Investigating the Underprovision of Public Goods: Does Segregation Matter?

Maria Burzillo

5/8/2020

Abstract

Trounstine (2016) suggests that high levels of residential segregation are associated with increased political polarization and decreased public goods spending. In this analysis, I was able to successfully replicate Trounstine (2016)'s main results. I then attempted to better deal with the large amounts of missing data in the datasets used in the original analysis by multiply imputing missing values and re-running the original models using the resulting multiply imputed datasets. My results suggest that segregation is associated with increased racial political polarization, although maybe not as strongly as Trounstine (2016) originally suggested. Furthermore, I find that Trounstine (2016)'s conclusion that increases in segregation are associated with decreases in public spending holds for large cities, but that diversity is a better explanatory factor for small cities.

Introduction

There is a large degree of variation in public goods spending across local governments. As a result, many scholars have worked to determine what factors may lead to the underprovision of public goods. Research in the past has associated racial diversity or changes in levels of diversity with the underprovision of public goods (Baqir, Easterly, and Alesina 1999; Hopkins 2009). However, Trounstine (2016) argues that it is racial segregation, not diversity in and of itself, that results in the underprovision of public goods. Trounstine's analysis consists of two main parts. First, she uses election and demographic data from 25 of America's largest cities between 1990 and 2010 to run a multilevel mixed-effects linear regression with fixed effects for region and year and with random effects for cities in order to show that political polarization increases with 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 an entire city. The second main part of Trounstine's analysis looks at the ability of the Theil's H Segregation Index to explain a variety of types of public expenditures by city using a sample of 2,637 cities with 13,742 city-year observations. Using linear regressions with fixed effects for cities and robust standard errors clustered by city, Trounstine finds a significant, negative relationship between segregation and public goods spending that is robust to the inclusion of a variety of relevant controls and an alternative specification in which the number of waterways is used as an instrument for segregation.

In this analysis, I first work to replicate the main results of Trounstine (2016) using R statistical software (R Core Team 2019). The original data and Stata code made publically available by the author were downloaded via the Harvard Dataverse (Trounstine 2015). I also make all of my code and analysis available on Github.[1] I was successfully able to replicate the main results of Trounstine (2016) in R with the exception of some of the marginal analyses, which I was nevertheless able to replicate in Stata.

One concern with Trounstine (2016) is the exclusion of a large proportion of observations from the regression analysis due to missing data for some values, which substantially constrains the sample size. For example, the regression analysis in the main specification using the the original racial polarization dataset exludes %55 of the observations, and the regression analysis in the main specification using the original financial segregation dataset excludes %95 of observations. As an extension of Trounstine (2016), I impute missing values in the original data using the mice package in R, which generates multivariate imputations using chained equations (van Buuren and Groothuis-Oudshoorn 2011). Then, I use the multiply imputed datasets to re-estimate the original models, pooling the results to generate final pooled regression coefficients and parameters.

Comparing the results of the original regressions and those done with the imputed data yields similar big picture results for the analyses concerning the relationship between political polarization and segregation. Like in Trounstine (2016), I find that segregation is positively associated with political polarization; however,

[1] Link to my Github repository for this project.

the magnitude and significance of the effect has diminsihed slightly. For the second set of analyses conducted with the financial segregation data, my initial results for running the models with the imputed data differ significantly from Trounstine (2016), suggesting that diversity, not segregation, is the more robust explanatory factor in patterns of public goods spending across cities. Nevertheless, this is likely because the data imputation exercise led to the inclusion of a much larger percentage of small cities in the sample, which were perhaps inadvertently largely excluded in Trounstine (2016) due to the list-wise deletion of observations with missing values, which were more frequently found in the observations of small cities. To explore whether or not city size impacted the results, I divided the original sample into two groups, large and small cities, based on the 25,000 population cutoff used to define cities in Baqir, Easterly, and Alesina (1999). Performing the data imputation exercise and re-running the models on these subsets revealed that for large cities, segregation was the most powerful explanatory variable for public goods spending and was associated with decreases in public spending, while for small cities, diversity remained the most powerful explanatory variable and was associated with increases in public spending.

Literature Review

Despite some progress made towards racial equity in the U.S. on other fronts, residential racial segregation continues to be deeply entrenched in American society (Fischer et al. 2004; Oliver 2010; Massey 1993). Research suggests that this kind of segregation has political consequences, as political cleavages in segregated cities tend to have racial as well as spatial dimensions (Massey 1993). Neighborhoods are often important actors within local politics because local governments provide many functions that are allocational in nature and concern geographical space (Trounstine 2016). Thus, when local areas are composed of competing neighborhoods that are divided on racial lines as well as spacial lines, it is natural to expect higher degrees of racial polarization as a result.

In order to understand what segregation is and why it is important as a political concept, it is critical to understand the literature on the effects of diversity at various geographic levels. Studying diversity can be somewhat difficult and confusing because its effects tend to differ by geographic level. On the neighborhood level, a lack of racial diversity has been associated with racial intolerance, resentment, and competition between racial groups (Oliver 2010). Living within non-diverse or segregated neighborhoods has also been associated with holding negative stereotypes and perceptions about out groups (Eric Oliver and Wong 2003). As a result, non-diverese, homogeneous neighborhoods have been associated with increased racial tension and political polarization in comparison to integrated, diverse neighborhoods. However, at the city or metropolitan level, the opposite seems to be true: when considering larger geographic areas, diversity, not homogeneity, is correlated with racial tension, competition, prejudice, lower levels of cooperation, and lower spending on public goods (Oliver 2010; Baqir, Easterly, and Alesina 1999; Hopkins 2009). Considering these theories jointly would predict that the geographic areas most ripe for racial tension and competition will be those that are diverse overall, but have many homogeneous, non-diverse neighborhoods, meaning that while people of different races co-exist within a larger metropolitan area, they live separately within their own neighborhoods. This situation corresponds to Trounstine (2016)'s definition of a severely segregated city. The differing effects of diversity at the neighborhood versus the larger geographic level provide an argument for considering segregation as opposed to simply diversity in local politics because measures of segregation (such as the Theils's H Index used by Trounstine (2016)) take into account the patterns of racial mixing at both the neighborhood and larger geographic level that, alltogether, are important in predicting whether a locality will be characterized by interracial cooperation or political polarization (Trounstine 2016; Oliver 2010; Bharathi et al. 2018).

Political polarization along racial lines may lead to decreased public spending and goods provision because groups may have different preferences, which can make compromise hard, and because groups may preceive a disutility in out-groups' receiving public goods expenditure (Baqir, Easterly, and Alesina 1999). Einstein (2012) found evidence that racial segregation predicts large political divisions at the metropolitan level and that these divisions can create a lack of willingness to compromise and collaborate on local policy problems. Trounstine (2016) finds similar results at the city level: that residential racial segregation is associated with both increased political division and decreased public spending. Thus, these authors suggest that it is the combination of homogenous neighborhoods within a much larger, diverse geographic area that can lead to increased political polarization and reduced public goods spending in local governments.

More recently, some scholars have called this hypothesis and its importance into question. For example, Lee (2018) finds evidence that larger inequalities within the political system favoring socially powerful groups, not local diversity patterns leading to decreased cooperation, may be a better explanation of failures in public goods provision in diverse areas. Others suggest that additional factors, such as income segregation, may be important confounding factors in public goods provision (An, Levy, and Hero 2018; Massey 1993). An, Levy, and Hero (2018), for example, suggests that the more closely related income inequality is to racial inequality, the less investment is made in public goods, and that this interaction may be a better predictor of public goods spending patterns than measures of diversity. Given the active debate in the literature over the relationships between diversity, segregation, public goods spending, and other factors, it is increasingly important to re-examine previously reported findings as a means of robustness checks.

Replication

In this analysis, I was able to successfully replicate all of the main results from Trounstine (2016). The regression analysis was fully replicated in R. As an example, in Table 5, which is presented in the Appendix, I replicate Table 1 from page 713 of Trounstine (2016), which is also included as Figure 7 for comparison. I was unable to successfully replicate Trounstine (2016)'s marginal effects analyses and margins plots using R. There does not yet appear to be a publically available R package or function to calculate marginal or predicted effects or to generate margins plots from the complicated multi-level models employed in the original paper, and creating such functions was outside of the scope of this project. Nevertheless, these results were successfully replicated in Stata.

There was one interesting outcome from my attempt to replicate the original Stata code in R. Due to the differences in R and Stata in dealing with missing values, the results of my first replication of the analyses using the financial segregation dataset were slightly different from the results in Stata presented in the original paper. Ultimately, I discovered that this discrepancy was due to differences in R and Stata in dealing with missing values during conditioning. In order to exclude cities from her analysis with only one census tract, Trounstine conditions her regressions in Stata such that the value for the number of census tracts is greater than one. In Stata, this does not remove missing values, whereas in R, it does. Since this variable is used only as a conditional filter and not as a regression variable (and thus, the observations with missing values for the number of census tracts are not dropped when running the regressions), Trounstine's analysis includes 14 cities and 58 observations with missing values for census tracts in addition to cities with two or more census tracts. This is a potential oversight on the part of the author, and I would suggest also dropping observations with missing census tract data, or else imputing them given that the aim of her condition was to exclude cities without a sufficient number of census tracts. Nevertheless, when I tried dropping these values from the analysis as a robustness check, I found that the change in the sample led to negligible differences in the results, likely because these observations represent a very small proportion of the overall sample.

Extension

One concern with the original analysis in Trounstine (2016) is the large amount of missing data values in the original datasets. In general, because both R and Stata by default drop any observations with missing values for any of the variables used in a regression, the existence of missing values can lead to the exlcusion of a large portion of the data from the analysis and potentially bias the results if the data is not missing completely at random. A large amount of data is missing in both of the main datasets used in the analyses of Trounstine (2016). The main model specification using the racial polarization dataset exludes 112 of the original 203 observations, or %55 of the data. Additionally, the main model specification using the financial segregation dataset excludes 288,291 of the original 302,033 observations, or %95 of observations from the original dataset.

In an attempt to better deal with the problem of missing data in Trounstine (2016), I impute missing values in the original data using the mice package in R, which generates multivariate imputations using chained equations (van Buuren and Groothuis-Oudshoorn 2011). While there are a variety of different imputation methods that could have been employed, multiple imputation (such as the multiple multivariate imputations generated by mice) is desirable because instead of inputting a single value such as the mean for missing values, it instead uses the distribution of the available data to estimate multiple potential values for the missing values. As a result, multiple imputation helps to account for the uncertainty inherent in the imputation process and allows for the calculation of standard errors around estimators. This allows the researcher to more accurately assess the uncertainty of the analysis in general.

Before performing the multiple imputations on the datasets, I first examined the missing data for patterns. To better understand any potential patterns in the missing data, I plotted the pattern of missingness and created a histogram showing the frequency of missing values for those variables with missing values in Figure 1 for the racial polarization data and in Figure 2 for the financial segregation data. Looking at Figure 1, the histogram shows that the dependent variable in the analysis, which denotes the largest vote split along racial lines, and the variable indicating whether or not the election had candidates of more than one race had by far the largest percentages of missing data, missing %54 and %51 of observations, respectively. From the plot on the right in Figure 1, we can see that approximately 42% of observations are complete, meaning they have no missing values. There seems to be a correspondence between missing a value for the largest vote split and missing a value for the biracial election indicator. There also seem to be about 2% of values for which most of the variables are missing. However, most observations are not missing more than 2-3 values.

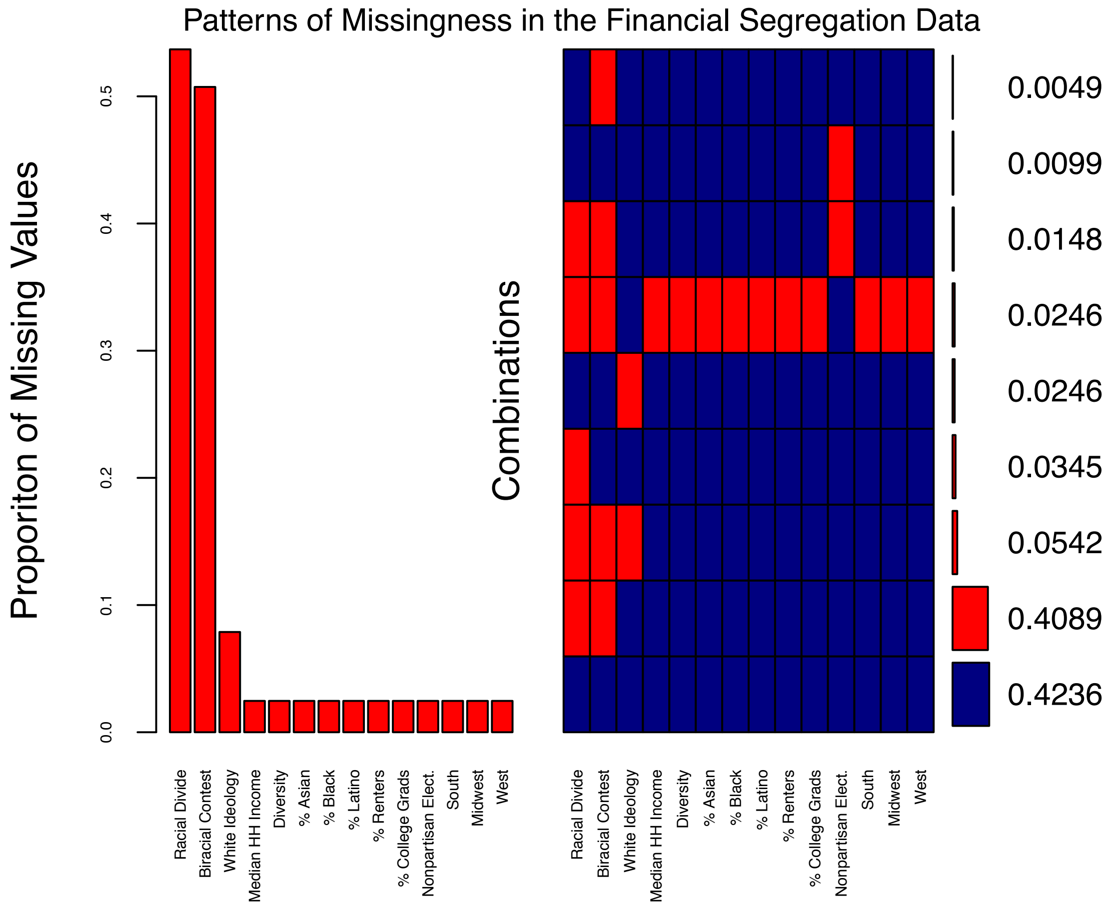


Figure 1: This plot examines the pattern of missing data in the the original racial polarization dataset from Trounstine (2016). The plot on the left shows the percentage of missing data for each variable with missing data and reveals that the dependent variable, which denotes the largest vote split along racial lines, and the indicator for whether or not the election had candidates of more than one race were missing most frequently. The plot on the right shows the pattern of missingness among the variables. We can see that approximately 42% of observations are complete (have no missing values). There seems to be a correspondence between missing a value for biggest split and missing a value for biracial. There also seem to be about 2% of values for which most of the variables are missing. However, most observations are not missing more than 2-3 values.

In Figure 2, we can similarly observe the trends for the financial segregation dataset. In this dataset, there is a very high proportion of missing values for most of the variables. For example, there are 282,334 missing values for the segregation index, which is the main independent variable, or %93 of the data. In this dataset, there are also many more observations for which values are missing for multiple variables in comparison to the racial polarization dataset, as can be seen in the plot on the right in Figure 2. In total, a mere %0.58 of observations are complete for all variables included in the main specification for the financial segregation dataset.

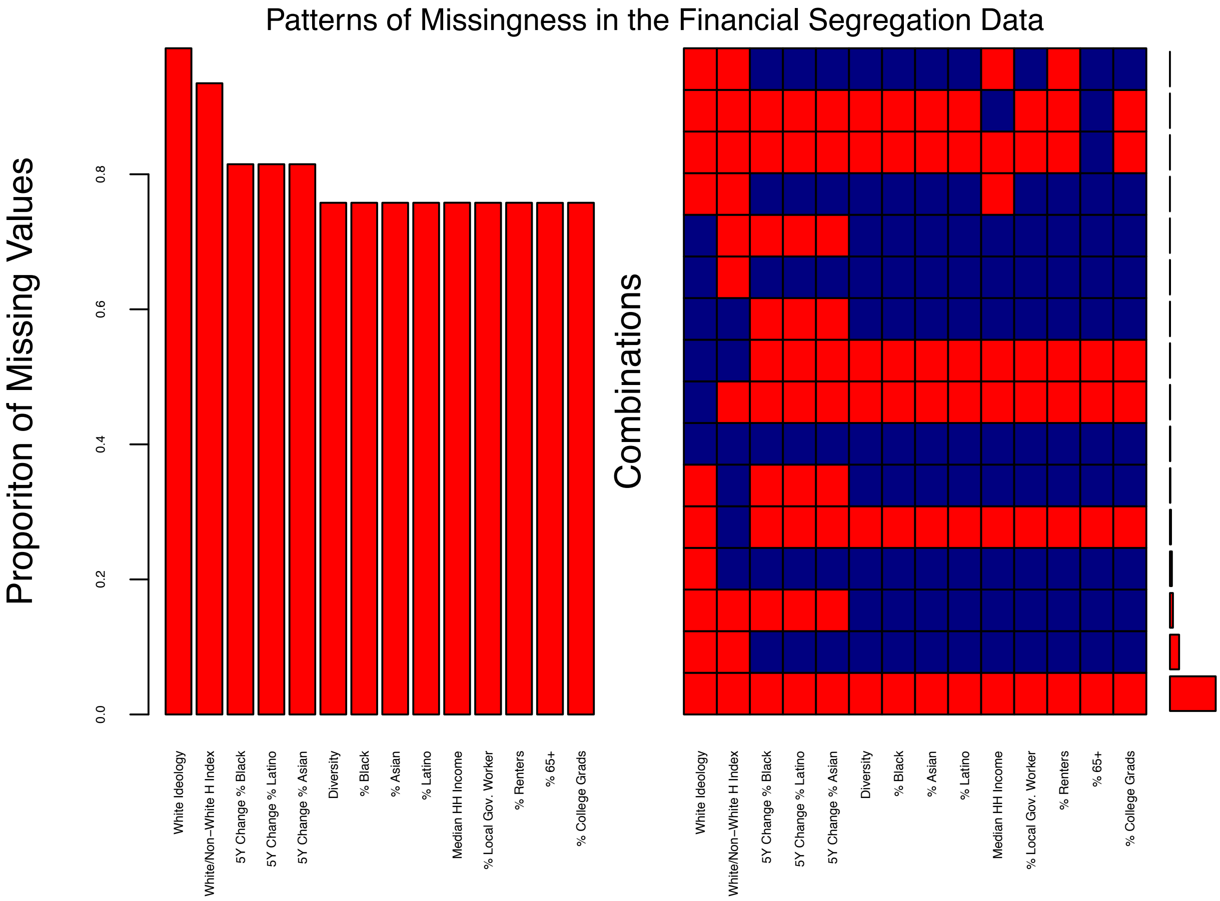


Figure 2: This plot examines the pattern of missing data in the the original financial segregation dataset from Trounstine (2016). The plot on the left shows the percentage of missing data for each variable with missing data and reveals that a large number of variables are missing the vast majority of data. The plot on the right shows the pattern of missingness among the variables. Only .58% of observations are complete (have no missing values), and a large majority of the observations are missing data for many variables.

Finding no clear patterns in the missingness of the data initially, I proceeded with the multiple imputations. For the racial polarization dataset, I performed multiple imputations with 20 iterations using mice(), while for the financial segregation dataset, I performed 5 iterations due to the large size of the dataset and computing limitations. A non-stochastic imputation method, Classification and Regression Trees (CART), was used instead of the default for the imputation for both datasets because of an error with matrix inversion caused by the data that prevented the use of the default method, Predictive Mean Matching. For the racial polarization dataset, I included all of the variables used in the analysis associated with the dataset and was able to impute all missing values. For the financial segregation dataset, however, I only imputed values for the main independent variable, the segregation index, and a few other variables, although all variables used in the regression analysis were included in the data subset input into the mice() function to inform the predictions. For those variables from the financial segregation dataset for which I did impute data, I was not able to impute all of the missing values. More of the data could not be imputed because of the very high proportion of missing values, the large size of the dataset in general, and computing and time constraints for this project.

Before examining the results of Trounstine's model using the multiply imputed data, I first ran some diagnostic tests on the imputation results to make sure that everything ran as expected. First, I checked the convergence of the algorithm used within mice() by plotting the trace lines as a function of the number of iterations for each of the variables. Then, I visually inspected the distributions of the imputed data in comparison to the original data with density and strip plots. All of these checks suggested that the imputed values were within a plausible range of the data and that their distribution fit the underlying distribution of the data relatively well. The only cases in which there was some cause for concern were for some of the variables with very few missing values in the racial polarization dataset, such as the indicators for region. However, since there are so few of these values missing in the actual dataset, this was not a major concern. Figure 3 presents the density plots of the imputed values in red overlaid on the density plots of the original variables in blue for the racial polarization data. The code to generate the rest of the plots and results of the diagnostic tests discussed here are included in the Appendix on GitHub.

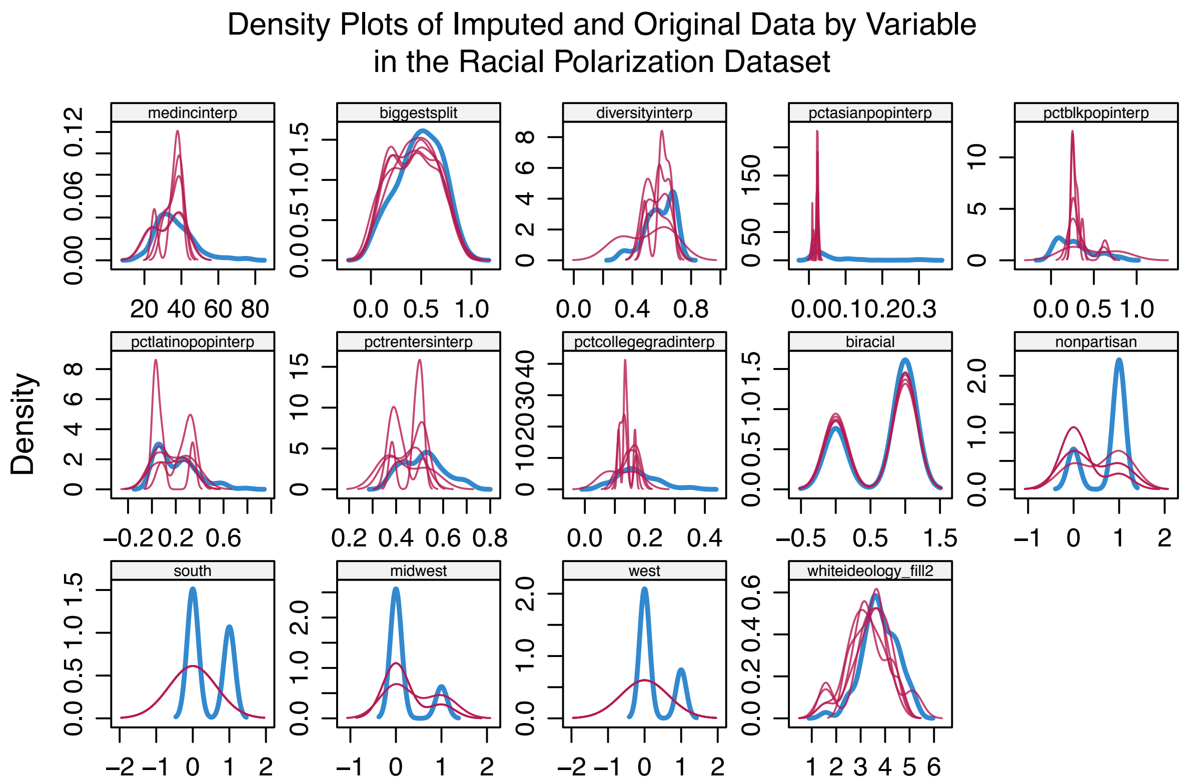


Figure 3: This figure presents the density plots of the imputed data (red) overlaid on the density plot for the pre-existing data (blue) for all those variables with missing values in the racial polarization dataset. For the most part, the imputed data seem to fit the pre-existing data relatively well. The only cases in which there is some cause for concern were for some of the variables with very few missing values in the racial polarization dataset (such as the indicators for region). However, since there are so few of these values missing in the actual datasets, this was not a major concern.

Since the results of the diagnostic checks were promising, I next proceeded to re-estimate the original model using the multiply imputed datasets, pooling the results to produce final pooled regression coefficients and parameters. The results of the analysis of the racial polarization dataset using Trounstine (2016)'s first model specification are presented in Table 1. The results of the model with the new dataset are slightly different compared to the original results from Trounstine (2016). With the imputed data, we now have a total of 203 observations in our model compared to the original model, which had only 91 observations. Interestingly, while the sign of the coefficient on the main variable of interest, the Theil's H Segregation Index, is the same and the standard error has decreased slightly, the result has become statisically insignificant and the 95% confidence interval contains zero. Thus while these results still suggest that segregation may be associated with increased political polarization, they confer a lesser degree of certainty than Trounstine's original analysis. All coefficients are statistically insignificant except for the indicator for a biracial election, which remains significant, and the indicator for a primary election, which has changed from being significant in the original analysis to being insignificant here. In general, the standard errors on the coefficients have decreased slightly.

Table 1: **Racial Polarization in Segregated Cities using the Multiply Imputed Data: Racial Divide with Multigroup Segregation Index**

Variable	**Estimate**	**Std. Error**	**P-Value**
Intercept	0.103	0.494	0.838
Multigroup Seg. Index	0.453	0.349	0.206
Diversity	0.124	0.312	0.694
% Asian	-0.074	0.445	0.869
% Black	0.208	0.193	0.288
% Latino	0.163	0.204	0.432
Median Income (1000s)	0.002	0.004	0.718
% Renters	-0.419	0.297	0.166
% College Grads	0.308	0.612	0.622
Biracial Elec.	0.188	0.039	0.001
Nonpartisan Elec.	-0.065	0.057	0.265
Primary Elec.	-0.049	0.028	0.094
Pop (logged)	0.010	0.042	0.820

* N = 203

Table 1:
This table shows the results of Trounstine's first model for the effect of segregation on political polarization using the multi-group Theil's H Segregation Index and the new, multiply imputed datasets. The sign of the coefficient on the main variable of interest, the segregation index, is positive as in Trounstine (2016), but the result has become statisically insignificant with the 95% confidence interval containing zero. Thus while these results still suggest that segregation may be associated with increased political polarization, they confer a lesser degree of certainty than Trounstine's original analysis. In general, the standard errors on the coefficients have decreased slightly.

For the second and third model specifications, which use the two-group Theil's H Segregation Index, the results were similar. The main coefficient on the segregation index decreased in magnitude, although it remained statistically significant. The other coefficients except for the biracial indicator are insignificant, and standard errors have reduced slightly in these specifications compared to the original analysis in Trounstine (2016). As a visual comparison of the model coefficients, I include coefficient plots for the diversity and segregation index variables as Figure 4. From this plot, it is clear that across all of the models, the coefficient on diversity is positive and insignificant, while the coefficient on the segregaiton index is positive and, in most cases, significant.

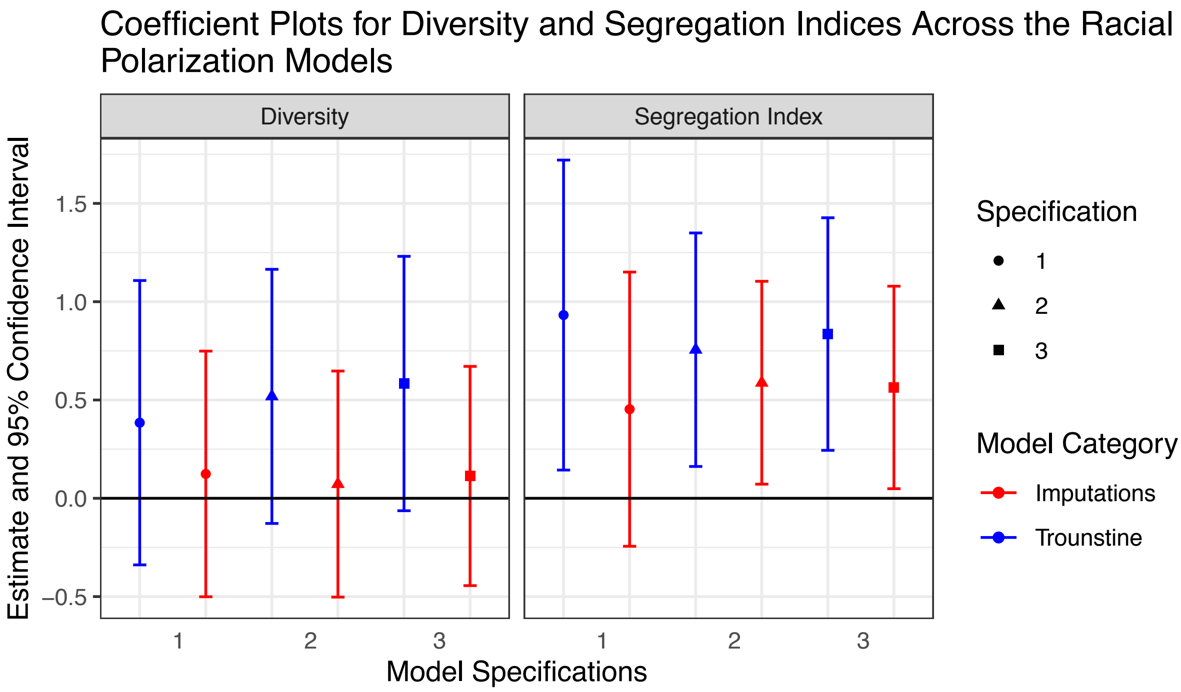


Figure 4: This plot demonstrates how the magnitude and significance of the coefficients on diversity and the segregation index vary among the models. Across all of the models, the coefficient on diversity is positive and insignificant. The coefficient on the segregation index is also consistently positive. In all 3 of the model specifications from Trounstine (2016) and in 2 of the 3 specifications using the multiply imputed data, the coefficient on the segregation index is significant.

For the financial segregation dataset, I similarly fit each of our 5 imputed datasets to the original models from Trounstine (2016) and then pooled the results for each. As a result of the imputation process, I was able to include an additional 59,377 observations in the analysis, increasing the sample size by %432. While this is an improvement, I was unable to impute all of the missing values because of the large percentage of missing values in the original dataset, which limited the performance of the algorithm, and due to time and computing constraints for this project.

The results of the first specification of model using the multiply imputed financial segregation data present a stark difference from the results of Trounstine (2016). Crucially, the effect size of the segregation index on public spending has essentially gone to zero and become statistically insignificant. Interestingly, the coefficient on the diversity variable has increased substantially from .106 to 0.76 and has become statistically significant. The results of the second and third specifications were similar, with the coefficient on segregation going almost to zero and becoming insignificant on both. For brevity, the results of the second and third specifications are not presented, but they are available in the Appendix on GitHub.

Because these results differed to such a degree from Trounstine (2016)'s original findings, I performed some checks on the data included in each analysis. I discovered that Trounstine (2016) excluded most cities below a certain population size because of the nature of the paper's list-wise deletion of observations containing missing values: as small cities were more likely to be missing data, it seems that they were inadvertently and

Table 2: **Effect of Segregation on Overall per Capita Direct General City Expenditures: Imputed Data for all Cities**

Variable	Estimate	Std. Error	P-Value
White/Nonwhite Seg. Index	-0.011	0.427	0.980
Diversity	0.760	0.294	0.010
% Black	0.473	0.350	0.177
% Asian	0.969	1.263	0.443
% Latino	1.574	0.440	0.000
Median Income (1000s)	0.006	0.007	0.399
% Local Gov. Worker	0.032	0.036	0.375
% Renters	0.094	0.336	0.779
% 65+	0.615	0.775	0.427
% College Grads	3.129	0.786	0.000
Pop (logged)	-0.588	0.147	0.000

* N = 73,119

Table 2:

This table shows the results of Trounstine's first model for the effect of segregation on per capita direct public expenditure using the two-group Theil's H Segregation Index and the new, multiply imputed datasets. The results present a stark contrast to Trounstine's original results: the main coefficient on segregation index has decreased to essentially zero and has become statistically insignificant. The coefficient on diversity has increased in magnitude and become statistically significant.

disproportionately dropped from Trounstine (2016)'s regression analysis. Given this development, I decided to test whether or not city population size affected the results. Making use of the population cutoff of at least 25,000 people used to define cities in Baqir, Easterly, and Alesina (1999), I divided the orginal sample into cities having populations above and below 25,000 people. I then again imputed missing values for each and used the multiply imputed datasets in re-runs of Trounstine's original models.

Before discussing the results of these two additional analyses, I first illustrate the differences in the city population sizes for the observations used in the various analyses in Figure 5, which displays the population distributions of the samples used in each. The figure makes it clear that the observations included in Trounstine (2016)'s original analysis were not representative of the original data sample with the missing values, as most of the cities with lower populations were excluded as a result of the list-wise data deletion. While the dataset generated from the data imputation on the whole sample also has a different distribution for city populaiton, the range of the data seems to be much more representative of the overall data range in comparison to the Trounstine (2016) sample. In the plot, I also include the distributions of city population size of the multiply imputed datasets above and below the 25,000 population cutoff. For cities above the cutoff, the city population distribution is similar to the distribution in Trounstine (2016).

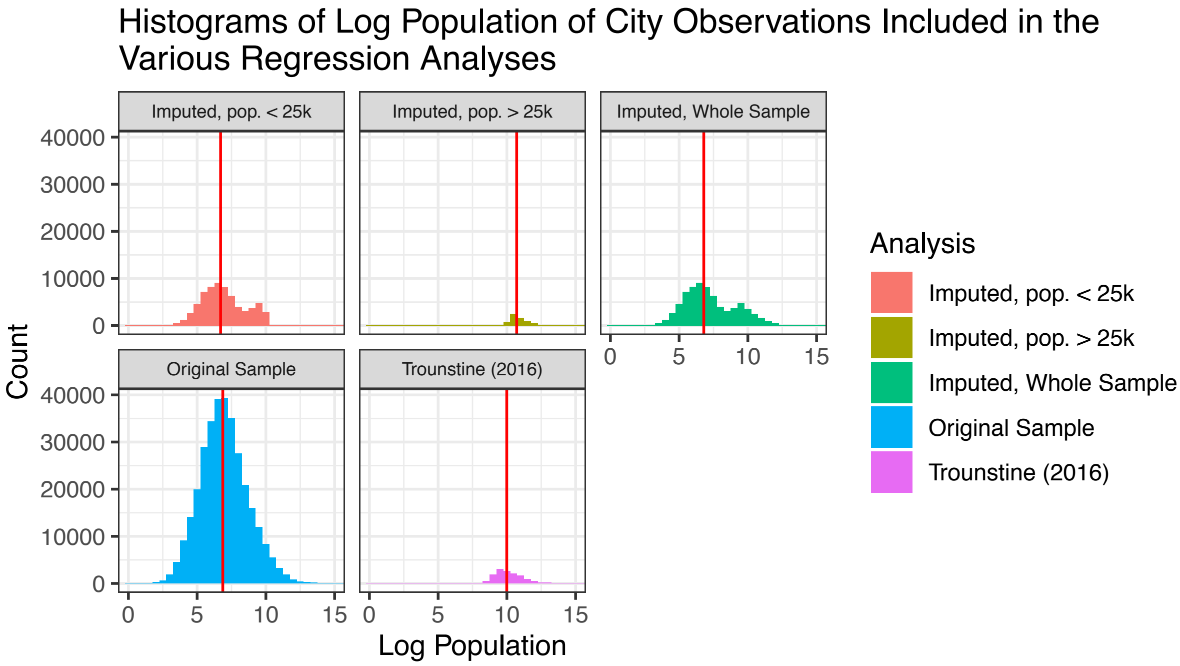


Figure 5: This figure makes clear the differences in city size (in terms of population) among the different analyses. It demonstrates how the list–wise deletion of observations with missing values resulted in a sample that almost entirely consisted of larger cities in Trounstine (2016). The distribution of city population in Trounstine (2016) more closely resembles the distribution of the original sample with a greater than 25,000 population cutoff imposed than the distribution of the original sample. The red lines indicate the median log population in each sample.

The results of the models run with the multiply imputed datasets above and below the cutoff are presented in the following two tables. In Table 3, the results for cities above the 25,000 person population cutoff are displayed. As in Trounstine (2016), the coefficient on the segregation index is negative and highly significant while the coefficient on diversity is insignificant and the magnitude of the coefficient is close to zero. In Table 4, the results for cities below the 25,000 person populaiton cutoff are displayed. Here, the coefficient on diversity is positive and significant while the coefficient on the segregation index is much smaller and insignificant.

As a summary of all of the model results for the two main coefficients of interest, segregation index and diversity, a coefficient plot is presented as Figure 6, with the models ordered by increasing median log population size of the sample on the x-axis. Looking first at the coefficients for diversity, the two models with low median city population size have significant, positive coefficients, while the models with the larger median city populations have insignificant coefficients. This suggestst that diversity is an important estimator of public goods spending for small, but not necessarily for large, cities. Next, looking at the plot for the segregation index, we see the oppposite effect. For the models where the median city population sizes were small, the coefficient on segregation is insignificant. However, for the two models with larger median city populations, the coefficients are significant and negative. This suggests that for small cities, segregation is not a very good estimator of public goods spending, but when cities are larger, increases in a city's segregation

Table 3: **Effect of Segregation on Overall per Capita Direct General City Expenditures: Imputed Data for Cities with Populations of 25,000+**

Variable	Estimate	Std. Error	P-Value
White/Nonwhite Seg. Index	-0.889	0.238	0.000
Diversity	-0.025	0.175	0.885
% Black	0.800	0.311	0.010
% Asian	-0.376	0.383	0.327
% Latino	1.963	0.227	0.000
Median Income (1000s)	0.004	0.002	0.028
% Local Gov. Worker	-0.006	0.026	0.810
% Renters	-0.132	0.443	0.767
% 65+	-0.830	0.484	0.086
% College Grads	5.834	0.616	0.000
Pop (logged)	-0.376	0.070	0.000

* N = 6,716

Table 3:

This table shows the results of Trounstine's model examining the effect of segregation on overall per capita direct general expenditure using the multiply imputed dataset for cities with populations above 25,000 people. As in Trounstine (2016), the effect of segregation on public goods spending is large, negative, and significant while the effect of diversity is small in magnitude and insignificant.

Table 4: **Effect of Segregation on Overall per Capita Direct General City Expenditures: Imputed Data for Cities with Populations of Less Than 25,000**

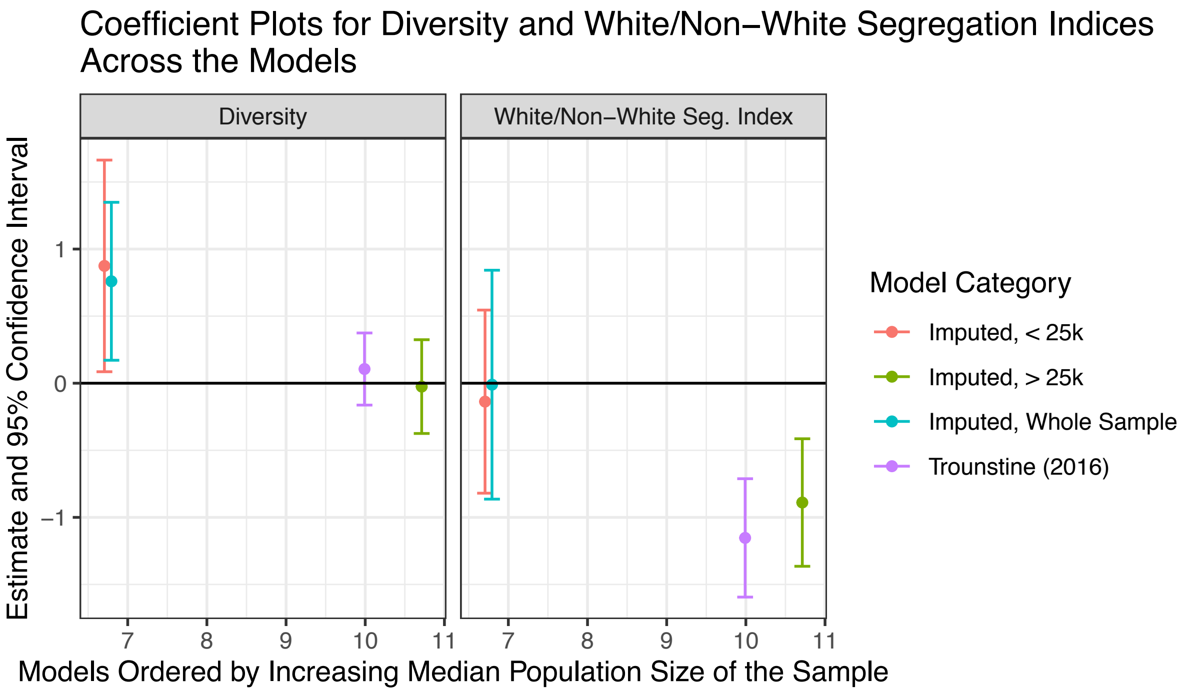
Variable	Estimate	Std. Error	P-Value
White/Nonwhite Seg. Index	-0.137	0.341	0.698
Diversity	0.874	0.394	0.027
% Black	0.350	0.397	0.377
% Asian	1.021	2.586	0.693
% Latino	1.253	0.606	0.039
Median Income (1000s)	0.005	0.007	0.450
% Local Gov. Worker	0.034	0.036	0.348
% Renters	0.106	0.350	0.763
% 65+	0.632	0.813	0.437
% College Grads	2.748	0.862	0.001
Pop (logged)	-0.659	0.170	0.000

* N = 66,403

Table 4:

This table shows the results of Trounstine's model examining the effect of segregation on overal per capita direct general expenditure using the multiply imputed dataset for cities with populations below 25,000 people. In contrast to Trounstine (2016) and the results of our analysis with the imputed datasets for cities over 25,000 people, the effect of segregation on public goods spending is negligible and insignificant, while the effect of diversity is large, positive, and significant.

index are associated with decreases in public goods spending.



This plot demonstrates how the significance of diversity and segregation in predicting public goods spending varies with city size. As city population size increases, diversity tends to have less explanatory power for public goods spending while segregation has more. The opposite is true when moving in the opposite direction: as city population size decreases, segregation becomes less important as diversity becomes increasingly important in explaining public goods spending.

Conclusion

Why cities differ in their amount of spending on public goods is an interesting question that is frequently debated in the literature. Trounstine (2016) posits that segregation may be an important underlying cause of the underprovsion of public goods. Finding evidence that residential segregation by race increases political partisanship, Trounstine (2016) argues that residential segregation can increase division and make cooperation amongst competing groups difficult, which results in decreased spending on public goods. The findings of Trounstine (2016) are particularly important because they suggest that segregation, as defined by many homogeneous neighborhoods within a larger, diverse geographical area, is the key factor in disinvestment in public goods. This contradicts previous research which has suggested that levels of diversity were instead most important (see for example Baqir, Easterly, and Alesina (1999) or Hopkins (2009)).

In this paper, I first worked to provide a check on the analysis presented in Trounstine (2016) by attempting to replicate the work in R. I was able to successfully replicate all of the main results, with the exception of the marginal analyses of the complex, multi-level models which, to my knowledge, are currently not supported by available R packages or functions. I was, however, able to successfully replicate these results in Stata.

I then tested the models in Trounstine (2016) by re-running them with additional data that I imputed from missing values in the original datasets. The large number of missing values in both of the datasets used in Trounstine (2016)'s main analysis created a concern for potential bias in the original results because observations with missing values were automatically excluded from the regression analysis. Using the mice package in R, I created multiply imputed datasets, upon which I re-ran the models and pooled the results to generate the final model coefficients and parameters.

The results of the data imputation exercise differed between the two datasets. For the analyses using the smaller, racial polarization dataset, re-running the models with the imputed data resulted in a smaller magnitude of the coefficient on the segregation index, and also made it statistically insignificant in the main model specification. This suggests that segregation is still positively associated with political polarization, but maybe not quite as strongly as Trounstine (2016) suggests.

On the other hand, for the analyses using the much larger financial segregation dataset, the results of the models using the multiply imputed data sharply contrasted the results of Trounstine (2016), suggesting that diversity, and not segregation, was the best explanatory variable for public goods spending across all cities. While these results originally seemed quite contradictory to Trounstine (2016), I was able reconcile this finding by discovering that the results of Trounstine (2016) seem to be representative only of larger cities. By dropping all observations in the original data with missing values for variables included in the regressions, it seems that Trounstine (2016) inadvertently restricted their sample to almost entirely larger cities (above 25,000 residents) whereas the full sample consisted of cities that had smaller populations on average. To test the hypothesis that diversity is more important for predicting public goods spending for small cities while segregation is more important for predicting public goods spending for larger cities, I divided the original dataset into two subsets of cities with populations above and below 25,000 people. I then multiply imputed the missing values for each data subset and re-ran the models. I found that the results presented initial evidence in support of my hypothesis. For the cities below the 25,000 person population cutoff, the coefficient for diversity was positive and significant while the coefficient for the segregation index was insignificant. For cities above the 25,000 person cutoff, the results were very similar to the results found in Trounstine (2016): the coefficient for diversity was insignificant, but the coefficient for the segregation index was significant and negative.

The replication exercise and the robustness check of Trounstine (2016)'s models using multiply imputed data provide evidence that residential racial segregation is associated with political polarization. Furthermore, I find evidence that supports the finding of Trounstine (2016) that increases in segregation are associated with decreases in public goods spending; however, I find that this is only the case for larger cities. For small cities, I find that diversity, not segregation, seems to be the most important explanatory variable for public goods spending. One potential explanation for why segregation does not seem to meaningfully explain public expenditures for small cities is that, when cities become too small, segregation becomes difficult to maintain.

The sort of segregation by neighborhood described in Trounstine (2016) as a diverse city broken down into homogenous, segregated neighborhoods requires that multiple neighborhoods exist and have the space to be separated geographically. When city population sizes are small, this may not be feasible, and interracial interactions may become more common. As a result, an increase in diversity may be associated with a larger degree of integration and increased interracial contact within a small city, just as the literature expects would happen on the neighborhood level (Oliver 2010). This explanation makes sense in the context of the fact that the coefficient on diversity was positive for small cities in my model, indicating that an increase in diversity in small cities was associated with an increase in public spending, perhaps due to this absence of racial polarization. Given the very small size of most "cities" in the original datasets in Trounstine (2016), I might even argue that these political entities would be more productively considered as "towns" or "neighborhoods," given that the overall median city population in the financial segregation dataset was a mere 958.

The underprovision of public goods may inhibit social mobility and unfairly disadvantage certain groups. Trounstine (2016) has identified an important relationship between segregation and the provision of public goods, which I suggest holds for larger cities, but not necessarily for smaller ones. Given the evidence presented here that the importance of segregation and diversity as predictors of the underprovision of public goods may differ based on city size, further research is needed to assess the nuance of these relationships. For example, at what city size or population does segregation become more relevant than diversity? And what implications does this have for how public policy should be designed and implemented to address these issues? In answering some of these questions, we may be able to improve our ability to identify, understand, and address the root causes of the underprovision of public goods.

Appendix

Table 5: **Racial Polarization in Segregated Cities**

	Racial Divide with:		
	Multigroup H Index	White/Nonwhite H Index	Ideology Controls
	(1)	(2)	(3)
Multigroup H Index	0.932^{**} (0.394)		
White/Nonwhite H Index		0.756^{**} (0.297)	0.835^{***} (0.296)
Diversity	0.385 (0.362)	0.518 (0.323)	0.584^{*} (0.323)
Percent Asian	−0.115 (0.527)	0.120 (0.558)	−0.004 (0.522)
Percent Black	−0.432 (0.269)	−0.237 (0.216)	−0.133 (0.212)
Percent Latino	−0.191 (0.257)	−0.059 (0.254)	0.095 (0.278)
Medain HH Income (1000s)	−0.004 (0.007)	−0.007 (0.007)	−0.002 (0.006)
Percent Renters	−0.580 (0.422)	-0.806^{*} (0.431)	−0.419 (0.454)
Percent College Degree	0.328 (0.711)	0.723 (0.729)	0.123 (0.869)
Biracial Contest	0.210^{***} (0.037)	0.208^{***} (0.037)	0.192^{***} (0.036)
Nonpartisan Election	−0.090 (0.066)	−0.089 (0.066)	−0.034 (0.065)
Primary Election	-0.092^{***} (0.032)	-0.090^{***} (0.032)	-0.071^{**} (0.030)
Population (logged)	0.035 (0.055)	0.048 (0.055)	−0.011 (0.061)
White Ideology			−0.051 (0.032)
Constant	−0.242 (0.569)	−0.393 (0.563)	0.236 (0.605)
N	91	91	86
Log Likelihood	62.774	63.190	67.461
AIC	−55.548	−56.381	−64.922
BIC	32.332	31.499	20.981

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 5: This table displays the results of the replication effort of the analyses conducted in Trounstine (2016) using the racial polarization dataset. The replication code can be found on my personal Github repository, which is linked in the introduction. The original table is displayed for comparison as Figure 7.

Figure 7

TABLE 1 Racial Polarization in Segregated Cities

	Racial Divide with Multigroup Segregation Index			Racial Divide with Two-Group Segregation Index			Racial Divide with Ideology Control		
	β	SE	P > |t|	β	SE	P > |t|	β	SE	P > |t|
Multigroup *H* Index	0.932	0.39	0.02						
White/Nonwhite *H* Index				0.756	0.30	0.01	0.835	0.30	0.01
Diversity	0.385	0.36	0.29	0.518	0.32	0.11	0.584	0.32	0.07
% Asian	–0.115	0.53	0.83	0.120	0.56	0.83	–0.004	0.52	0.99
% Black	–0.432	0.27	0.11	–0.237	0.22	0.27	–0.133	0.21	0.53
% Latino	–0.191	0.26	0.46	–0.059	0.25	0.82	0.095	0.28	0.73
Median HH Income (1,000s)	–0.004	0.00	0.52	–0.007	0.00	0.32	–0.002	0.00	0.81
% Renters	–0.580	0.42	0.17	–0.806	0.43	0.06	–0.419	0.45	0.36
% College Degree	0.328	0.71	0.65	0.723	0.73	0.32	0.123	0.87	0.89
Biracial Contest	0.210	0.04	0.00	0.208	0.04	0.00	0.192	0.04	0.00
Nonpartisan Election	–0.090	0.07	0.18	–0.089	0.07	0.18	–0.034	0.06	0.60
Primary Election	–0.092	0.03	0.00	–0.09	0.03	0.01	–0.071	0.03	0.02
Population (logged)	0.035	0.06	0.53	0.048	0.05	0.38	–0.011	0.06	0.86
White Ideology							–0.051	0.03	0.11
Constant	–0.242	0.57	0.67	–0.393	0.56	0.49	0.236	0.61	0.70
Wald χ^2	187.12		0.00	189.68		0.00	222.92		0.00
N		91			91			86	

Note: Multilevel mixed-effects linear regressions with fixed effects for region and year, and random effects for cities are presented.

Figure 4: This is the original table displaying the results of the models for the effects of residential segregation on political polarization from page 713 of Trounstine (2016). My replication of these results can be found in Table 5 in the Appendix of this paper.

Bibliography

An, Brian, Morris Levy, and Rodney Hero. 2018. "It's Not Just Welfare: Racial Inequality and the Local Provision of Public Goods in the United States." *Urban Affairs Review* 54 (5): 833–65. https://doi.org/10.1177/1078087417752476.

Baqir, Reza, William Easterly, and Alberto Alesina. 1999. "Public Goods and Ethnic Divisions" 114 (4): 1243–84. https://doi.org/10.1162/003355399556269.

Bharathi, N. Manjula, Deepak V. Malghan, Sumit Mishra, and Andaleeb Rahman. 2018. "Public Goods, and Nested Subnational Units: Diversity, Segregation, or Hierarchy?" In. https://doi.org/10.2139/ssrn.3254095.

Einstein, Katherine Levine (Ph D. 2012. "Divided Regions: Race, Political Segregation, and the Fragmentation of American Metropolitan Policy." PhD thesis.

Eric Oliver, J., and Janelle Wong. 2003. "Intergroup Prejudice in Multiethnic Settings." *American Journal of Political Science* 47 (4): 567–82. https://doi.org/10.1111/1540-5907.00040.

Fischer, Claude, Gretchen Stockmayer, Jon Stiles, and Michael Hout. 2004. "Distinguishing the Geographic Levels and Social Dimensions of U.S. Metropolitan Segregation, 1960–2000." *Demography* 41 (1): 37–59. https://doi.org/10.1353/dem.2004.0002.

Hopkins, Daniel J. 2009. "The Diversity Discount: When Increasing Ethnic and Racial Diversity Prevents Tax Increases." *The Journal of Politics* 71 (1): 160–77. https://doi.org/10.1017/S0022381608090105.

Lee, Alexander. 2018. "Ethnic Diversity and Ethnic Discrimination: Explaining Local Public Goods Provision." *Comparative Political Studies* 51 (10): 1351–83. https://doi.org/10.1177/0010414017740604.

Massey, Douglas S. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Democracy and Urban Landscapes. Cambridge, Mass.: Harvard University Press. http://nrs.harvard.edu/urn-3:hul.ebookbatch.ACLS_batch:20170426MIU01100000000000000000544.

Oliver, J. Eric. 2010. *The Paradoxes of Integration:Race, Neighborhood, and Civic Life in Multiethnic America*. Chicago ; London: University of Chicago Press.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Trounstine, Jessica. 2015. "Replication Data for: Segregation and Inequality in Public Goods." Harvard Dataverse. https://doi.org/10.7910/DVN/4LZXTY.

———. 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.

van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2011. "mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45 (3): 1–67. https://www.jstatsoft.org/v45/i03/.