Demographics and the Implications for Voter Participation in US Counties

by

Julia Lewis

A thesis submitted in satisfaction
of the Honors Program option in pursuit of the

Bachelor of Arts

in

Economics

in the

COLLEGE OF LETTERS AND SCIENCE

of the

UNIVERSITY OF CAIFORNIA, BERKELEY

Faculty Advisor:

Professor Hilary Hoynes

April 2019

I would like to extend sincere thanks to Professor Hilary Hoynes, whose time and mentorship were critical to the completion of this project. I would also like to thank Jim Church and Jesse Silva in the UC Berkeley Library Data Lab for their invaluable assistance in the data collection process.

Abstract

This paper examines the effects of inequality, racial fragmentation, and educational fragmentation on voter participation in US counties during national elections spanning from 2010-2016. Data was collected from the American Community Survey 5-year estimates as well as the Dave Leip Atlas of US Elections. An OLS model that incorporates state and year fixed effects is used to estimate the relationship between the three demographic measures and voter participation. The findings suggest that inequality, racial fragmentation, and educational fragmentation have a negative and statistically significant relationship with voter participation.

1. Introduction

American society still suffers from the persistent problems of inequality and social fragmentation (Ryscavage 1999). Average income for the top 1 percent grew by 242% from 1979 to 2015 while incomes for the bottom 2 percent grew only by 79% over the same period (Sherman et al 2018). While the nation grows more diverse, it is not difficult to observe groups living, worshiping, and socializing with others who share similar backgrounds (Costa and Kahn 2003). The haves and the have-nots remain in separate spheres of socialization and interaction (Ryscavage 1999). This phenomenon has implications for both individual relationships within communities and society as a whole, but what do inequality and fragmentation mean for American democracy?

Academics in a variety of fields such as economics, sociology, and political science have all approached the question of how a society can function when there exist social divisions along the lines of income, race, and ethnicity. The theoretical concept of social capital serves to explain why these divisions might be so detrimental to the functioning of society. Social capital refers to the network of relationships and social linkages that enable individuals, groups, and communities to work together and achieve socially efficient outcomes (Keeley 2007). Social capital has been shown as key component to such essential public functions as economic development (Casey and Christ 2005), effective political institutions (Putnam 1993), and the provision of public goods (Anderson et al. 2004). The theory then follows that fragmentation and inequality prevent the social channels and cohesion necessary for a well-functioning society. For example, Alesina and La Ferrara (2000) construct an index of social capital that incorporates measures on trust, membership in groups, and voting behavior and examine the distribution across US states. They find that the index is highest in the North/Northwest, regions generally characterized by their racial homogeneity and lower levels of income inequality, and lowest in the South/Southwest,

where racial heterogeneity and income inequality are higher (Alesina and La Ferrara 2000). The variation in levels of social capital follows similar patterns as the demographic characteristics across the US. It is therefore hypothesized that through the avenue of social capital, community heterogeneity influences social, political, and economic outcomes.

Of course, one cannot fail to acknowledge the benefits of heterogeneity in communities. Diversity in classrooms often provokes an elevated level of discourse that draws from a wide variety of backgrounds and bolsters the learning outcomes of students. Cities often appreciate the presence of different cultures in the diversity of culinary options available and in the presence of varied architectural style that give cities charm and artistic flare. There have also been connections drawn between diversified cities and innovation, leading to economic success (Duranton and Puga 2001). Despite the benefits of a diverse community, however, there still exists the persistent tendency for people to prefer being around others like themselves. In heterogenous communities, this preference can lead to fewer linkages among groups. Fewer social linkages then prevent the engagement necessary for optimal social outcomes.

In this paper, I will build on previous research that considers how inequality and community fragmentation impact social outcomes. In this case, I will use measures studied previously, such as the Gini index of income inequality and an index of racial fragmentation, as well as new measures, such as an index of educational fragmentation. I will examine US counties in election years from 2010-2016, using a fixed effects model to estimate how voting behavior changes based on the demographic characteristics of a community. The findings suggest that racial and educational fragmentation lower voter participation by a notable degree, by approximately 2 percentage points for interquartile range increases in inequality and racial fragmentation and by 2.6 percentage points for an interquartile range increase in educational fragmentation.

The remainder of the paper is organized in the following structure. Section 2 will give an overview of the literature on the subject of heterogeneity and community effects. Section 3 describes the sources and nature of the data used in this study and describes the methodology used to construct the relevant indices and build the fixed effects regression model. Section 4 presents and interprets the regression results from the fixed effect model. Finally, Section 5 draws conclusions from these results, suggests further avenues for research, and provides suggestions for overcoming the negative effect of fragmentation on voter participation.

2. Literature Review

The literature surrounding community effects on social capital and civic engagement has expanded widely in the last 30 years. The literature both links social capital to positive outcomes and demonstrates how characteristics of communities can prevent the linkages necessary for social capital. The focus in the literature then extends to capture how demographic characteristics, such as ethnic and racial fragmentation as well as inequality, affect social outcomes. The outcomes under investigation span from participation in groups (Alesina and La Ferrara 2000, Costa and Kahn 2003) to rates of volunteering (Costa and Kahn 2003, Rotolo and Wilson 2014), to public trust (La Porta et al. 1997, Glaeser et al. 2000, Costa and Kahn 2003), the provision of public goods (Alesina et al. 1999, Vigdor 2004), economic outcomes (Casey and Christ 2005), and voting (Costa and Kahn 2003, Branton 2004, Cho et al. 2006). Generally, much of the research concludes that fragmentation and inequality lead to less efficient outcomes with detrimental effects in the public sphere. The remainder of this section summarizes the studies that are closest to my analysis and consider how fragmentation and inequality affect civic engagement.

Alesina and La Ferrara (2000) investigated community heterogeneity by studying how participation in communities varies with different measures of fragmentation. They hypothesize

that more heterogeneous communities participate less in groups, with groups including a wide array of membership-based activities, such as church groups, service groups, and unions. Alesina and La Ferrara (2000) note that participation in groups is correlated with political participation, and in this way fragmentation effects can impact civic engagement and society more generally. In this study, they examine "Metropolitan Sampling Areas," using data from the General Social Survey and the Current Population Survey for the years 1974-1994 as well as the 1990 Census (Alesina and La Ferrara 2000). Additionally, the measures for community heterogeneity include the Gini index and fragmentation indices for race and ethnicity, calculated by the equation

$$Race_i = 1 - \sum_k s_{ki}^2$$

which I also adopt here. Alesina and La Ferrara (2000) find negative and statistically significant effects for each of the three key measures of inequality and fragmentation on group participation. The findings support their hypothesis that people in more unequal and more heavily fragmented communities participate less in groups.

Costa and Kahn (2003) find similar results in their study of civic engagement and community heterogeneity. They argue that while there are noted benefits of community diversity, heterogeneity also imposes costs and leads to variation in social interactions depending on whether the surrounding environment is heterogeneous or homogenous. They again focus on income, race, and ethnicity but widen the scope of measures of civic engagement to include volunteering, membership, trust and voting. They look at metropolitan areas for the measures of volunteering, membership, and trust, and census tracts in the 1998 and 2000 general elections for the data on voting. This study found that the Gini index was a statistically insignificant predictor of participation and trust in the community, and racial fragmentation was a significant predictor only of lower rates of volunteering (Costa and Kahn 2003). Ethnic fragmentation, what the authors refer to as birthplace fragmentation, was the most robust measure in the study and

predicted lower rates of volunteering, group membership, and trust. Costa and Kahn (2003) then turn their attention to voter participation as a measure of community participation, arguing that since voting is costly and it is hard to determine the effect of one's individual vote, it is a wonder why everyone does not free ride in elections. They expect to find that in communities with more social capital, that is, in more homogenous communities, there is a greater civic sense as residents pursue goals beyond their own self-interest and vote more consistently. They find results to support this theory. In their model, voter turnout rates decline by two percentage points for a one standard deviation increase in ethnic fragmentation. They do not consider the other measures for inequality and racial fragmentation in the analysis of voting trends.

My own study expands on the current body of research in a few ways. Firstly, although many studies have considered inequality on the basis of income and fragmentation on the basis of age, sex, race, and ethnicity, none have considered the effect of educational fragmentation. Fragmentation on the basis of education presents an alternate way of measuring inequality in a population since educational attainment is so closely tied with future potential earnings (Autor 2014). Additionally, it has been indicated that people associate and form networks with those who share similarities along common dimensions, including education (Bottero 2007). Therefore, fragmentation on the basis of education not only indicates inequality, but also fewer linkages with others in the community if it is heavily fragmented by education (Bottero 2007). Since social linkages are argued to be the characteristic of homogenous communities that prompt them to participate more in elections (Costa and Kahn 2003), it is therefore hypothesized that educational fragmentation depresses participation in elections.

This study also builds upon past research by considering county level data in more recent elections. After the initial burst of studies in the late 1990s and early 2000s, comparatively less research has been done in recent years to examine demographic effects on community

engagement, especially on the dimension of voting. Additionally, a majority of studies focus on cities or metropolitan areas (Alesina and La Ferrara 2000, Costa and Kahn 2003, Rotolo and Wilson 2014), while my analysis includes counties from 49 of the 50 states, excluding Alaska, and from counties that the US Census classifies as "small"[1]. County level data gives the additional benefit of capturing metropolitan and rural areas and are defined consistently across all years. Thus, this study expands on the findings in the literature that examine whether heterogeneity and fragmentation lead to less civic engagement, as measured by voter participation.

3. Data and Empirical Methodology

a. Data Sources and Description

Because this study combines variables in a new way and looks at a different geographical unit, US counties, the dataset had to be self-constructed using various sources of voting and demographic data. The dataset draws election data from Dave Leip's Atlas of U.S. Elections and demographic data from the American Community Survey in the Congressional elections in the years 2010, 2012, 2014 and 2016.

Dave Leip's Atlas provides election data dating back as far as 1912 until 2018, and data for the four recent years under investigation was extracted from the master database. The datasets collected provide total number of votes, margin of victory totals and percentages, and party vote counts for states, counties, towns, and congressional districts. The number of votes in each county in each year were essential in the construction of the independent variable voter participation. To determine the voter participation rate, I divided total votes for the House of

[1] The US Census classifies "small" counties as those with 48 people per square mile or fewer. Small counties represent roughly 95% of total counties while less than half of US citizens reside in small counties. (US Census Bureau 2017).

Representatives elections by the total voting age population in each county, which was takes from the American Community Survey data.

Some limitations arose in the use of the Dave Leip Atlas data as there were 29 data points across the four years of data for which the vote counts were not listed. The missing data came mostly from the year 2010 and were clustered predominantly in one state, Oklahoma. There was no explanation provided for the missing data, but since the number of counties for which data is missing represents a very small portion of the overall sample, it was reasonable to drop the counties in every year for which any voting data was missing.

The data for the key independent variables as well as the control variables came from the American Community Survey 5-year estimates, which began to be released in 2009. The American Community Survey (ACS) is an ongoing survey administered yearly by the US Census Bureau that provides detailed demographic data about people and US geographies. The ACS 5-year estimates aggregate over the entire 5-year period of available data and represent an average across the period[2]. As a result, the 5-year estimates do not provide the highest degree of currency compared with the 1- and 3-year estimates. However, the ACS 5-year estimates include all US geographies, even those with populations below 20,000, which is critical to the analysis of voting in counties because many have populations below that margin. The ACS 5-year estimates also give the most reliable estimates of all variables since aggregating over 5 years provides a larger sample size. The ACS provides all of the key variables necessary for examining income, racial, and educational heterogeneity and fragmentation as it provides data for the Gini index, population by race, and educational attainment. Additionally, the ACS gives the population by age necessary for the calculation of the voting age population and the voter participation rate.

[2] Because the ACS 5-year estimates average across the preceding 5-year period and there are only 2 years between the elections under investigation, the county measures overlap across years. For example, the 2014 estimates included data from 2010-2014. Therefore, the data may not precisely pick up year to year variation, but the 5-year estimates still provide the benefit of including accurate measures counties with population under 65,000 persons.

Because the ACS provides so many demographic measures, it also offered the information for median household income and population sizes necessary for control variables.

To create the working dataset, the ACS and Dave Leip election data were merged based on county, identified by each county's unique fips code, and sorted by year. One limitation arose in the merging of both datasets as it became apparent that the geographic units of measurement in the ACS and election data for Alaska did not align. As a result, the state of Alaska was dropped from the dataset. Similar to the counties with missing voting data, the number of counties and number of people in Alaska is very small in comparison with the remainder of the dataset. For this reason, it was reasonable to exclude Alaska from the study.

With the data acquired from the Dave Leip Atlas and the American Community Survey and the dataset constructed, I then constructed the indices for the measures of racial and educational fragmentation. Because the ACS already provided data for the Gini index of income inequality, I did not have to conduct further calculations for that index. The racial and educational fragmentation indices were created for each county in every year considered by this study. The racial fragmentation index (*Race*) was constructed using ACS data and the following formula, which appears in numerous studies on the topics of heterogeneity and fragmentation (Alesina and La Ferrara 2000, Costa and Kahn 2003, etc):

$$Race_i = 1 - \sum_k s_{ki}^2$$

where *k* represents the following races: a) White, b) Black or African American, c) American Indian and Alaska Native, d) Asian, e) Native Hawaiian and Other Pacific Islander, f) Other, and g) Two or More Races. Each term *s* represents the share of race *k* in the county *i* population. The racial fragmentation index tells the "probability that two randomly drawn individuals in county *i* belong to different races" (Alesina and La Ferrara 2000). Therefore, higher measures of the

index signal more racial fragmentation, with 1 representing full fragmentation and 0 representing total homogeneity.

The educational fragmentation index (*Education*) was constructed also using data from the ACS that gives the educational attainment for the population 25 years and older, which it breaks into the categories a) less than high school, b) high school graduate, c) some college, d) bachelor's degree, e) master's degree, f) professional school degree, and g) doctorate degree. I employed a similar formula as used for the racial fragmentation index, except replacing race with educational attainment. The formula for the educational fragmentation index is the following:

$$Education_i = 1 - \sum_k s_{ki}^2$$

where *k* represents the following levels of educational attainment: a) less than high school, b) high school graduate and some college, c) college or more, including masters, doctoral and professional school degrees, combining categories in b) and c) to reflect similar levels of educational achievement. Each term *s* represents the share of the people in the county population *i* who have attained that level of education. The educational fragmentation index tells the "probability that two randomly drawn individuals" in county *i* have achieved different levels of education (Alesina and La Ferrara 2000). Similar to the index of racial fragmentation, a higher measure for the index implies higher educational fragmentation.

b. Methodology

This section will give a detailed account of the construction of the fixed effects regression model. From the theoretical claims that inequality and social fragmentation along demographic lines produce worse community outcomes and less public engagement, I hypothesize that higher income inequality, racial fragmentation, and education fragmentation produce lower voter participation in counties. Under this hypothesis, more homogenous counties should see higher

participation in elections. To evaluate this claim and measure the effect of inequality and fragmentation, I estimate voter participation as a function of the three measures of inequality and fragmentation, first separately and without control variables. To control for time independent variation in counties on account of their location in a particular state, state fixed effects were included. Similarly, to control for year to year variation, year fixed effects were also included. The model also recognizes the fact that counties can vary widely in population size, and thus a minority of counties holds the majority of voters. To account for this variation in county size, I weight the model by county population, placing larger weight on more populated counties. The basic regression equations for each of the three key measures are of the following form:

$$(1)\ VoterPartic_{it} = \alpha_0 + \beta_1 Gini_{it} + \gamma_s + \delta_t + \beta_2 X_{it} + u_{it}$$

$$(2)\ VoterPartic_{it} = \alpha_0 + \beta_1 Race_{it} + \gamma_s + \delta_t + \beta_2 X_{it} + u_{it}$$

$$(3)\ VoterPartic_{it} = \alpha_0 + \beta_1 Education_{it} \gamma_s + \delta_t + \beta_2 X_{it} + u_{it}$$

where $VoterPartic_{it}$ is the voter participation rate in a given county *i* in a given year *t*, $Gini_{it}$ in equation (1) represents the Gini index of income inequality, $Race_{it}$ is the racial fragmentation index in equation (2), and $Education_{it}$ the educational fragmentation index in equation (3), and u_{it} is the error term. State fixed effects are represented by the γ_s [3] terms while year fixed effects are represented by the δ_t terms. If inequality and fragmentation lower civic engagement and therefore participation in elections, the β_1 term on each of the indices in the models above should be negative.

In some models, I control for other county by time variables, represented by X_{it} in equations (1)-(3). I introduce control variables for total population and median household

[3] While the panel data is by county, the model includes fixed effects for states rather than counties as there are only four periods in the panel data and controlling for county fixed effects strips the variation out of the data needed for analysis. While using state rather than county fixed effects is most sensible for this analysis, it still presents a limitation. Further research should strive to control for county fixed effects.

income, expressed in logs, to examine the sensitivity of the regression results in the presence of these controls and fixed effects. I incorporate a control for population size, as has previous research, under the assumption that fragmentation may be higher in larger counties, especially urban counties that may be already fragmented with the existence of segmented neighborhoods (Alesina et al 1999). The β_2 term on the log of the total population should also be negative under this assumption. I also control for median household income[4], another common demographic control in the literature (Alesina et al 1999, Alesina and La Ferrara 2000, Costa and Kahn 2003). In this case, I expect that, since wealthier counties likely have more resources to go into voting efforts and wealthier citizens may have more time to follow elections and vote, the coefficient on the median household income should be positive. In effect, wealthier counties should vote more.

I also consider regressions that incorporate two or all three of the key independent variables in order to determine the robustness of each index in the presence of others. Table 1 gives summary statistics for each of the indices and each control variable incorporated in the models.

Table 1: Summary Statistics

	N	**Mean**	**Standard Deviation**	**Min**	**Max**	**25th Percentile**	**75th Percentile**
Gini	12416	0.438	0.036	0.207	0.652	0.414	0.460
Race	12416	0.230	0.169	0.000	0.745	0.083	0.367
Education	12416	0.767	0.045	0.292	0.958	0.738	0.797
Voter Participation	12416	0.469	0.127	0.081	1.590	0.376	0.560
Total Population	12416	10.278	1.46	3.714	16.124	9.329	11.114
Median Household Income	12416	10.782	.247	9.851	11.85	10.625	10.921

* Total Population and Median Household Income expressed as logs

[4] The ACS gives median household income in inflation adjusted dollars for the corresponding year, i.e. median household income reported in the 2010 ACS 5-year estimates are expressed in 2010 inflation adjusted dollars. Median household income for the purposes of this study were converted to 2016 inflation adjusted dollars using data from the US Inflation Calculator, a tool of the US Department of Labor Bureau of Labor Statistics.

4. Results

Table 2 presents the OLS regression results from equations (1)-(3) that look one at a time at the key independent variables without control variables but still including fixed effects for state and year.

Table 2: Regression Results on Voter Participation, without Control Variables

	(1)	(2)	(3)
Gini Index	-0.281***		
	(0.036)		
Race Fragmentation		-0.174***	
		(0.006)	
Education Fragmentation			-0.687***
			(0.026)
Constant	0.537***	0.455***	0.947***
	(0.016)	(0.002)	(0.020)
Obs.	12416	12416	12416
R-squared	0.609	0.632	0.652
State Fixed Effects	YES	YES	YES
Year Fixed Effects	YES	YES	YES
Robust standard errors are in parenthesis *** p<0.01, ** p<0.05, * p<0.1			

The initial results from Columns (1)-(3) in Table 2 support the hypothesis that inequality and fragmentation on the basis of race and education lower the rate of voter participation. Coefficients for each of the three measures are negative and statistically significant at the 1% significance level. In order to interpret these findings, it is necessary to understand what a reasonable change in each variable is as a change in any one of the indices by 1 would represent an unrealistic jump from total equality to total inequality or total integration to total fragmentation. To ascertain a relevant scale, I consider the movement of each of the variables during the 6-year period under investigation and also the interquartile range. Table 3 provides the relevant information for this analysis.

Table 3: Summary Statistics, 2010, 2016 & All Years

	Mean	Standard Deviation	25th Percentile	75th Percentile
2010				
Gini	0.432	0.036	0.407	0.454
Race	0.228	0.170	0.078	0.372
Education	0.774	0.044	0.745	0.804
2016				
Gini	0.445	0.035	0.421	0.465
Race	0.235	0.168	0.088	0.370
Education	0.762	0.045	0.731	0.791
All Years				
Gini	0.438	0.036	0.414	0.460
Race	0.230	0.169	0.083	0.367
Education	0.767	0.045	0.738	0.797

Over the six-year period from 2010-2016, the mean of the Gini index of income inequality increased by 0.013 from 0.432 to 0.445. Also, over the entire period, the interquartile range of the Gini index is 0.046. Rounding for convenience, it reasonable to expect a change in the Gini of this magnitude since 0.05 represents an approximate for the interquartile range and a change in the Gini index that could realistically occur over a 30-year period if it continues to increase at the current rate of 0.013 over 6 years. Column (1) in Table 2 reveals that a change in the Gini index of 0.05 gives a change in the voter participation rate of $0.05 * -0.281 = -0.014$, or a predicted 1.4 percentage point decrease in the voter participation rate for a 0.05 increase in the Gini index. Compared to the mean of voter participation of 0.469, as shown in Table 1, the 1.4 percentage point decrease in the reflects a 2.98% change compared to the mean of voter participation.

Conducting a similar analysis for the racial fragmentation index, one can see that over the six-year period of the study that there has been only a small increase in the mean of the index, from 0.228 to 0.235. There is, however, wider variation between the 25th and 75th percentiles over the entire period, giving an interquartile range of 0.284. Rounding conservatively to 0.25 for

ease of calculation, I will use a change in the racial fragmentation index of 0.25 to interpret the corresponding regression coefficients. Column (2) in Table 2 shows that a change in the racial fragmentation index of 0.25 gives a change in the voter participation rate of $0.25 * -0.174 = -0.044$, or a predicted 4.4 percentage point decline in the voter participation rate for a .25 increase in racial fragmentation index. This percentage point decline represents 9.38% of the mean of voter participation of 0.469.

Applying the same methodology to the educational fragmentation index, one can observe a decline in the mean of the index by 0.012 over the six-year period of the study. The index's interquartile range is 0.059. Rounding down to 0.05 for convenience of calculation, I will use a change in the educational fragmentation index of 0.05 to interpret the corresponding coefficients. Much like for the Gini index, it is reasonable to expect a change in the educational fragmentation index of this scale because 0.05 represents not only the interquartile range but also a realistic change over a 30-year period if the increases keep the current rate of 0.012 over 6 years. Using this means of interpretation, Column (3) in Table 2 tells that a change in the educational fragmentation index of 0.05 gives a change in the voter participation rate of $0.05 * -0.687 = -0.034$, or a predicted 3.4 percentage point decline in the voter participation rate. Considering the mean of voter participation, there is a 7.2% decline in the voter participation rate relative to the mean of 0.469, or 46.9%, for an interquartile increase in educational fragmentation.

Table 4 gives the regression results for the model while introducing control variables and while considering the indices jointly.

Table 4: Regression Results on Voter Participation, with Controls and Key Variables Together

	(1)	(2)	(3)	(4)
Gini Index	0.145**			0.249***
	(0.062)			(0.074)
Race Fragmentation		-0.081***		-0.073***
		(0.013)		(0.015)
Education Fragmentation			-0.516***	-0.475***
			(0.042)	(0.050)
Total Population	-0.020***	-0.012***	-0.011***	-0.009***
	(0.001)	(0.002)	(0.001)	(0.002)
Median Household Income	0.185***	0.165***	0.165***	0.172***
	(0.008)	(0.007)	(0.006)	(0.008)
Constant	-1.471***	-1.271***	-0.891***	-1.112***
	(0.092)	(0.064)	(0.065)	(0.090)
Obs.	12416	12416	12416	12416
R-squared	0.798	0.801	0.807	0.812
State Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1
Total Population and Median Household Income expressed as logs

All measures are significant at either the 1% or 5% significant levels. The coefficient on racial fragmentation, Column (2) in Table 4, weakens in the presence of controls to -.081, reflecting a predicted 2.02 percentage point decline in voter participation for an interquartile change in the racial fragmentation index, compared to a 4.4 percentage point decline without controls. The 2.02 percentage point decline still represents a predicted 4.31% decline in relation to the mean of voter participation. The coefficient on educational fragmentation also weakens in the presence of controls, although both racial and education fragmentation indices remain significant at the 1% significance level both with and without control variables included in the model. In the presence of control variables for total population and median household income, the coefficient on the

educational fragmentation index is -0.516, down from -0.687 with controls. This coefficient represents a predicted 2.58 percentage point decline in voter participation for an interquartile increase in educational fragmentation, corresponding to a predicted 5.50% decrease relative to the mean of voter participation. When all three key independent variables are considered together in Column (4) in Table 4, the coefficients on racial and educational fragmentation decrease again but by a smaller degree than when the control variables were introduced. In this case, the racial fragmentation index predicts a 3.89% decrease in the voter participation rate for an interquartile increase in the index relative to the mean of voter participation while the educational fragmentation index predicts a 5.06% decrease in the voter participation rate relative to the mean for an interquartile increase in the index. While the effects of fragmentation decline in each variation of the model, they still impact voter participation to a notable degree, especially when considering turnout in the 2010 midterm election was only 37% (Desilver 2014).

Noticeably, the coefficient on the Gini Index becomes positive in the presence of control variables in Column (1) in Table 4 and remains positive in the presence of the other key variables in Column (4). A positive sign on the coefficient for the Gini index would suggest that higher levels of inequality would prompt increases voter turnout, which conflicts with the theory that inequality and heterogeneity prevent the creation of social capital necessary for civic engagement. One possible explanation for the positive coefficient on the Gini in the presence of controls is its relationship with median household income. One could imagine that counties with higher median household income have lower inequality as people in the county are more effectively sharing in the wealth. A negative correlation of -0.407 between the Gini index and the logarithm of median household income would support this claim. To see how the results differ for the Gini index when the control for median household income is removed, Table 4

regressions that include the Gini index are recreated without the control variable for median household income. The new regressions presented in Table 5.

Table 5: Regression Results, Considering Gini without Median Household Income

	(1)	(2)
Gini Index	-0.392***	-0.203***
	(0.053)	(0.061)
Race Fragmentation		-0.103***
		(0.017)
Education Fragmentation		-0.588***
		(0.049)
Total Population	-0.005***	0.008***
	(0.001)	(0.001)
Constant	0.620***	0.879***
	(0.024)	(0.051)
Obs.	12416	12416
R-squared	0.734	0.757
State Fixed Effects	YES	YES
Year Fixed Effects	YES	YES
Robust standard errors are in parenthesis		
*** p<0.01, ** p<0.05, * p<0.1		

Reconsidering the relevant regressions for the Gini index without the control for median household income corrects the positive sign that was previously present. Results for the Gini index in Table 5 are not only negative but also statistically significant at the 1% significance level, even when considered alongside the indices for racial and educational fragmentation. Column (1) in Table 5 now suggests that for an interquartile range increase in the Gini index, there is a predicted 1.96 percentage point decrease in the voter participation rate, all else equal. Additionally, while the Column (2) in Table 5 suggests a smaller negative effect of the Gini index, it still reflects a predicted 2.16% decline in voter participation relative to the mean of participation for an interquartile increase in the Gini index. Therefore, the regression results without controlling for median household income strengthens the conclusion drawn for the Gini index that inequality does lower participation on the behalf of voters.

Column (4) in Table 4 and Column (2) in Table 5 both combine all of the three key indices. The magnitude of each index is smaller in the regressions in which they are considered together than when they are considered separately. The correlations among the three indices, reported in Table 6, could help explain this phenomenon. However, because the correlations among the key indices are relatively weak, it is likely not necessary to consider each index separately. Examining the effects of the indices when used jointly in regressions may give a better indication of which one has the largest impact. In this case, Column (2) in Table 5 suggest that educational fragmentation may have the largest impact on voter participation as it tells that for an interquartile increase in educational fragmentation there is a predicted 2.94 percentage point decline in voter participation, compared to a predicted 1.02 percentage point decline in participation for an interquartile increase in the Gini index and a predicted 2.58 percentage point decline for an interquartile increase in racial fragmentation, all else equal. This result is interesting as this paper is among the first to consider education as a dimension for fragmentation in communities.

Table 6: Correlations

Variables	Voting	*Gini*	*Race*
Gini	-0.155		
Race	-0.281	0.365	
Education	-0.359	0.281	0.468

Additionally, while the effects of the fixed effects have not been reported in the regression tables, the year fixed effects are statistically significant at the 1% significance level and exhibit a trend as one might expect when considering election years. In each regression, the coefficient terms for 2012 and 2016 are positive, indicating that there was an additional increase in the voter participation rate in 2012 and 2016 relative to the omitted. Elections in which voters decide the presidency historically have higher turnout than in midterm years, for example the

37% national turnout among eligible voters in 2010 compared with the 54% national turnout in 2012 (DeSilver 2014). The magnitude of the predicted effect of being in 2012 or 2016 ranges from 0.12 to 0.15 in each regression, which remains in line with the roughly 10% to 20% variation between midterm and presidential election years.

Overall, the regression results support the hypothesis that fragmentation on the basis of race and education lower civic engagement, as measured by voter participation. Each of the regression results for the indices alone confirm that there is both a statistically significant negative and sizeable effect on voting. Most notably, the Column (3) in Table 4 predicts a 2.6 percentage point decrease in the voter participation rate for a 0.05 increase in the educational fragmentation index, all else equal. This result is noteworthy not only because this study is among the first to consider educational fragmentation as a measure of heterogeneity and inequality in this context, but also because it had the most sizeable effect on voting when compared with the other two indices, as shown in Column (4) in Table 4 and Column (2) in Table 5.

5. Conclusion

There is much to celebrate as America grows increasingly diverse, yet this study demonstrates that growing heterogeneity in the US could have negative implications for the future of voting and democracy. In this paper, I used inequality, race, and education as three measures along which the US has been and continues to be fragmented. The results are clear that communities participate less in the public sphere when they face fragmentation along demographic lines.

There are some questions that remain, however. What are the personal determinants of an individual's decision to participate? How might these individual level factors interact with

community level measures of fragmentation to impact civic engagement? Would voting patterns change along the lines of the model if, say, a large number of immigrants entered a once homogenous community? Further research on this topic could provide insight on not only these questions but also could look more closely at the relationship between heterogeneity and social capital. There must be situations in which people of different income levels, races, and educational backgrounds interact to produce beneficial outcomes in the public domain. Under what circumstances could these interactions take place to form social networks, produce social capital, and boost participation in the community? While many questions have been posed in the literature about heterogeneity and fragmentation in general, further investigation into mechanism behind social capital formation, or lack thereof, in heterogenous communities could reveal more as to why this phenomenon exists. Additionally, this study presents the finding that educational fragmentation has a greater effect on voter participation than income inequality or racial fragmentation. Since examining fragmentation on the basis of education is new in the literature, further studies that consider the effects of educational fragmentation on broader measures of civic engagements would be helpful to determine why and how it impacts communities in such a way.

While many questions remain that are outside of the scope of this study, the results from this paper do help us draw conclusions about voting and government today. As campaigns often try to mobilize voters and "get out the vote," these results inform where they may want to concentrate more of their efforts. If "get out the vote" efforts can increase turnout by roughly 2 percentage points in the most diverse and fragmented communities, they may be able to overcome the tendency in those communities to stay home from the voting booths. Since diverse communities are also the ones suffering from lower turnout, efforts in these spaces would have the additional benefit of giving voting power and an electoral voice to many diverse groups that

are not always represented in government. Additionally, if governments are interested in tackling low turnout, the findings in this paper suggest that they may want to focus efforts on easing inequality and bridging the gap between segmented communities. I have presented evidence that heterogenous communities participate less in elections, which can go on to have negative implications for equal representation. Armed with this knowledge, it is the responsibility of elected officials, campaigns, and citizens themselves to acknowledge these effects and work toward resolving these differences across communities.

References

Alesina, Alberto, and Eliana La Ferrara. "Participation in Heterogeneous Communities." *The Quarterly Journal of Economics* 115, no. 3 (August 2000): 847-904. doi:10.3386/w7155.

Alesina, Alberto, Reza Baqir, and William Easterly. "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics* 114, no. 4 (November 1999): 1243-284. doi:10.1162/003355399556269.

Anderson, Lisa R., Jennifer M. Mellor, and Jeffrey Milyo. "Social Capital and Contributions in a Public-Goods Experiment." *American Economic Review* 94, no. 2 (May 2004): 373-76. doi:10.1257/0002828041302082.

Autor, D. H. "Skills, Education, and the Rise of Earnings Inequality Among the 'Other 99 Percent.'" Science 344, no. 6186 (May 22, 2014): 843–851.

Bottero, W. (2007), Social Inequality and Interaction. Sociology Compass, 1: 814-831. doi:10.1111/j.1751-9020.2007.00030.x

Branton, Regina P. "Voting in Initiative Elections: Does the Context of Racial and Ethnic Diversity Matter?" *State Politics & Policy Quarterly* 4, no. 3 (2004): 294-317. doi:10.1177/153244000400400303.

Casey, Terrence, and Kevin Christ. "Social Capital and Economic Performance in the American States*." *Social Science Quarterly* 86, no. 4 (December 2005): 826-45. doi:10.1111/j.0038-4941.2005.00358.x.

Cho, Wendy K. Tam, James G. Gimpel, and Joshua J. Dyck. "Residential Concentration, Political Socialization, and Voter Turnout." *The Journal of Politics* 68, no. 1 (February 2006): 156-67. doi:10.1111/j.1468-2508.2006.00377.x.

Coleman, James S. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press, 1990.

"Consumer Price Index Data from 1913 to 2019." US Inflation Calculator. April 10, 2019. Accessed April 25, 2019. https://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/.

Costa, Dora L., and Matthew E. Kahn. "Civic Engagement and Community Heterogeneity: An Economists Perspective." *Perspectives on Politics* 1, no. 1 (March 2003). doi:10.2139/ssrn.311839.

DeSilver, Drew. "Voter Turnout Always Drops off for Midterm Elections, but Why?" Pew Research Center. July 24, 2014. Accessed April 15, 2019. https://www.pewresearch.org/fact-tank/2014/07/24/voter-turnout-always-drops-off-for-midterm-elections-but-why/.

Duranton, Gilles, and Diego Puga. "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products." *American Economic Review* 91, no. 5 (December 2001): 1454-477. doi:10.1257/aer.91.5.1454.

Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman, and Christine L. Soutter. "Measuring Trust." *Quarterly Journal of Economics* 115, no. 3 (August 2000): 811-46. doi:10.1162/003355300554926.

Keeley, Brian. *Human Capital: How What You Know Shapes Your Life*. Paris: Organisation for Economic Co-operation and Development, 2007.

La Porta, Rafael, Florencio Lopez-De-Silane, Andrei Shleifer, and Robert W. Vishny. "Trust in Large Organizations." *American Economic Review* 82, no. 2 (May 1997): 333-38. doi:10.3386/w5864.

Leip, Dave, “Dave Leip’s Atlas of U.S. Presidential elections, datasets,” accesses through UC Berkeley Library, Department of Government Information, California & United States.

Putnam, Robert D., Robert Leonardi, and Raffaella Nanetti. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton Univ. Press, 1993.

Rotolo, Thomas, and John Wilson. "Social Heterogeneity and Volunteering in U.S. Cities." *Sociological Forum* 29, no. 2 (June 2014): 429-52. doi:10.1111/socf.12091.

Ryscavage, Paul. *Income Inequality in America: An Analysis of Trends*. Armonk, NY: M.E. Sharpe, 1999.

Sherman, Arloc, Chad Stone, Roderick Taylor, and Danilo Trisi. "A Guide to Statistics on Historical Trends in Income Inequality." Center on Budget and Policy Priorities. December 11, 2018. Accessed April 19, 2019. https://www.cbpp.org/research/poverty-and-inequality/a-guide-to-statistics-on-historical-trends-in-income-inequality.

US Census Bureau. *Selected demographic characteristics*, 2006-2016 *American Community Survey 5-year estimates*. Retrieved from https://www.socialexplorer.com/explore-tables.

US Census Bureau. "Big and Small America." The United States Census Bureau. October 24, 2017. Accessed April 10, 2019. https://www.census.gov/library/stories/2017/10/big-and-small-counties.html.