This is a draft version. A final publication version will be available forthcoming in the Annals of Regional Science, 2024. Please reach out to the author for additional details.

AN UNEVEN LANDSCAPE OF PUBLIC SERVICES FOR PEOPLE OF COLOR: IDENTIFYING ENDOGENEITY IN THE RELATIONSHIP BETWEEN LOCAL RACE COMPOSITION AND PUBLIC EXPENDITURE

AUSTIN LANDINI
UNIVERSITY OF MISSOURI

[1]I would like to thank my colleagues, advisors Stephan and Anita and friends for the creative inspiration and my parents Frann and Greg for their continuing love and support. I would also like to thank the conference discussants and editors who have provided feedback on my work.

[2]Contact Austin Landini, University of Missouri-Columbia, 216 Mumford Hall
austin.landini@missouri.edu (319)541-3471

An Uneven Landscape of Public Services for People of Color: Identifying Endogeneity in the Relationship Between Local Race Composition and Public Expenditure

Abstract

Previous economic literature finds that race diversity is associated with increased tax collection and expenditure per capita, but relative under-provision of important public goods and services. Using a panel data set covering Census 'Places' of over 25k population 1980-2010, I verify this conclusion in a cross-section for the 2010 US Census data. I then show that the measure of racial heterogeneity commonly used throughout the literature is endogenous over time due to omitted variables, and potential reverse causality or simultaneity. As a result, cross-section and pooled results are likely to be biased. Using predicted race composition from national trends as an instrumental variable, I show that the relationship between changes in race heterogeneity and both tax and spending per capita is negative. This result suggests a potential reversal of the findings following of the influential paper by Alesina, Baqir and Easterly (1999).

Keywords

Regional Economics

Public Expenditure

Taxation

Geospatial Analysis

Demographics

JEL Classifications

H41

H71

R23

Highlights

- I replicate results from a canonical work of Alesina, Baqir, and Easterly (1999) using most recent available US Census data.
- In a cross-section, it appears that US Census 'Places' of over 25k population tax and spend more as they become more diverse.
- This relationship is likely biased due to omitted variables, and correcting for the bias may reverse the conclusion.
- Instead of taxing and spending more, but not on productive public goods, US Places are taxing and spending relatively less as they diversify.

I Introduction

Local quality of education, health care, and environmental conditions are important quality of life determinants. Previous economic literature identifies inequality in the provision of public goods at the local community level. For example, educational providers are of higher quality in wealthier areas, and of lower quality in predominantly non White areas– especially inner cities (Darling-Hammond 2007; Lleras 2008). Access to quality health care is limited in low income and racially diverse areas. (Fiscella 2000; Kawachi 2005; Williams 2005).

Alesina, Baqir and Easterly (1999) provide a seminal framework to measure the impact of racial heterogeneity on public goods provision. The authors measure race heterogeneity as a Herfindahl index- the probability that two randomly drawn people from a geographic area will be of different races. While this measure of race heterogeneity is positively associated with expenditures and tax collection per capita, it is also negatively associated with spending shares on "productive" public goods and services: education, roads, and sewage/trash pickup.

Following this influential paper, additional literature has emerged attempting to identify mechanisms explaining why racial heterogeneity might reduce provision of public goods. While theories such as preference, technology or communication, and strategy selection have been proposed (Habyarimana 2007), these findings generally rely on the assertion that places which are more racially heterogeneous tax and spend more per capita, but are unable to utilize these resources toward the provision of important public services. The negative relationship between race heterogeneity and spending on public projects has been shown to hold across nations (Alesina, et al. 2003; Baldwin and Huber 2010; Lieberman 2012) and at other levels of geographic analysis including world regions (ex. Banerjee, et al. 2005; Jackson 2013), local districts and municipalities (Bardhan 2000; Habyarimana 2009).

In recent literature, the assertion that race heterogeneity is negatively related to spending shares on public goods been challenged. Since 1980, the US has undergone a considerable demographic change. In US Cities, the Herfindahl

index of race diversity, or the average probability that two randomly selected citizens will be of different races as defined by the US Census, has increased from around 23% in 1980 to 43.5% in 2010. Inequality has been the focus of much of the literature studying the relationship between the population dynamics of an area and the public goods and services that area provides. Corcoran and Evans (2010) find that inequality which reduces the median voter's tax share induces higher educational spending. Likewise, Boustan et al. (2013) show that growing income inequality is associated with the expansion of government revenues and a variety of services. Still, it remains unclear that the expansion of services occurs uniformly over the entirety of the local area. Concerns remain that expansion of services may occur primarily in wealthier areas.

In this regard, there is a positive correlation between changes in income inequality and changes in residental segregation (Fogli 2019). Previous literature studies location decisions among those relocating and finds that "generally, high income or highly educated households, or households belonging to ethnic groups with high levels of ethnic capital, choose to live in less segregated neighborhoods." (Borjas pg. 25). Migrants into cities tend to seek out ethnic enclaves for more favorable job market outcomes (Edin 2003, Damm 2009). For these reasons, race composition in a community as measured by the standard heterogeneity measure is likely to be endogenously determined over time, in part by public goods and services provision. As Benabou notes, "Most cities are segregated along income and occupational lines. People with high skill, high wage jobs live together in certain select areas or suburbs, and those with low skill, low wage jobs, or no job at all reside in different parts" (Benabou pg. 1).

If there are differences in preferences over public goods selection between groups, or if the groups are less successful in collective action to achieve public goods provision, community formation is important in determining these spending outcomes. Low income earners may be drawn to communities which are relatively cheaper and provide less public goods, thereby further diluting the public goods provided, an important concern in the Tiebout Hypothesis. Endogeneity in the determination of race or income composition would bias contemporaneous results, potentially resulting in an overstated impact of heterogeneity on public goods provision in cross sectional analysis. The possibility that demographic distributions within cities might be in part determined over time by public goods provision and previous racial composition remains largely unaddressed (Banerjee, et al. 2005, Singh & vom Hau 2016). Furthermore, it is possible that accounting for endogeneity in the relationship, more diverse cities, especially those with a high degree of segregation, collect less tax and spend less, rather than the opposite (Trounstine 2016).

Boustan, et al., (2013) employ an instrumental variables approach to address endogeneity in the relationship between inequality and public services. The authors implement a shift-share instrument for income inequality based on an initial 1970 value and growth of median income in Census microdata. Their results show that growing inequality is associated with increases in tax revenues and faster growth in public expenditures at municipality and school district levels. Likewise, race heterogeneity may be endogenous since omitted and unobservable features of a residential area are likely to impact demographic change, taxation, and spending over time. I implement a similar instrumental

variables approach for race heterogeneity to study the impact this endogeneity.

This is not the first study to reconsider the findings of Alesina, et al., (1999) and similar work using panel analysis. For example, from three periods of Census data, An, et al., (2018) examine the relationship between the racial structure of local income inequality and government spending on public goods and services. Using a local measure of racial inequality developed by Hero and Levy (2016 and 2017), the authors conclude that racial diversity and inequality are jointly two important influences on policy. While the authors argue that diversity is impacting local public goods spending, their racial diversity measure is found to be statistically significant only in predicting changes in school district level spending, and sanitation service spending.

Both Hopkins (2009) and Lee, et al., (2016) employ panel analyses of US cities to show that the impact of diversity on spending shares of individual public goods varies by type of good/service, rather than being consistent across all types. The focus on spending shares for public goods as cities diversify is important, especially with regard to increases in policing (Hopkins 2009) and decreases in important services such as Medicaid coverage (Lee, 2013), or educational spending (An, et al., 2018) which have become socially and politically salient in recent years. I expand upon these works by highlighting the importance of demographic change over time in its relationship to trends in overall tax and expenditure in an area over time. While the impact on demographic change on individual spending shares may vary, a higher degree of ethnic diversification is statistically associated with reductions in overall expenditure and tax collection over time. These changes may in turn create a feedback effect which impacts the distribution of households who choose to reside in the area. I contribute to the notion that the racial composition of a geographic area may be statistically associated with spending on public goods and services in that area by employing an instrumental variables approach similar to Boustan (2013). In doing so, I am able to provide evidence that the relationship is endogenous. Furthermore, in models such as first difference and instrumental variables which correct for the endogeneity, I show that changes in the racial heterogeneity of a Census Place are statistically associated (negatively) with both tax and expenditure per capita in real dollars.

The results of this paper provide evidence on a broader scale in support of Hopkins (2011) who suggests, using a study of Massachusetts and Texas data, that diversity reduces localities' willingness to raise taxes only when localities are undergoing sudden demographic changes. As a result, instead of understanding diversity as differential preferences, it is important to understand how demographic changes can destabilize residents' expectations and influence local public policy. Hopkins concludes that to understand how diversity influences public good provision, it is best to look at towns that are diversifying, not necessarily the towns that are most diverse. I consider the sample of US Census 'Places' with population 20K or greater to identify national trends in the impact of diversification on local tax and spending over time.

To demonstrate the importance of intertemporal considerations in this relationship, I first replicate the results of Alesina, Baqir and Easterly (1999) with more recent 2010 Census data. I show that in a cross-section, race diversity is

positively associated with tax and expenditure per capita, but negatively associated with spending shares on important public services such as road maintenance and sewage. Then, using a shift-share instrument in a panel of decennial Census data 1980-2010, I argue that the relationship between race heterogeneity and both public tax and spending is endogenous. Analysis correcting for the endogeneity bias suggests that areas which diversify collect relatively less tax and spend less per capita. This result suggests further reconsideration of ethnic conflict model following Alesina, Baqir and Easterly (1999), who argue that "ethnic conflict is an important determinant of local public finances" (Alesina, et al., 1999, pg. 1). Previous literature has critiqued this argument by highlighting inconsistencies in the conclusion that diversity negatively impacts spending shares (Hopkins 2009 and 2011; Lee 2013; Lee, et al., 2016). However, the ethnic conflict argument requires as an assumption that more diverse areas have higher tax collection and expenditure. I do not find this assumption to be true as cities diversify over time, thereby providing another type of critique to a growing literature which reconsiders the ethnic conflict argument.

The remainder of the paper is organized as follows: Section II describes the 1980-2010 panel of US Census Places of over 25K population used for analysis. Section III replicates Alesina, Baquir and Easterly's (1999) results updating from 1990 Census Data to 2010. Section IV introduces the shift-share instrumental variable analysis. Section V discusses robustness checks and data concerns, and Section VI concludes.

II Data

This study builds on the present literature by utilizing data featuring public goods expenditures and race composition at the city level over time, from 1980 to 2010. In addition, I use predicted race composition as an instrument to address endogeneity. The year 1980 is chosen as a starting point due to data quality for measuring racial heterogeneity in cities. Before 1980, Census data on race is limited to "White, Black, and other".

Decennial city data 1980-2000 are taken from 1983, 1994, and 2000 City Data Books (Haines 2010). Since the Census Bureau discontinued the City Data books series following 2007, 2010 data has been scanned using OCR from the 2013 County and City Extra: Annual Metro, City, and County Data Book (Gaquin and Ryan 2013). Data in each from each City Book corresponds to the most recent decennial census demographic information, combined with U.S. Census Survey of Governments as well as labor statistics provided by the Bureau of Labor Statistics. Tables 1 and 2 present a selection of important summary statistics over the 40-year panel.

Given changing demographics, the number of cities over 25k population varies over the years in the sample. There are 1422 cities considered in 2010, 1083 in the 2000 and 1990 sets, and 957 in 1980. To reproduce the Alesina, et al. style contemporaneous analysis, I use the 2010 cross-section. I also conduct panel analysis with the variables available for the complete 1980-2010 sample, for which there are 734 observations in the continental United States. Figure 1 demonstrates the spatial distribution of these panel observations and their racial heterogeneity in the 2010 Census,

while Figure 2 maps the 30-year change in heterogeneity.

City level provision should most directly reflect the needs and preferences of community members, but quality and quantity of public goods and services provided often varies widely within close proximity. Study at the city level over time allows for a detailed analysis of the relationship between community race heterogeneity and public goods provision. The US Census defines race categorizations which may be arbitrary or socially constructed. Nonetheless, these classifications broadly represent groups which are politically salient in the U.S. economy (Alesina 1999). Given proportions of each race in a population, the standard definition of race heterogeneity:

$$h_i = 1 - \sum_i (race_i)^2$$

$$i = [\text{White, Black, Asian, Pacific Islander, Other}]$$

is a common measure used throughout the literature. Due to the introduction of multi-race categories in the 2000 survey, I use only one-race data, representative of 97% of the population in the 2010 Census.[1]. While authors of previous studies argue that the correlation between "Other" and "Hispanic" is high enough to be used as a proxy, the correlation between the two may be highly variable between cities. The race heterogeneity measure tells gives only a general measure of diversity in a community, however, during recent US history, increased diversity has most commonly meant relative decreases White population, and relative increases in non-White populations. For this reason in later sections, I provide additional first difference estimates for White, Black and Hispanic, and consider changes in the group identifying as "White" compared to "non-White".

III Empirical Specification

To test the relationship between race heterogeneity and public expenditure and tax, I will use variations of the basic empirical specification:

$$Spending_i = \alpha_i + \beta_1 h_i + \Theta X_i + \varepsilon_i$$

Using the main outcome variables Tax and Expenditure per capita. Table 1 presents population and demographic summary statistics while Table 2 presents summary statistics for main dependent public finance variables from the 1980-2010 panel. Public goods provision data are taken from the City Data books and supplemented with more detailed Census government finance data 1967-2012 from the 5-year Census of Governments aggregated by Pierson, et al (2015). While government finance data is available for all townships, racial composition data from the census is only widely available for cities over 25000 population.

[1] Hispanic is not mutually exclusive with other race categories. Other and Hispanic have .898 correlation in the 2010 data, and about .870 overall across the panel

X is a vector of controls specific to city i. For each regression, in addition to the measure of race heterogeneity, I include control variables, following as closely to the literature as possible. As controls, I include logged income per capita city population since wealthier, more developed cities may provide more public goods and there scale effects might exist in large cities. I use educational attainment as defined by percent of population over age 25 with at least a college degree since this group is likely to demand better quality public goods. Since inequality could motivate disparity in public goods provision (An, et al 2018), I include inequality as measured by the ratio of per capita to median income[1], with poverty rate as a robustness check. I include the percentage of population over age 65 since the empirical local public good literature has emphasized the role of age structure as a determinant of preferences for public goods. Finally, I include intergovernmental transfers as a measure of state level policy.

At this point, I focus my analysis on the relationship between the race heterogeneity measure chosen, h_i and tax collection and expenditure per capita. A contemporaneous analysis suggests that there is something about race heterogeneity today which impacts tax and spending today. Race should only impact these public measures via its effects on other variables; some of which, like income, are observed, others of which are not. The impact of race on observed and unobserved variables in the model is more likely to be dynamic, meaning a panel analysis may provide more robust conclusions.

To highlight this issue present first difference estimates of the form:

$$\Delta Spending_{it} = \beta \Delta h_{it} + \Theta \Delta X_{it} + \Psi Year_t + \varepsilon_{it}$$

There is also reason to believe that the relationship between Race and Expenditure/Tax at the local public level is endogenous. To address this issue, I create an instrument from predicted Race compositions based on 1980 starting point and national trends. The use of a shift-share approach IV is a common empirical strategy within regional economics. One potential justification for this approach is when base shares are endogenous, but affected by exogenous external shifts (Borusyak 2022).

The approach is a two stage IV specification using the predicted values:

$$h_{it} = \beta_1 Z_{it} + \Theta X_{it} + \Psi Year_t + \varepsilon_{it}$$

$$Spending_{it} = \theta_1 \hat{h_{it}} + \Gamma X_{it} + \varphi Year_t + u_{it}$$

Where the instrument Z_{it} is the local predicted values from the national trend. Using this approach allows for a more formal testing of endogeneity in the relationship between race composition and local public taxation and spending per capita.

This instrumental approach follows from Boustan, et al. (2013) who use the same process to create predicted GINI

[1] Mean income data to calculate mean to median ratio was not available for 1980. Small/large households would inflate/deflate this measure, but the bias should be negligible in large samples.

coefficients at the municipal level. More broadly, Bartik (1991) type instruments popularized by Blanchard and Katz (1992) predict local employment growth rate by interacting local industry employment shares with national industry employment growth rates (Goldsmith-Pinkham, et al. 2020). Bartik-style instruments have been used widely in the Regional Science literature. Gould, et al. (2002) use Bartik shocks to examine the causal effect of income on local crime rates. Additional literature has used these instruments to explore the impacts of income inequality on happiness (Luttmer 2005), and how local labor market shocks impact skill-specific amenities (Diamond 2016). Saiz (2010) and the literature following it estimate housing supply elasticities using local housing market shocks generated using a Bartik-style procedure (Baum-Snow and Ferreira 2015).

Recent literature also highlights potential shortcomings of the shift-share instrument. Goldsmith-Pinkham, et al., (2020) argue that there are situations in which the Bartik type instrument will and will not be valid. In typical specifications, the assumption is that the shares are exogenous to changes in the error term (i.e, changes in the outcome variable), rather than levels of the outcome variable (pg. 14). To assess the plausibility of the Bartik type instrument, Goldsmith-Pinkham et al., suggest first considering correlation between the initial period characteristics and factors which predict changes (and not just levels) of the outcome variables. Neither the initial condition of the race heterogeneity measure, nor any of the individual race shares as of 1980 appear to be highly correlated with 10-year changes in the tax or expenditure per capita. All correlation coefficients fall between 0 and 0.2 in absolute value. The instrument should also satisfy a monotonicity assumption. Of the 734 Census Places studied in the sample, 709 became more racially heterogeneous in the observation period. These measures suggest that the shift-share approach, while imperfect, is a potentially valid instrument.

Jaeger, et al (2018) further discuss the use of the shift-share instrument in immigration literature. Immigrants' location choices are not random, and the economy may adjust in many different ways to a change in local factor supplies. The conventional IV approach captures the short term impact, but also the longer term adjustment process to previous flows. In the case of the immigration model, authors argue that estimates are biased upwards (Jaeger 2018) or toward zero (Borjas 1999; 2006). Implementation of the shift-share instrument in the context of local or intra-country migration may suffer from some of the same types of biases.

Finally, in order to implement the shift-share procedure, it is assumed that the base shares in 1980 are endogenous, and shocks over time are close to exogenous (Borusyak, Jaravel and Hull 2022). It is therefore implied that changes in the predicted racial composition of a US Census Place will only impact the outcomes of interest, real tax and expenditure per capita, through its relationship with the observed racial composition of the Census Place. Goldsmith-Pinkham, et al., (2020) among others keenly point out that this exogeneity assumption is likely violated because dynamic adjustments over time may reflect institutional factors which are correlated with the distribution at the time of the initial period. In these cases, use of the 'past-settlement' instrument which interacts past shares with aggregate level growth, may exacerbate potential biases (Jaeger, et al., 2018).

In pursuit of more precise causal estimation, future work could implement a multiple instrumentation process as suggested in Jaeger, et al. (2018), to address the fact that outcomes related to migration and immigration will feature both shorter and longer term adjustment processes. In this paper, I demonstrate that the sign of the coefficient on Race heterogeneity for both expenditure and tax per capita is negative. It is important to emphasize that for the purposes of this paper, the magnitudes of the coefficents are of less interest than the sign. Previous literature which discusses mechanisms behind the finding that ethnic conflict is an important determinant of public expenditure relies on race heterogeneity being positively related to tax and expenditure, but negatively related to spending shares on certain public services. Even if the estimates produced by the shift-share instrumental approach are imperfect, they may have the power to cast some doubt upon the ethnic conflict hypothesis. If the direction of the bias of these estimates is either upward or toward zero, the true impact of diversification on tax and spending over time may be understated.

IV Results

The 2010 cross section contains 1422 cities across the continental United States. Alaska and Hawaii are not considered due to data inconsistency. Table 3 presents the results of OLS analysis. Table 4 compares these findings to the baseline results presented in Alesina, et al. (1999), converted to 2010 dollars where applicable. In a cross-section, the impact of race across time is remarkably consistent. The race heterogeneity measure is positively related to tax collection and expenditure per capita.

Using the same specification as the literature, I find coefficients which are categorically similar to those in the Alesina paper, adjusted for inflation. The coefficients on both tax and expenditure per capita are positive and statistically significant at the 99% percent confidence level. Further, the h_i measure is negatively related to percent spending shares on highways and sewage as well as highway spending per capita, but positively related to police share. These results verify the main relationships found the literature, for the most modern available Census data. Despite apparent increases in tax collection and expenditure per capita, expenditure shares on productive public goods aside from police decline. In addition to these main results, I also find the negative relationship to hold for spending share on parks, but a positive relationship between race heterogeneity and both housing and welfare spending shares.

This type of result has been used to validate the claim that racial conflict is an important factor in public service provision, and a literature which examines why that could be, suggesting a variety of mechanisms. Cities do appear to be collecting more tax revenue and spending more, but not spending on productive public goods. Cross-sectional estimates assume fundamentally that there must be something about race composition in the present which impacts public funding contemporaneously. It is more likely that the impact of changes in race distribution in a community's public funding measures unfold over time. Cross-sectional estimates are likely to be biased by omitted variables, sorting, and simultaneity.

Instead of viewing the relationship as a contemporaneous one, it may be more instructive to think about how changes in composition over time will impact tax and spending in that time period. Although racial heterogeneity likely evolves slowly over time, change in the United States has occured rapidly enough that over 30 years, there is significant variance both within and between Census Places to analyze. These panel statistics are illustrated in Table 5. Table 6 presents the results of a first difference model. This model is useful for two reasons. First, decennial Census data of any period is a snapshot in time of a changing landscape, not a steady state. Given the rapid changes in United States population, it is important to consider the impact of changes in demographics on these important public outcomes. Second, the cross-sectional analysis likely suffers from bias due to omitted variables. If any of these are time invariant, the first difference process will remove the impact, producing more precise estimates.

Interestingly, in the first difference model, the coefficients of interest on real tax per capita and real spending per capita are reversed from the cross-section. Changes in per capita expenditures and tax collection are now negatively related to changes in the race heterogeneity measure. A percent increase in race heterogeneity leads to -$13.81 decreases in real spending and -$7.20 tax revenue per capita over each 10 year period. Since it is difficult to conceptualize the impact of this value in a per person measurement over a decade's time, I find it more instructive to consider, given these results, how the *average* change in the heterogeneity measure would impact spending and tax. The mean change in race heterogeneity would correspond to an -6.4% decrease in real spending and -7.8% decrease in tax revenue per capita over each 10 year period.

The reversal of this coefficient is damaging to any theory of the mechanism linking race heterogeneity and public goods provision which relies on higher tax and spending per capita in more diverse areas but lower spending shares on public services. We now see that if a more racially diverse community is providing relatively lower levels of public goods, that it may actually be due to a dearth of tax revenue available to fund public measures, rather than disagreements or inability to pass public funding measures. Instead, future research should focus on attempting to explain statistically why more diverse places feature relatively *less* tax and spending per capita. One possible link is via income per capita, which is included as a control in the model. As the residents of Census Places become wealthier in terms of income per capita, there is a higher tax pool for income taxes, in addition to the possibility of greater sales and other types of taxes in the community. This in turn could be used to finance higher levels of public expenditure. As expected increases in income per capita predict higher expenditure and tax, but the relationship in the model is weak and only significant at the 10% significance level for expenditures.

But does this negative coefficient persist over time? Table 6 also presents results from an Arellano-Bond (1991) type specification which includes lagged values of the main outcome and response variables of interest. The addition of lags does not significantly alter the magnitude of the coefficients of interest on race h_i. While the coefficient on the lagged race measure remains negative (although statistically insignificant), the coeffcient on the lagged race measure for tax is positive. This result is amenable with the literature on ethnic enclaves, which suggests that over time

as communities diversify, cultural enclaves form amongst the subgroups of residents which then leads to formation of business and social networks. Originating in Bonacich and Modell (1980), an ethnic economy is one in which "immigrant and ethnic minorities create their own employment opportunities in the general labor market [...] The constitutive businesses are networked together in various ways depending on the type of industry and on the ethnic group (Zhou 1998), but an integral component of all ethnic economies is the operation of some form of socio-relational ties among economic agents and firms." (Mitchell 2000 p. 394). Over time, growing minority groups within racially heterogeneous communities are able to slowly develop the political and social capital required to coordinate public goods expansion and pass tax collection and spending measures. In other words, the group's efficacy, or ability to achieve together a socially optimal result, is increasing as the prevailing ethnic mix ingrains itself into the structure of a community.

Still, it is likely that institutional factors are important in this relationship (Lee, 2018). Omitted variables in the regression, such as bargaining power, culture, and political representation, and neighborhood preference are actually time variant, and so will continue to cause endogeneity bias in the parameters. In an attempt to correct this endogeneity, I instrument the race heterogeneity measure using predicted values from US national trend. To create the instrument following Boustan (2013), I use race composition of each Census 'Place' in the balanced panel in 1980 as the starting point. From 1980, I predict 1990 values using the national trends for each race, then use those to create a predicted race heterogeneity measure in 1990. Predictions for years 2000 and 2010 are computed using the previous years' prediction and national trends for that decade. National trends should not be influenced by the local policy or institutional decisions which influence local Place-level community composition.

To demonstrate this issue, I implement a Durbin-Wu-Hausman (Wooldridge 2010) test to determine whether the race heterogeneity measure is endogenous. Table 7 presents the results of this test. In the first stage, the potentially endogenous race heterogeneity measure is regressed on the set of controls and the predicted race instrument to obtain the residuals v_2. In the second stage, I include the stored residuals v_2 in the original models from Table 6. The t-statistic values on the stored residual v_2 are 7.41 and 8.55 for real expenditure and tax, respectively. From this test, we can reject the strict exogeneity of the race heterogeneity at the $<1\%$ confidence level. Instrumental variable modeling is probably a good idea, provided that we can trust the specification of the instruments.

Table 8 presents the results of the panel instrumental variables regression with a linear time trend, while Table 9 presents the first stage and diagnostic testing. The predicted race heterogeneity measure is highly correlated with the true values, but should not be directly influenced by the omitted variables and sorting which are causing problems for the measurement chosen. Diagnostic testing using the Stock-Yogo method demonstrates that the instrument is sufficiently strong. Further, a test using the difference in Sargan-Hansen statistics between models with and without the endogenous regressor indicate that the race heterogeneity measure is endogeneous at the $>99\%$ confidence level for both tax and spending.

Instrumental variables results confirm the findings from the first difference model, that the relationship between race heterogeneity and local tax and spending per capita over time is negative. A percent increase in race heterogeneity leads to -$9.36 decrease in real spending and -$6.35 decrease in tax revenue per capita over each 10 year period. Interpreted as an impulse-response at the mean, the average change in race heterogeneity is associated with a -4.3% decrease in real spending and -6.9% decrease in tax revenue per capita over each 10 year period in the IV model.

Table 8 also illustrates the importance of considering potential endogeneity in the relationship between race heterogeneity and tax or expenditure. Panels (2) and (4) present OLS specifications of the same model, which report that race heterogeneity is positively and significantly associated with both expenditure and tax per capita. However, these coefficients are reversed in both the IV results presented in that table and the first differences model from Table 6.

Minority groups sorting into cities are frequently low income, and do not have sufficient bargaining power or representation in local policy to pass spending measures targeted at their demographic or neighborhood. As a result, in the short run, the *a priori* belief should be that the coefficients on tax and especially spending might become more negative accounting for the endogeneity. More simply, minority groups are most frequently sorting into areas which are cheaper to live in. These areas often collect less tax and spend less, and because the influx of residents does not provide a sufficient tax basin to fund spending measures, these communities are most likely to become economically disenfranchised. As US Census Places diversify, there may also be spillover effects to consider. Changes in the socio-demographic distribution of an area might impact housing and job availability, causing lower income residents in high cost of living areas to migrate to lower cost of living areas. I consider geographic spillover effects, and the potential for another type of endogeneity, spatial spillovers, at a finer geographic (Census Tract) unit of analysis in forthcoming research.

Demographics statistics demonstrate that in the United States over time White populations have been decreasing while Black, Asian, Other, and the Hispanic ethnicities have increased. As a result, increases in the race heterogeneity measure generally correlate with larger non-White populations. Still, from a policy perspective, it disaggregate the race diversity measure to provide evidence that correct conclusion is more-White places spend and tax more while less-White places do the opposite. If this is the case, the economic conclusions shift to address 1) how to convince city planners and public policymakers not to pursue predominantly White non-Hispanic communities as a means of maximizing government revenues, tax and expenditure and 2) how to promote economic opportunity in Places which are rapidly diversifying and often feature high-density low income neighborhoods.

Tables 13 and 14 demonstrate that the negative relationship is driven by non-White groups in the race heterogeneity measure. The mean change in White population over each 10-year period is -5.64%. Since the coefficient on White percent is positive, increases in the White population would increase both expenditure and tax per capita. A one percent change in the White population increases expenditure per capita by $17.62, while percent changes in Black and Hispanic reduce expenditure per capita by -$8.40 and -$12.44, respectively.

Similar patterns emerge for taxation where one percent changes in the White population increase tax per capita by $7.67, while percent changes in Black and Hispanic reduce tax per capita by -$2.83 and -$3.93 respectively. A possible explanation lies in the labor market fragmentation literature, which ultimately could explain some degree of residential segregation along income and race lines. An alternative explanation could revolve around gentrification. Especially in larger cities, many communities which were largely Black and Hispanic in 1980 and 1990 underwent a revitalization, or gentrification which led to higher spending, and a higher income base from which to collect tax. The pre-gentrification population is likely to have consumption patterns of a lower social class in addition different ethnic and racial constitution as compared to the post-gentrification population (Spain 1980, Zukin 1987, LeGates & Hartman 2013). The results of Tables 13 and 14 demonstrate that the negative effect of the race heterogeneity measure on expenditure and tax appears to be driven by increases on non-White heterogeneity.

Labor market polarization, or a gap between concentration of higher and lower-skilled workers may arise following population migration and demographic change could in part explain the negative relationship between growth in non-White populations and declines in expenditure and tax. If people who work in traditionally minimum wage jobs such as services cluster in some areas while those who work in non-routine, higher skill positions live elsewhere (Rendall and Weiss 2016, Fan and Pena 2020). This type of polarization is likely to occur between US Census Places which are close together, such as clusters around large cities like Los Angeles and New York. Additional research could consider case studies of these closely clustered places to determine whether labor market polarization is occurring between communities by income and racial composition. A large degree of the labor market polarization will also occur within the Places themselves at the neighborhood level. This potentially justifies future research about the relationship between demographic composition and local expenditure, tax, and other economic indicators such as income and employment, at an even finer geographic level such as Census Tracts.

The idea that over time racially diversifying communities could experience business and capital growth and ultimately lead to wealthier communities with higher tax and expenditure is another potential justification for the conclusion of this paper, that economic policymakers should study the relationship between race diversity and local tax/spending. The goal should not be to reduce diversity or prevent the formation of ethnic enclaves, but rather, to understand why over time some communities become economically disenfranchised, while others thrive, and to promote progressive policies to improve standards of living in Places which suffer from lower tax/spending and thereby lower quality public goods and services. While public policymakers could choose to interpret the results of this and similar works as being supportive of demographic segregation in favor of Whiteness in order to increase tax and spending in the community, this is not the desired policy prescription. The formation of “Ghettos” characterized by low income per capita and often high concentration of racial minorities is associated with reductions in a variety of economic quality of life indicators from education and human capital formation to crime, income, health and life expectancy (Cutler and Glaeser 1997).

V Discussion and Robustness

The panel instrumental variables procedure used to demonstrate that race composition is endogenously related to tax and expenditure per capita uses a created variable, the predicted race composition in a Place. This procedure should create consistent estimates. In some well cited cases of using a predicted variable in the first stage, the covariance matrix for the second stage is adjusted to take into account the variability of estimates (Hole 2006). Efficiency of standard errors may be improved using the Murphy-Topel procedure (see: Murphy and Topel 1985, Hardin 2002).[1]

Important omitted variables in the model might include culture, representation or policy which varies at the state level. To account of this concern, Table 10 presents the first difference and instrumental variables models from Tables 6 and 8, with the addition of state level fixed effects which are suppressed in the output. The addition of state level fixed effects does not alter the interpretation of the coefficients in either model. It may be noteworthy if the addition of state level effects change the magnitude of the coefficents in the FD model by more than in the IV model. This potentially highlights the existence of time varying, state level omitted variables which are biasing the pooled, and to a lesser extent, the FD model.

If the sign of the coefficients on race heterogeneity for real spending and tax outcomes are actually negative, more research is warranted into mechanisms by which the variation in race could be impacting public finance. The first is the potential for reverse causality. To illuminate this, Table 11 presents the FD framework used previously, but with race h_i as the outcome variable, and the spending outcomes real expenditure and tax per capita used as independent variables in the model.

In the specification with race heterogeneity as an outcome, changes in both real expenditure per capita and real expenditure per capita are negatively, and significantly related to changes race heterogeneity. Interestingly, the coefficients remain negative when a lag of the spending or tax measure is included, however, the coefficient on that lagged measure is reversed. This result might further suggest the presence of 'ethnic enclaves'. As over time (20 years may be sufficient) these communities become more establish, there is small business and income growth which results in more tax and spending within the community. This process is attractive to members of that race or ethnicity group who then move to or close to the community over time. (Note that aggregation of a certain race would actually reduce race heterogeneity, evidence of which is in the negative lagged heterogeneity coefficient.)

Further research should consider the potential causal impact of race heterogeneity over time these public finance measures. One possibility is a supply/demand model for local government economic activity. After identifying the factors which affect tax and spending per capita, the underlying structural model is estimable with and without a race heterogeneity measure to determine whether the variables have a statistically significant causal impact on tax and

[1] I argue that this adjustment is not necessary because unlike the canonical cases where the first stage predictions are generated from an econometric model $E[y_1|X_1, \theta_1]$, the first stage predictions in this method are calculated using population parameters: Racial composition in 1980, and the observed growth rate of Races in the US in 10-year spans from 1980 to 2010.

spending per capita.

In later work, I intend to specify a structural public finance model and fit it to local Colorado data. Even so, such a structural model would not likely be able to account for all of the unobserved variables which are impacting tax and spending. Likewise, the dynamic choice framework of public goods games may be incomplete; "[. . .]if residential choice is hampered by wealth constraints or interacts with other inherited characteristics, the long-run equilibrium (of a dynamic choice model) may depend on the history of the communities" (Benabou pg. 25). While it is not generally possible for the researcher to observe a history of racial interaction and policy within a community, it may be possible to create a valid proxy by studying employment and earnings by location and race composition. The a priori expectation is that if individuals are sorting into ethnic enclaves in pursuit of more favorable employment outcomes (Edin 2003, Damm 2009), the amount of segregation chosen will have important impacts on income per capita, and thus tax and spending per capita.

The impact of outliers in real spending/tax, or the race heterogeneity measure could be a potential cause for concern. In the case of real spending and tax, I look for outliers by first differencing and identifying individual observations at the extremes. These extreme observations were compared to the Willamette College aggreates from the Census of Governments publically available. While the data is drawn from the same place, the data set was published independently from the city and county factbooks which compose my dataset, so makes a good consistency check. I found the results in my dataset to be generally valid, even those which showed extreme changes in tax and spending per capita over the 10 year periods.

Censorship of the data to remove observations for tax and spending at the extremes may result in further bias if there is endogeneity in the determination of whether a place receives the treatment (being dropped from the sample). Cursory results which may be biased due to the assignment of treatment do not impact the magnitude or significance of important results and so are omitted here.

Data validation checks for the race heterogeneity measure are more complicated due to the limitations on collecting and publishing this data. The ACS surveys provide a potential validation source, but are only small sample estimates, especially for the smaller places included in my sample. As with the case of real spending/tax, censorship of the data may result in bias. To check for the impact of outliers in the heterogeneity measure, I censor the data in Table 12 to include only values of race heterogeneity between 5% and 70%. This restriction is mostly arbitrary, except that it results in the exclusion of about 100 data points, about half at each extreme. Results are presented with the sole purpose of demonstrating that extreme values of the race heterogeneity measure do not appear to be impacting the magnitude or significance of published coefficients in any meaningful way.

As a test of instrumental validity, I include a specification in Table 15 with three IVs instead of one. The additional IVs are lagged variables as suggested in Alesina (1999). Wooldridge (2010) suggests that when there are more instruments than are required to identify an equation, we can implement a Sargan-Hausman test to determine whether the

additional instruments are valid, in the sense that they are uncorrelated with the error term. Under the null hypothesis, the potential instruments are uncorrelated. Rejection of the null hypothesis would be cause to reevaluate the choice of IVs. The test does not indicate which of the IVs fail the exogoeneity requirement. It could be one or all of them. Using the Bartik shock race instrument as well as the two lags, we reject the null hypothesis of the Sargan-Hausman test. To test further, I also use % foreign born as a contemperaneous instrument. For all combinations using any lagged variable the null hypotheis is rejected. For the combination of only contemperaneous variables, the Bartik shock and % foreign born, in this case, we fail to reject the null hypothesis. This result leads the author to believe that it is the lagged variables which are potentially problematic in a panel of only four time periods. I also provide disaggregated race measures in Tables 13 and 14 in order to demonstrate that more diversity as measured by growth of Black or Hispanic is driving the relationship in the Herfindahl measure. An additional relevance test for weak instruments is reported using the method of Stock and Yogo (2005) in Table 9.

As an alternative instrumental approach, Lewbel (2012) suggests that it is possible to identify structural parameters in regression models with endogenous or mismeasured regressors in the absence of traditional identifying information. Given concerns about the validity of the Bartik style shift-share approach, I provide an alternative specification in Table 18 using the Lewbel Instrumental Variables approach. In this case, I assume that the traditional approaches such as external instruments or repeated measurements are unavailable and instead identify parameters by having regressors that are uncorrelated with the product of heteroskedastic errors. In this approach, the sign of the main coefficients of interest, on real expenditure and tax both remain negative, but lose their significance. Future work could improve upon the accuracy of these estimates by implementing alternative instrumental approaches. These results continue to suggest that changes in the race heterogeneity measure are not positively related to changes in real tax and expenditure, as would be required for the ethnic conflict hypothesis on race and public spending to hold.

A final type of concern surrounds the size or location of the places included in the sample. Since I observe only Census Places greater than 25k in population, we might want to know whether the size of the city is a meaningful determinant of the relationship between race heterogeneity and real spending/tax. Likewise, unobserved variables at the state or regional level such as culture or political representation could vary the results geographically. The impacts of spatial and size distribution of the Places in the sample on the results is an interesting question which warrants further investigation. To test for issues related to sampling I conduct robustness checks on the sample chosen, which are displayed in the Appendix.

First, migration patterns are not random and individuals are likely sorting into Places based on factors such as the public goods provided. Theory suggests Tiebout sorting works best in urban areas with a large number of small Census places. As a robustness check, in Table 16, I reduce the sample to only Census Places contained within the largest quartile of MSAs by population. This restriction will partially allow for labor market sorting which occurs less frequently between smaller Census Places spaced further apart. The results from the restricted, large-MSA sample are

virtually identical to those in Table 6.

As another robustness check on labor market area sorting, I restrict the sample to only Census Places which are located within a certain distance of other Census Places in the dataset. I use GIS to calculate the mileage distance and travel time to the nearest other Census Place, and present results for all Places within 30, 60 and 90 miles in Table 17. For both real tax and expenditure per capita, the estimated coefficients start slightly smaller than in the full sample, although similar, and trend toward their full sample coefficients as the distance range becomes wider.

Despite these robustness checks, there may still be concerns about demographic change and labor market sorting which occurs in less populated areas. Data which only include larger towns and cities which are often located far apart geographically cannot fully capture sorting which occurs based on taxation and public services. Additional research should carefully consider less populated areas to highlight similarities and differences in patterns related to diversification over time. In forthcoming work, I consider demographic change at the Census Tract level to comment on urban/rural differences in measurable economic indicators.

VI Conclusion

Alesina, Baqir and Easterly (1999) show that ethnic heterogeneity is associated with lower shares of public goods provision. Literature following that work takes as given the fact that there is a positive relationship between race heterogeneity and tax/spending per capita, but a negative relationship between race and spending shares on important public goods/services, and posits mechanisms behind this relationship summarized in Habyarimana (2009). While this result does hold for an updated cross-section of 2010 data, it does not appear that the relationship holds over time as Census Places of over 25K population diversify. Due to the slowly evolving nature of the relationship between race heterogeneity and public tax and expenditure at the local level, recent literature suggests that it is important to consider the time dynamic (An, et al., 2018) and institutions which may be present in this relationship (Lee, 2018).

While there has been extensive analysis of the conclusion that diversity reduces spending shares on important public goods (Hopkins 2019 and 2011, Lee 2013, Lee, et al., 2016), there has been limited analysis of the assumption that diverse places tax and spend more per capita (which then does not materialize as higher spending on public goods due to ethnic conflict). Following previous approaches by Boustan, et al. (2013) and Lee, et al., (2016), I implement a shift-share instrumental approach to account for endogeneity due to omitted and time variant factors. I find evidence in contrast to a long-standing stylized fact from Alesina, et al. (1999); that race heterogeneity is an important (positive) determinant of public spending. Conversely, I show that over time the coefficent for race heterogeneity on both tax and spending per capita is negative and not positive. This result provides further evidence in opposition to the racial conflict hypothesis on the relationship between diversity and public spending.

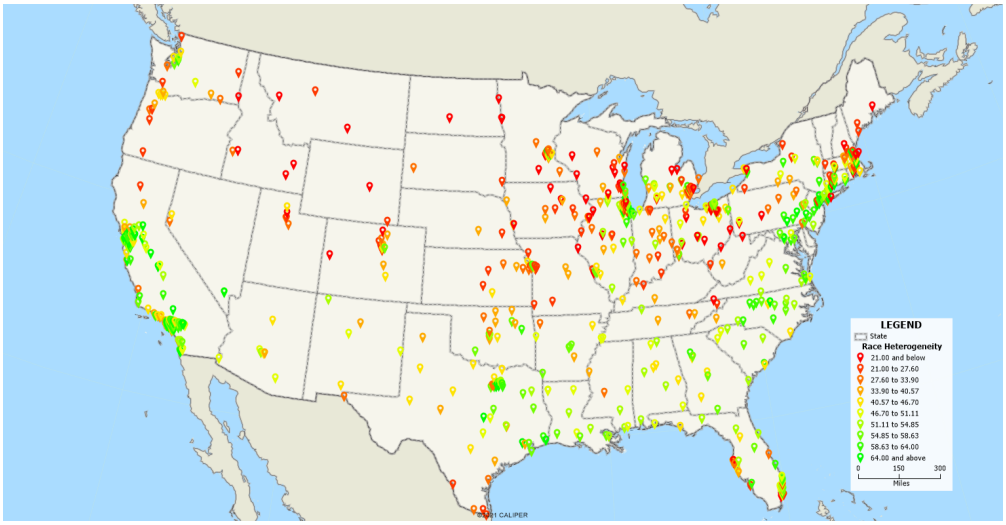

Figure 1: Race Heterogeneity in 2010

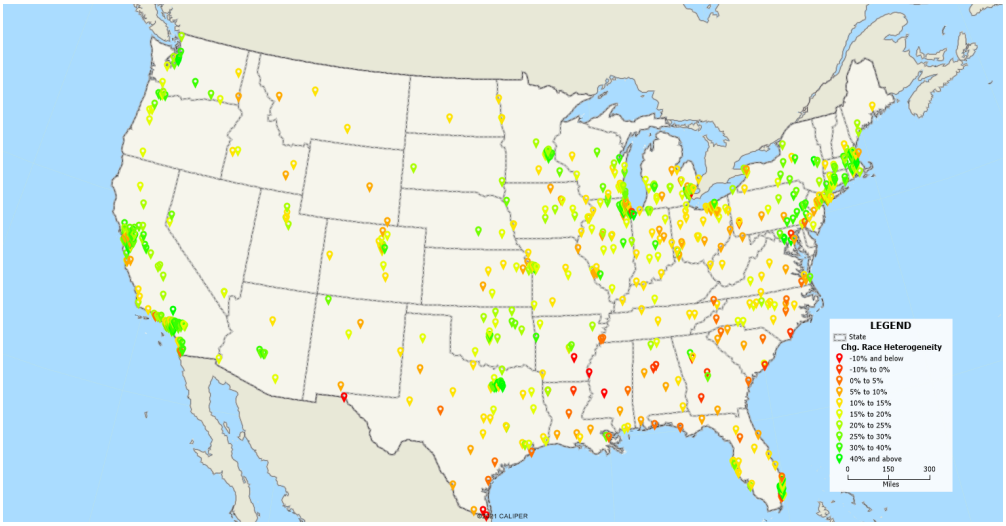

Figure 2: Change in Heterogeneity 1980-2010

Table 1: Selected Demographic Measures 1980-2010

Variable	year	mean	sd	min	max	N
Population						
	2010	141139	390356	24751	8336697	734
	2000	130280	373914	20681	8008278	734
	1990	118111	343740	25098	7322564	734
	1980	108881	330714	25112	7071639	734
log(Population)						
	2010	11.305	.8267	10.125	15.936	734
	2000	11.230	.8042	9.937	15.896	734
	1990	11.137	.7881	10.130	15.806	734
	1980	11.035	.7892	10.131	15.772	734
% White						
	2010	67.238	18.470	10.6	96.5	734
	2000	71.554	18.463	12.3	97.9	734
	1990	78.994	16.772	17.4	99.3	734
	1980	84.169	14.623	30.5	99.5	734
% Black						
	2010	14.974	16.753	0.30	82.7	734
	2000	13.669	16.277	0.20	81.6	734
	1990	12.205	15.018	0.08	75.7	734
	1980	10.594	14.040	0.01	66.8	734
% Asian/Pac. Islander						
	2010	6.331	9.242	0.10	66.9	734
	2000	4.847	7.455	0.20	61.8	734
	1990	3.713	5.704	0.07	57.46	734
	1980	1.922	3.039	0.03	34.76	734
% "Other"						
	2010	7.180	7.742	0.20	45.4	734
	2000	9.218	9.121	0.70	56.0	734
	1990	4.448	7.242	0.04	65.4	734
	1980	2.771	4.366	0.00	26.6	734
% Hispanic						
	2010	18.420	18.840	0.90	97.1	734
	2000	14.209	17.180	0.50	95.6	734
	1990	10.160	14.791	0.23	93.9	734
	1980	7.518	12.317	0.14	93.01	734
Race Heterogeneity h_i						
	2010	43.942	16.273	6.86	77.21	734
	2000	40.073	18.921	4.15	86.87	734
	1990	30.571	18.608	1.44	90.22	734
	1980	23.792	17.075	0.91	72.99	734

Table 2: Selected Provision Measures and Controls 1980-2010

Variable	year	mean	sd	min	max	N
Real Exp/Capita						
	2010	1713.30	1043.88	292	10907	734
	2000	1377.76	801.62	344	7106	734
	1990	1469.11	841.70	269	7652	734
	1980	1248.54	721.71	169	4914	734
Real Tax/Capita						
	2010	781.23	505.62	141	7261	734
	2000	573.16	351.92	68	3063	734
	1990	610.15	378.43	84	3654	697
	1980	512.18	338.91	73	2786	734
Real Intergov/Capita						
	2010	446.08	610.57	17	6565	734
	2000	789.25	553.14	113	3955	734
	1990	851.31	576.31	132	4246	697
	1980	441.68	402.13	15	2957	734
log(Income/Capita)						
	2010	10.15	.295	9.39	11.31	734
	2000	10.17	.278	9.28	11.36	734
	1990	10.08	.268	9.25	11.25	734
	1980	9.94	.208	9.25	10.82	734
Per Capita/Median Income Ratio						
	2010	53.01	8.67	28.42	99.22	734
	2000	51.44	8.71	26.59	101.94	734
	1990	48.36	7.83	25.71	107.78	734
	1980	44.49	6.84	25.91	104.72	734
% BA or higher						
	2010	29.72	14.12	4.6	79.4	734
	2000	26.51	13.22	4.0	74.4	734
	1990	22.92	11.61	1.7	71.2	734
	1980	10.88	5.54	1.1	37.4	734
% over 65						
	2010	12.63	3.94	4.6	27.9	734
	2000	12.64	4.01	3.2	37.7	734
	1990	12.78	4.74	2.7	47.6	734
	1980	11.38	5.20	1.7	51.7	734

Table 3: Coefficients on Race Heterogeneity h_i for Expenditure Shares 2010

Variable name	1 (no controls)	6 (all controls)	#obs	Adj. R^2
Share of spending on:				
Roads	-0.077^{***}	-0.057^{***}	1422	0.06
	(0.014)	(0.014)		
Sewerage	-0.065^{***}	-0.095^{***}	1422	0.08
	(0.015)	(0.016)		
Police	0.066^{***}	0.083^{***}	1422	0.05
	(0.012)	(0.014)		
Parks	-0.050^{***}	-0.024^{**}	1362	0.07
	(0.011)	(0.012)		
Housing	0.119^{***}	-0.119^{***}	1135	0.06
	(0.014)	(0.016)		
Health	-0.001	-0.001	854	0.02
	(0.013)	(0.015)		
Welfare	0.027^{**}	0.015	252	0.08
	(0.011)	(0.013)		
Totals:				
Roads Spending/Capita	-68.18^{***}	-43.41^{*}	1422	0.02
	(23.81)	(25.02)		
Intergov. Rev/Capita	362.99^{***}	134.45	1422	0.06
	(93.73)	(83.19)		
Taxes/Capita	378.76^{***}	483.72^{***}	1422	0.16
	(82.01)	(93.79)		
Expenditure/Capita	864.62^{***}	636.38^{***}	1422	0.09
	(167.93)	(167.61)		

Table 4: Coefficient on Race Heterogeneity 2010 vs Alesina 1990, adjusted for inflation

Year	exp/capita	tax/capita	hwy/capita	hwy%	police%	sewer%
1990	544.41	256.88	-63.36	-0.083	0.099	-0.079
2010	636.38	483.72	-43.41	-0.057	0.083	-0.095

Table 5: Panel Within and Between Summary Statistics

	Mean	SD		Mean	SD
Race Heterogeneity	34.595	19.427	Δ Race Heterogeneity	6.717	7.082
Between		16.787	Between		4.257
Within		9.793	Within		5.661
Real Exp/Capita	1452.18	876.66	Δ Real Exp/Capita	154.92	540.81
Between		782.57	Between		256.62
Within		395.90	Within		476.12
Real Tax/Capita	617.04	392.83	Δ Real Tax/Capita	87.33	194.98
Between		362.11	Between		104.35
Within		153.90	Within		169.94

Table 6: First Difference Regression Results on (real) Per Capita Outcomes

	(1) Δ Exp/Capita	(2) Δ Exp/Capita	(3) Δ Tax/Capita	(4) Δ Tax/Capita
Δ Race	-13.81***	-11.63***	-7.196***	-6.091***
	(-8.57)	(-7.23)	(-11.86)	(-9.61)
Δ log(pop)	-122.6	-121.0	-30.64	-29.82
	(-1.61)	(-1.60)	(-1.00)	(-0.99)
Δ Intergov/Cap	0.422***	0.422***	-0.0162	-0.0162
	(7.27)	(7.27)	(-1.38)	(-1.40)
Δ Inc/Cap	0.00116	0.00198	-0.00195	-0.00154
	(0.29)	(0.49)	(-1.12)	(-0.88)
Δ Inequality	-25.29***	-22.48***	-9.232***	-7.805***
	(-6.15)	(-5.57)	(-5.97)	(-4.93)
Δ % BA or higher	-1.237	0.0124	6.002***	6.636***
	(-0.49)	(0.01)	(6.32)	(6.99)
Δ % over 65	-16.60**	-13.44*	2.618	4.220
	(-3.01)	(-2.44)	(1.11)	(1.70)
Lag. Race		2.439***		1.239***
		(3.76)		(4.76)
year	0.172***	0.117***	0.0647***	0.0367***
	(11.55)	(6.80)	(15.20)	(5.35)
N	2128	2128	2128	2128

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Durbin-Wu-Hausman (1978) Test for Endogeneity

	(1) Δ Exp/Capita	(2) Δ Tax/Capita
Δ Race	-7.989***	-4.217***
	(-4.70)	(-6.32)
Δ log(pop)	53.29	59.41*
	(0.72)	(1.99)
Δ Intergov/Cap	0.460***	0.00301
	(7.85)	(0.26)
Δ Income/Cap	0.00358	-0.000714
	(0.91)	(-0.42)
Δ Inequality	-32.63***	-12.99***
	(-7.93)	(-8.21)
Δ % Bachelor or higher	-1.527	5.853***
	(-0.62)	(6.61)
Δ % over 65	-26.16***	-2.277
	(-4.69)	(-0.95)
v2	-29.99***	-15.35***
	(-7.41)	(-8.55)
year	0.256***	0.108***
	(12.53)	(14.92)
N	2128	2128

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: XTIV Estimation with Linear Time Trend

	(1) Exp/Capita	(2) Exp/Capita	(3) Tax/Capita	(4) Tax/Capita
Race h_i	-9.358**	3.819***	-6.345***	2.161***
	(-2.83)	(6.71)	(-4.92)	(6.31)
log(pop)	-73.75	-2.709	-2.444	-6.252
	(-1.50)	(-0.34)	(-0.11)	(-1.23)
Intergov	0.666***	1.098***	0.0396**	0.262***
	(7.39)	(22.99)	(2.78)	(12.46)
Income/Capita	0.0267***	0.0108***	0.00985***	0.0113***
	(6.87)	(4.93)	(5.33)	(7.51)
Inequality	-20.29***	9.366***	-7.089***	2.208
	(-5.82)	(4.84)	(-5.10)	(1.74)
% BA or Higher	0.207	0.228	2.524**	0.547
	(0.09)	(0.16)	(2.63)	(0.67)
% over 65	-9.442*	-6.150*	2.054	3.838*
	(-2.52)	(-2.00)	(0.99)	(2.08)
N	2899	2899	2899	2899
R^2	0.425	0.890	0.430	0.790
F	56.02	3114.3	121.61	1737.8
Specification	IV	OLS	IV	OLS
Weak ID F-test	584.54	–	584.54	–
Weak ID Robust P-value	0.001	–	0.001	–
Endogeneity χ^2	8.74	–	13.15	–
Endogeneity P-value	0.003	–	0.001	–

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 9: IV First Stage Results

	(1) Race
$\hat{Race}$	0.835***
	(21.60)
Intergov	0.00221***
	(6.31)
log(pop)	3.907***
	(4.02)
Income/Capita	-0.000138**
	(-2.31)
Inequality	-0.749***
	(-12.19)
% over 65	-0.0452
	(-0.60)
% BA or Higher	0.102***
	(3.01)
N	2899
Weak ID F	466.48
Weak ID P	.0001
Under ID χ^2	169.5
Under ID P	.0001
Weak ID 10%	16.38
Weak ID 25%	5.53
Robust Inference F	22.64
Robust Inference P	.0001

t statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 10: FD Model with State FE

	(1) Δ Exp/Capita	(2) Δ Tax/Capita
Δ Race	-9.540***	-5.902***
	(1.618)	(0.989)
Δ Intergov	0.573***	0.0398***
	(0.081)	(0.014)
Δ log(pop)	-51.43	-58.33*
	(83.859)	(35.037)
Δ Income/Capita	0.00572	-0.000966
	(0.004)	(0.002)
Δ Inequality	-19.50***	-6.145***
	(5.086)	(1.962)
Δ % BA or Higher	9.416***	11.20***
	(2.278)	(1.236)
Δ % over 65	1.248	8.422***
	(6.796)	(2.857)
FE	State & Year	State & Year
N	2128	2128
R^2	0.347	0.290
F	10.67	14.93

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 11: Reverse Causality Tests

	(1) Δ Race	(2) Δ Race
Δ Exp/Capita	-0.00309***	-0.00276***
	(0.000)	(0.000)
L.Exp/Capita		0.000843***
		(0.000)
Δ Intergov	0.00369***	0.00395***
	(0.001)	(0.001)
Δ log(pop)	9.038***	9.764***
	(1.183)	(1.210)
Δ Income/Capita	0.000118**	0.0000681
	(0.000)	(0.000)
Δ Inequality	-0.234***	-0.298***
	(0.060)	(0.061)
Δ % BA or Higher	0.131***	0.115***
	(0.027)	(0.028)
Δ % over 65	-0.386***	-0.363***
	(0.086)	(0.085)
Year	1.928***	2.566***
	(0.104)	(0.140)
L.Race		-0.0901***
		(0.009)
N	2128	2128
R^2	0.495	0.524
F	260.2	265.3

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

	(1) Δ Race	(2) Δ Race
Δ Tax/Capita	-0.00720***	-0.00632***
	(0.002)	(0.002)
L.Tax/Capita		0.00247***
		(0.000)
Δ Intergov	0.00217***	0.00240***
	(0.000)	(0.000)
Δ log(pop)	9.078***	10.27***
	(1.175)	(1.197)
Δ Income/Capita	0.0000981**	0.0000150
	(0.000)	(0.000)
Δ Inequality	-0.228***	-0.264***
	(0.058)	(0.059)
Δ % BA or Higher	0.181***	0.160***
	(0.028)	(0.028)
Δ % over 65	-0.310***	-0.291***
	(0.086)	(0.084)
Year	1.868***	2.332***
	(0.104)	(0.130)
L.Race		-0.0829***
		(0.008)
N	2128	2128
R^2	0.502	0.531
F	267.3	273.4

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 12: FD Results with Censored Race Heterogeneity (compare to Table 6)

	(1) Δ Exp/Capita	(2) Δ Tax/Capita
Δ Race	-14.55***	-7.205***
	(1.683)	(0.587)
Δ Intergov	0.517***	0.0195
	(0.0717)	(0.0130)
Δ log(pop)	-186.1**	-41.41
	(76.06)	(30.32)
Δ Income/Capita	0.00676	-0.000369
	(0.00412)	(0.00174)
Δ Inequality	-29.65***	-9.897***
	(4.422)	(1.511)
Δ % over 65	-7.671	5.653**
	(5.819)	(2.577)
Δ % BA or Higher	0.260	6.645***
	(2.568)	(0.929)
Year	121.2***	43.13***
	(10.74)	(2.451)
N	2020	2020
R^2	0.295	0.303
F	35.64	102.0

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 13: FD Expenditure Results for Individual Races (compare to Table 6)

	(1) Δ Exp/Capita	(2) Δ Exp/Capita	(3) Δ Exp/Capita
Δ % White	17.62***		
	(2.297)		
Δ % Black		-8.401**	
		(3.050)	
Δ % Hispanic			-12.44***
			(3.383)
Δ Intergov	0.524***	0.497***	0.507***
	(0.0696)	(0.0675)	(0.0685)
Δ log(pop)	-243.8**	-330.1***	-242.6**
	(74.02)	(75.41)	(74.57)
Δ Income/Capita	0.00485	0.00396	0.00546
	(0.00405)	(0.00411)	(0.00407)
Δ Inequality	-31.77***	-26.81***	-30.67***
	(4.467)	(4.144)	(4.687)
Δ % BA or Higher	-0.0808	-1.462	-2.646
	(2.482)	(2.472)	(2.600)
Δ % over 65	-7.610	-5.365	-6.215
	(5.604)	(5.714)	(5.779)
Year	126.8***	100.5***	113.9***
	(11.25)	(9.279)	(11.69)
N	2128	2128	2128
R^2	0.292	0.266	0.270

Standard errors in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: FD Tax Results for Individual Races (compare to Table 6)

	(1) Δ Tax/Capita	(2) Δ Tax/Capita	(3) Δ Tax/Capita
Δ % White	7.672***		
	(0.787)		
Δ % Black		-2.833**	
		(1.073)	
Δ % Hispanic			-3.928**
			(1.202)
Δ Intergov	0.0167	0.00403	0.00682
	(0.0128)	(0.0127)	(0.0126)
Δ log(pop)	-82.27**	-119.2***	-91.38**
	(29.70)	(30.05)	(31.30)
Δ Income/Capita	-0.00109	-0.00132	-0.000805
	(0.00174)	(0.00180)	(0.00178)
Δ Inequality	-11.03***	-8.967***	-10.21***
	(1.541)	(1.517)	(1.598)
Δ % BA or Higher	6.555***	5.913***	5.530***
	(0.933)	(0.919)	(0.913)
Δ % over 65	6.112*	7.333**	7.118**
	(2.448)	(2.520)	(2.570)
Year	44.29***	32.50***	36.64***
	(2.644)	(2.256)	(2.771)
N	2128	2128	2128
R^2	0.272	0.237	0.239

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Coefficients using three instruments: Predicted Race, Lagged Race, Lagged Income

	(1) Exp/Capita	(2) Tax/Capita
Race	-5.573	-4.346***
	(-1.83)	(-3.56)
Income/Capita	-0.109***	-0.0485***
	(-6.31)	(-7.18)
Intergov	0.387***	-0.0443**
	(3.97)	(-2.87)
log(pop)	206.3*	61.70
	(2.02)	(1.50)
Inequality	19.08**	5.925*
	(3.22)	(2.35)
% BA or Higher	75.01***	35.19***
	(7.34)	(10.85)
% over 65	-14.60*	5.869*
	(-2.08)	(2.18)
N	2165	2165
R^2	.0623	.1575
Weak ID F-Test	64.466	64.466

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 16: First Difference Regression Results on (real) Per Capita Outcomes restricted to largest MSA Quartile

	(1) Δ Exp/Capita	(2) Δ Exp/Capita	(3) Δ Tax/Capita	(4) Δ Tax/Capita
Δ Race	-16.822***	-15.139***	-7.283***	-6.337***
	(2.965)	(2.921)	(1.034)	(1.250)
Δ log(pop)	3.455	13.761	9.992	15.786
	(127.420)	(126.393)	(63.700)	(63.650)
Δ Intergov/Cap	0.531***	0.532***	-0.097**	-0.096**
	(0.169)	(0.169)	(0.046)	(0.046)
Δ Inc/Cap	-0.012	-0.011	-0.007*	-0.007*
	(0.008)	(0.008)	(0.004)	(0.004)
Δ Inequality	-12.840	-10.790	-7.947**	-6.794
	(8.295)	(8.539)	(3.892)	(4.223)
Δ % BA or higher	12.354	13.919*	12.605***	13.485***
	(7.782)	(7.449)	(2.745)	(2.780)
Δ % over 65	-11.730	-11.291	-0.645	-0.398
	(8.794)	(8.893)	(4.514)	(4.612)
Lag. Race		1.590		0.894
		(1.470)		(0.680)
year	0.160***	0.114**	0.055***	0.028
	(0.039)	(0.045)	(0.010)	(0.022)
N	592	592	592	592

Robust standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 17: First Difference Regression Results on (real) Per Capita Outcomes restricted by Proximity

	(30mi) ΔExp/Capita	(60mi) ΔExp/Capita	(90mi) ΔExp/Capita	(30mi) ΔTax/Capita	(60mi) ΔTax/Capita	(90mi) ΔTax/Capita
ΔRace	-10.889***	-11.516***	-12.078***	-5.781***	-6.136***	-6.256***
	(1.806)	(1.662)	(1.641)	(0.729)	(0.679)	(0.650)
Δlog(pop)	-148.711*	-155.569**	-110.253	-31.285	-41.207	-27.125
	(85.170)	(76.948)	(75.188)	(36.116)	(32.380)	(30.795)
ΔIntergov/Capita	0.427***	0.415***	0.430***	-0.006	-0.016	-0.016
	(0.074)	(0.059)	(0.059)	(0.014)	(0.013)	(0.012)
ΔIncome/Capita	-0.000	0.001	0.001	-0.002	-0.001	-0.002
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
ΔInequality	-22.053***	-22.342***	-21.583***	-7.101***	-8.159***	-7.689***
	(4.900)	(4.248)	(4.096)	(2.245)	(1.942)	(1.626)
Δ% Bachelor or higher	3.995	2.772	0.959	6.973***	7.086***	6.696***
	(2.920)	(2.440)	(2.391)	(1.229)	(1.048)	(0.971)
Δ% over 65	-8.170	-10.241*	-13.121**	5.053*	4.475*	4.111
	(6.177)	(5.682)	(5.554)	(2.901)	(2.644)	(2.509)
Lag Race	2.841***	2.817***	2.618***	1.265***	1.214***	1.219***
	(0.785)	(0.696)	(0.663)	(0.326)	(0.293)	(0.269)
year	0.105***	0.109***	0.113***	0.036***	0.037***	0.038***
	(0.021)	(0.018)	(0.018)	(0.009)	(0.008)	(0.007)
Observations	1487	1848	2046	1487	1848	2046

Robust standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 18: Lewbel (2012) Instrumental Variables

	(1) ΔExp/Capita	(2) ΔTax/Capita
ΔRace	-7.681	-2.080
	(8.721)	(3.298)
Δlog(pop)	314.267*	79.768
	(168.825)	(63.850)
ΔIntergov/Capita	0.607***	0.055***
	(0.029)	(0.011)
ΔIncome/Capita	0.005	-0.001
	(0.006)	(0.002)
ΔInequality	-16.109***	-4.887**
	(5.063)	(1.915)
Δ% Bachelor or higher	16.429***	13.652***
	(3.682)	(1.393)
Δ% over 65	34.452**	24.951***
	(14.908)	(5.638)
Lag Race	-5.572	1.066
	(5.473)	(2.070)
year	41.253***	12.913***
	(3.113)	(1.177)
Observations	2128	2128
R^2	0.290	0.225
F	61.47	42.80

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Works Cited

[1] A. Alesina, R. Baqir, and W. Easterly. Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284, nov 1999.

[2] Alberto Alesina, Arnaud Devleeschauwer, William Easterly, and Sergio Kurlat. Fractionalization. *Journal of Economic Growth*, 8(2):155–194, 2003.

[3] Brian An, Morris Levy, and Rodney Hero. It's Not Just Welfare: Racial Inequality and the Local Provision of Public Goods in the United States. *Urban Affairs Review*, 54(5):833–865, 2018.

[4] Manuel Arellano and Stephen Bond. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297, 1991.

[5] Kate Baldwin and John D. Huber. Economic versus cultural differences: Forms of ethnic diversity and public goods provision. *American Political Science Review*, 104(4):644–662, 2010.

[6] Abhijit Banerjee, Lakshmi Iyer, and Rohini Somanathan. History, social divisions, and public goods in rural India. *Journal of the European Economic Association*, 3(2-3):639–647, 2005.

[7] Pranab Bardhan. Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation Communities in South India. *Economic Development and Cultural Change*, 48(4):847–865, 2000.

[8] Timothy J Bartik. Boon or boondoggle? The debate over state and local economic development policies. 1991.

[9] Nathaniel Baum-Snow and Fernando Ferreira. Causal inference in urban and regional economics. In *Handbook of regional and urban economics*, volume 5, pages 3–68. Elsevier, 2015.

[10] R. Benabou. Workings of a City: Location, Education, and Production. *The Quarterly Journal of Economics*, 108(3):619–652, 1993.

[11] Olivier Jean Blanchard, Lawrence F Katz, Robert E Hall, and Barry Eichengreen. Regional evolutions. *Brookings papers on economic activity*, 1992(1):1–75, 1992.

[12] Edna Bonacich and John Modell. *The economic basis of ethnic solidarity*. University of California Press, 1980.

[13] George J. Borjas. To Ghetto or Not to Ghetto: Ethnicity and Residential Segregation. *Journal of Urban Economics*, 44(2):228–253, 1998.

[14] Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213, 2022.

[15] Leah Boustan, Fernando Ferreira, Hernan Winkler, and Eric M. Zolt. The effect of rising income inequality on taxation and public expenditures: Evidence from U.S. Municipalities and school districts, 1970-2000. *Review of Economics and Statistics*, 95(4):1291–1302, 2013.

[16] Sean Corcoran and William N Evans. Income Inequality, the Median Voter, and the Support for Public Education. *Public Choice: Analysis of Collective Decision Making*, 2010.

[17] Anna Pill Damm. Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2):281–314, 2009.

[18] Linda Darling-Hammond. Race, inequality and educational accountability: the irony of 'No Child Left Behind'. *Race Ethnicity and Education*, 10(3):245–260, sep 2007.

[19] Rebecca Diamond. The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524, 2016.

[20] Per-Anders Edin, Peter Fredriksson, and Olof Aslund. Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment. *The Quarterly Journal of Economics*, 118(1):329–357, 2003.

[21] Maoyong Fan and Anita Alves Pena. Decomposing US Political Ideology: Local Labor Market Polarization and Race in the 2016 Presidential Election. *Journal of Economics, Race, and Policy*, 4(1):56–70, 2021.

[22] Kevin Fiscella, Peter Franks, Marthe R. Gold, and Carolyn M. Clancy. Inequality in Quality. *Jama*, 283(19):2579, 2000.

[23] Alessandra Fogli and Veronica Guerrieri. The End of the American Dream? Inequality and Segregation in US Cities. *NBER Working Papers*, 2019.

[24] Dierdre A. Gaquin and Meghan Ryan. *2013 County and City Extra: Annual Metro, City, and County Data Book*. Bernan Press, 21st edition, 2013.

[25] Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624, 2020.

[26] Eric D Gould, Bruce A Weinberg, and David B Mustard. Crime rates and local labor market opportunities in the United States: 1979–1997. *Review of Economics and statistics*, 84(1):45–61, 2002.

[27] James Habyarimana, Macartan Humphreys, Daniel N Posner, M Jeremy, Jeremy M Weinstein, and James Habyarimana. Why Does Ethnic Diversity Undermine Public Goods Provision? *American Political Science Review*, 101(4):709–725, 2007.

[28] Michael R. Haines. No TitleHistorical, Demographic, Economic, and Social Data: The United States, 1790-2002. *Inter-university Consortium for Political and Social Research*, 2010.

[29] James W. Hardin. The Robust Variance Estimator for Two-stage Models. *The Stata Journal: Promoting communications on statistics and Stata*, 2(3):253–266, 2002.

[30] Arne Risa Hole. Calculating Murphy-Topel variance estimates in Stata: A simplified procedure. *Stata Journal*, 6(4):521–529, 2006.

[31] Daniel J Hopkins. The diversity discount: When increasing ethnic and racial diversity prevents tax increases. *The Journal of Politics*, 71(1):160–177, 2009.

[32] Daniel J. Hopkins. The limited local impacts of ethnic and racial diversity. *American Politics Research*, 39(2):344–379, 2011.

[33] Ichiro Kawachi, Norman Daniels, and Dean E. Robinson. Health disparities by race and class: Why both matter. *Health Affairs*, 24(2):343–352, 2005.

[34] Alexander Lee. Ethnic diversity and ethnic discrimination: Explaining local public goods provision. *Comparative Political Studies*, 51(10):1351–1383, 2018.

[35] Soomi Lee. Racial heterogeneity and medicaid expenditure in the us states: A longitudinal analysis. *The Journal of socio-economics*, 45:28–37, 2013.

[36] Soomi Lee, Dongwon Lee, and Thomas E. Borcherding. Ethnic Diversity and Public Goods Provision: Evidence from U.S. Municipalities and School Districts. *Urban Affairs Review*, 52(5):685–713, 2016.

[37] Richard T LeGates and Chester Hartman. The anatomy of displacement in the United States. In *Gentrification of the City*, pages 194–219. Routledge, 2013.

[38] Evan S Lieberman and Gwyneth H Mcclendon. The Ethnicity – Policy Preference Link in Sub-Saharan Africa. *Comparative Political Studies*, 46(5):574–602, 2012.

[39] Christy Lleras. Race, racial concentration,and the dynamics of educational inequality across urban and suburban schools. *American Educational Research Journal*, 45(4):886–912, 2008.

[40] Erzo F P Luttmer. Neighbors as negatives: Relative earnings and well-being. *The Quarterly journal of economics*, 120(3):963–1002, 2005.

[41] Katharyne Mitchell. Networks of ethnicity. *A companion to economic geography*, page 392, 2000.

[42] Kevin M. Murphy and Robert H. Topel. Estimation and inference in two-step econometric models. *Journal of Business and Economic Statistics*, 20(1):88–97, 2002.

[43] Kawika Pierson, Michael L. Hand, and Fred Thompson. The government finance database: A common resource for quantitative research in public financial analysis. *PLoS ONE*, 10(6):1–22, 2015.

[44] Michelle Rendall and Franziska J Weiss. Employment polarization and the role of the apprenticeship system. *European Economic Review*, 82:166–186, 2016.

[45] Albert Saiz. The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296, 2010.

[46] Prerna Singh and Matthias vom Hau. Ethnicity in Time: Politics, History, and the Relationship between Ethnic Diversity and Public Goods Provision. *Comparative Political Studies*, 49(10):1303–1340, 2016.

[47] Daphne Spain. Indicators of urban revitalization: Racial and socioeconomic changes in central-city housing. In *Back to the City*, pages 27–41. Elsevier, 1980.

[48] Jessica Trounstine. Segregation and Inequality in Public Goods. *American Journal of Political Science*, 60(3):709–725, 2016.

[49] David R. Williams and Ronald Wyatt. Racial bias in health care and health: Challenges and opportunities. *JAMA - Journal of the American Medical Association*, 314(6):555–556, 2015.

[50] Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

[51] Yu Zhou. How do places matter? A comparative study of Chinese ethnic economies in Los Angeles and New York City. *Urban Geography*, 19(6):531–553, 1998.

[52] Sharon Zukin. Gentrification: culture and capital in the urban core. *Annual review of sociology*, 13(1):129–147, 1987.