***
	(0.022)	(0.041)	(0.155)
diversityinterp	0.062***	−0.023	0.094
	(0.012)	(0.020)	(0.085)
pctblkpopinterp	0.001	0.004	0.261**
	(0.017)	(0.037)	(0.120)
pctasianpopinterp	−0.113***	0.143*	−0.143
	(0.034)	(0.076)	(0.233)
pctlatinopopinterp	0.064***	0.137***	1.187***
	(0.016)	(0.024)	(0.120)
medinc_cpi	0.00000***	−0.00000	0.00000***
	(0.00000)	(0.00000)	(0.00000)
pctlocalgovworker_100	−0.004**	−0.007***	0.001
	(0.002)	(0.002)	(0.013)
pctrentersinterp	0.141***	0.082***	0.567**
	(0.034)	(0.032)	(0.263)
pctover65	−0.003	−0.061	0.449
	(0.043)	(0.052)	(0.474)
pctcollegegradinterp	0.232***	0.306***	4.306***
	(0.037)	(0.063)	(0.350)
logpop	−0.010***	−0.010**	−0.118***
	(0.003)	(0.004)	(0.032)
Observations	13,684	13,684	13,684
R^2	0.669	0.821	0.886
Adjusted R^2	0.590	0.779	0.859
Residual Std. Error (df = 11050)	0.050	0.056	0.420

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5: **Effect of Segregation on Public Goods**

	Dependent variable:				
	highwayspercapNC_cpi	policepercapNC_cpi	parkspercapNC_cpi	sewerspercapNC_cpi	welfhoushealthNC_cpi
	(1)	(2)	(3)	(4)	(5)
H_citytract_NHW_i	−0.041***	−0.215***	−0.047***	−0.159***	−0.130***
	(0.016)	(0.023)	(0.017)	(0.022)	(0.041)
diversityinterp	0.003	0.060***	−0.002	0.062***	−0.023
	(0.010)	(0.013)	(0.012)	(0.012)	(0.020)
pctblkpopinterp	0.052***	0.138***	0.034**	0.001	0.004
	(0.013)	(0.018)	(0.016)	(0.017)	(0.037)
pctasianpopinterp	−0.036	−0.054	−0.061***	−0.113***	0.143*
	(0.026)	(0.035)	(0.023)	(0.034)	(0.076)
pctlatinopopinterp	0.025*	0.330***	0.049***	0.064***	0.137***
	(0.014)	(0.019)	(0.014)	(0.016)	(0.024)
medinc_cpi	0.00000**	0.000	−0.00000	0.00000***	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
pctlocalgovworker_100	−0.0004	−0.001	0.001	−0.004**	−0.007***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
pctrentersinterp	0.007	0.082***	0.030	0.141***	0.082**
	(0.022)	(0.028)	(0.020)	(0.034)	(0.032)
pctover65	0.148***	0.140***	0.110***	−0.003	−0.061
	(0.031)	(0.045)	(0.039)	(0.043)	(0.052)
pctcollegegradinterp	0.209***	0.769***	0.421***	0.232***	0.306***
	(0.026)	(0.039)	(0.036)	(0.037)	(0.063)
logpop	−0.014***	−0.049***	−0.003	−0.010***	−0.010**
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
Observations	13,684	13,684	13,684	13,684	13,684
R^2	0.570	0.835	0.749	0.669	0.821
Adjusted R^2	0.468	0.796	0.689	0.590	0.779
Residual Std. Error (df = 11050)	0.039	0.043	0.034	0.050	0.056

Note: *p<0.1;

8

Table 6: **Effect of Segregation on City Expenditures, IV Approach A**

	Dependent variable:			
	dgepercap_cpi	highwayspercapNC_cpi	policepercapNC_cpi	parkspercapNC_cpi
	(1)	(2)	(3)	(4)
H_citytract_NHW_i	−2.676***	−0.363***	−0.350***	−0.034*
	(0.935)	(0.056)	(0.109)	(0.019)
dgepercap_cpilag	1.472***			
	(0.007)			
highwayspercapNC_cpilag		0.477***		
		(0.003)		
policepercapNC_cpilag			0.955***	
			(0.004)	
parkspercapNC_cpilag				0.869***
				(0.006)
diversityinterp	0.264	−0.032	−0.020	0.004
	(0.355)	(0.022)	(0.042)	(0.007)
pctblkpopinterp	0.376	0.085***	0.096**	0.003
	(0.325)	(0.020)	(0.038)	(0.007)
pctasianpopinterp	0.143	−0.111**	−0.064	−0.022
	(0.940)	(0.056)	(0.110)	(0.019)
pctlatinopopinterp	0.087	0.088***	0.088***	0.007
	(0.284)	(0.017)	(0.033)	(0.006)
medincinterp	−0.004	0.001***	0.001***	0.0002***
	(0.003)	(0.0002)	(0.0004)	(0.0001)
pctlocalgovworker_100	−0.104***	0.021***	0.026***	0.003***
	(0.032)	(0.002)	(0.004)	(0.001)
pctrentersinterp	−0.553	0.165***	0.187***	0.035***
	(0.350)	(0.021)	(0.041)	(0.007)
pctover65	0.301	0.267***	0.238***	0.072***
	(0.730)	(0.044)	(0.086)	(0.015)
pctcollegegradinterp	0.248	−0.044	−0.101	0.028**
	(0.567)	(0.034)	(0.067)	(0.012)
Constant	0.328	−0.125***	−0.183***	−0.029***
	(0.278)	(0.017)	(0.033)	(0.006)
Observations	21,145	20,704	20,627	19,056
R^2	0.685	0.615	0.789	0.540
Adjusted R^2	0.685	0.615	0.789	0.539
Residual Std. Error	4.877 (df = 21125)	0.290 (df = 20684)	0.566 (df = 20607)	0.093 (df = 19036)

Note: *p<0.1; **p<0.05; ***p<0.0

Table 7: **Effect of Segregation on City Expenditures, IV Approach B**

	Dependent variable:		
	sewerspercapNC_cpi	genrevownpercap_cpi	welfhoushealthNC_cpi
	(1)	(2)	(3)
H_citytract_NHW_i	−0.363***	−1.873**	−0.115**
	(0.060)	(0.789)	(0.054)
sewerspercapNC_cpilag	0.064***		
	(0.008)		
genrevownpercap_cpilag		1.235***	
		(0.006)	
welfhoushealthNC_cpilag			0.893***
			(0.005)
diversityinterp	0.080***	0.047	−0.047**
	(0.024)	(0.300)	(0.022)
pctblkpopinterp	0.058**	0.360	0.076***
	(0.025)	(0.274)	(0.023)
pctasianpopinterp	−0.223***	−0.029	0.009
	(0.068)	(0.793)	(0.053)
pctlatinopopinterp	−0.050***	0.206	0.078***
	(0.019)	(0.240)	(0.017)
medincinterp	0.0002	0.0003	0.001**
	(0.0003)	(0.003)	(0.0002)
pctlocalgovworker_100	0.001	−0.003	0.016***
	(0.002)	(0.027)	(0.002)
pctrentersinterp	0.073***	0.263	0.098***
	(0.024)	(0.295)	(0.023)
pctover65	0.287***	0.782	0.127**
	(0.051)	(0.616)	(0.050)
pctcollegegradinterp	0.029	−0.035	−0.038
	(0.040)	(0.478)	(0.038)
Constant	0.004	−0.174	−0.093***
	(0.019)	(0.234)	(0.018)
Observations	16,616	21,148	14,711
R^2	0.006	0.681	0.699
Adjusted R^2	0.005	0.681	0.698
Residual Std. Error	0.284 (df = 16596)	4.115 (df = 21128)	0.252 (df = 14691)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8: **TABLE A1 Summary Statistics: Racial Polarization Data**

Statistic	N	Mean	St. Dev.	Min	Max
Largest Racial Divide	91	0.481	0.213	0.016	0.934
H Index: Multigroup	91	0.376	0.119	0.183	0.635
H Index: Two-Group	91	0.353	0.114	0.156	0.614
Diversity	91	0.623	0.088	0.323	0.736
% Asian	91	0.067	0.074	0.008	0.318
% Black	91	0.275	0.181	0.030	0.815
% Latino	91	0.229	0.155	0.009	0.605
Median HH Income	91	36,724.530	10,113.880	17,267.500	75,981.730
% Renters	91	0.535	0.092	0.368	0.718
% College Degree	91	0.167	0.056	0.049	0.359
Biracial Contest	91	0.725	0.449	0	1
Nonpartisan Election	91	0.714	0.454	0	1
Primary Election	91	0.352	0.480	0	1
Population (logged)	91	14.166	0.826	13.065	15.921
White Ideology	86	3.835	0.648	2.667	5.250

TABLE A3 Summary Statistics: Census of Government Finance and Population

Statistic	N	Mean	St. Dev.	Min	Max
dgepercap_cpi	24,630	1.077	1.277	0.0004	96.668
H_citytract_NHW_i	19,609	0.079	0.110	0.000	0.885
diversityinterp	13,757	0.309	0.188	0.007	0.772
pctblkpopinterp	13,757	0.098	0.151	0.000	0.980
pctasianpopinterp	13,757	0.032	0.054	0.000	0.674
pctlatinopopinterp	13,757	0.103	0.160	0.000	0.987
medinc_cpi	13,757	54,407.450	22,067.850	15,642.800	240,938.000
pctlocalgovworker_100	13,757	3.361	0.950	0.677	8.365
pctrentersinterp	13,757	0.361	0.140	0.014	0.871
pctover65	13,757	0.125	0.050	0.012	0.771
pctcollegegradinterp	13,757	0.160	0.099	0.003	0.587
logpop	24,630	9.937	1.107	4.745	15.921