Civic Identity and Library Proximity: Measuring the Impact of Libraries in California on Political Participation and Engagement

Matthew L. Sexton[†]

Abstract

As institutions that provide individuals with opportunities for communication and the ability to research, libraries help to promote civic engagement within their communities. Utilizing branch library and voting data from the State of California, this study examines the effect that the introduction of a new library in a county has on political participation, focusing specifically on voter registration, presidential election voting, and political party affiliation. Through a fixed effects methodology, the main results of this study show a statistically significant connection between libraries and political participation with regard to registration data. In the average county, a new library corresponds to a 0.737 percentage point increase in the percent of registered voters, a 0.471 percentage point increase in the percent of registered Democrats, and a 0.254 percentage point decrease in the percent of registered Republicans. Considering the closeness of recent national elections, these registration choices have the potential to alter America's political landscape. Though political socialization and education have been examined in the past for their impact on political participation, libraries have not been explored through a similar lens. As such, these results not only provide new information to the discipline, but also provide a way to increase a community's civic engagement and affiliation with a political ideology. Thus, understanding this impact both furthers research efforts regarding political participation and allows us to have a better grasp on how we might control legislation within a given region.

* †

*I would like to thank the entire Department of Economics and Accounting for providing me with the opportunity to pursue research on a deeper level through this thesis. In particular, I would like to thank Professors Robert Baumann, Katherine Kiel, and Justin Svec, whose courses provided me with the empirical understanding required of economics research. I would also like to give my appreciation to Professors Monica Carney and Joshua Congdon-Hohman for helpful ideas while establishing a topic. Finally, I would like to extend my greatest thanks to my advisor, Professor Melissa Boyle, for constant guidance, emotional support, and invaluable advice. I am proud to call her both a mentor and a friend, as both this thesis and my education would be incomplete without her.

†Bachelor of Arts, Department of Economics and Accounting, College of the Holy Cross, 1 College Street, Worcester, MA 01610, USA, E-Mail: mlsext22@g.holycross.edu

I Introduction

Libraries, by nature, serve to provide the public with free access to knowledge that would be inaccessible otherwise. These institutions not only create "opportunities for learning" and support, but also "help [to] shape the new ideas and perspectives that are central to a creative and innovative society" (White, 2012). Having access to libraries and obtaining this support is critical, especially to those from low-income backgrounds. As of 2020, 37.2 million individuals in the United States were living at or below the poverty line, equating to over 11 percent of the population (*National Poverty in America Awareness Month*, 2022). With such a large amount of the US being affected, ensuring that individuals in this group have access to the resources they need is crucial to the country's overall wellbeing. On a similar note, individuals in this low-income bracket tend be less likely to vote than those from higher income brackets, with Hartley (2020) finding that eligible low-income voters are 22 percentage points less likely to vote in elections at the national level compared to their higher-income peers. While this partially occurs due to transportation and accessibility problems, a major source of this lowered voter turnout is due to a lack of interest in the issues or candidates on the ballot (Hartley, 2020). In theory, making sure that these individuals have access to the knowledge necessary to inform themselves about the issues should push help them to engage more civically. Considering this, understanding the impact of library accessibility on an individual's political participation is of utmost importance to improving civic engagement in the US.

To fully comprehend why libraries are so influential, it is necessary to consider the vast number of resources that these institutions make available to the public. Typically, when thinking about library resources, the first thing that comes to a person's mind is books. While it is true that libraries have a plethora of books available—varying by age, genre, theme, and discipline—these institutions have so much more to offer. In recent years, libraries have begun offering services to help expand access to the internet. These include in-house computers for the community to use, laptop rentals for individuals to work on either at home or in the library, free Wi-Fi, and a multitude of other resources (*Library Services in the Digital Age*, 2020).

Libraries also provide a variety of services meant to benefit the community for children and adults alike. While children might be able to access early education programs, daily read-alouds, or afterschool daycare sessions, adults can attend book clubs, search for employment opportunities, and even access programs directly targeted at senior or immigrant populations (*Libraries Foster Community Engagement*, 2022). It is also pertinent to remember that libraries foster community; they allow individuals to ask questions and get to know one another, rather than just serving as mere book-lenders. With the various programs, services, and communication opportunities libraries provide, it is easy enough to view them as being the pillars of a community (*Libraries Foster Community Engagement*, 2022).

Some have argued that the US has witnessed a declining need for libraries. A 2015 Pew Research study noted that while 78% of Americans surveyed had been to a public library during their life, just 44% had visited a public library in the prior year, compared to 53% in a similar survey given three years earlier (*Library Services in the Digital Age*, 2020). Data from the Institute for Museum and Library Services further confirms these findings, noting a continued decline in library visits of roughly 3% per year from 2012 to 2019 in the US (*Healthier, More Equitable America*, 2021). These statistics, however, completely disregard the non-physical aspects of libraries used by the population. In the current world, most libraries have begun offering digital components to their arsenal including the devices discussed previously and digital e-books. In a country where over 13 percent of households lacked an internet subscription and 26 percent did not have a desktop or laptop in 2019, this provision of digital resources and device accessibility is a necessity (*Desktop Ownership Among Adults in the US*, 2020;*U.S. Households with Internet Subscriptions*, 2021). Libraries also assist individuals with research, providing online chat services for answering questions, access to databases, and posting digital archives of historical documents (*About Us*, 2022). While individuals may not necessarily be going into the physical buildings as much, they are still accessing the resources provided by them. As such, the existence of these institutions is still necessary in today's world, if

not more so than before.

Due to their free provision of knowledge, libraries are some of the most democratic institutions in the US. By providing equal information access to all individuals "regardless of age, education, ethnicity, gender, language, [and] income," libraries help to strengthen communities and improve literacy (*Libraries are America's Most Democratic Institutions*, 2022). Illiteracy has been linked to higher rates of unemployment, lower paying jobs, decreased access to health services, and even higher rates of incarceration (*Child Illiteracy in America: Statistics, Facts, and Resources*, 2021). Libraries help to eliminate these issues by informing individuals through technology training, job preparation, and filling out government forms (*Libraries are America's Most Democratic Institutions*, 2022). By doing so, they start a conversation regarding important issues individuals face in their day-to-day lives. According to Eshelman (1987), the civic aim of the free public library is to offer the means to become more informed citizens, ultimately providing "a range of views on [important] issues so that people may make up their own minds about where they stand." As such, part of the importance of libraries is serving as a modern pillar of what political science deems the political socialization process.[1] Thus, their use should result in a better-informed public who regularly engages in political participation.

The rest of this paper will continue as follows. First, an examination of the literature and a determination of the research questions of focus will be presented in Sections II and III, respectively. Next, section IV will provide a breakdown of the model and methodology being used. This will start with an examination of the data and continue into explaining the regression frameworks being employed. Following this explanation, Section V will analyze the data using the model and provide relative results. Section VI will then discuss potential directions for future research. Finally, I will conclude in Section VII with a summary of the overarching results and contextualize their significance.

[1]Political socialization refers to how individuals "develop the attitudes, values, beliefs, opinions, and behaviors that are conducive to becoming good citizens in their country" (*Political Socialization*, 2016).

II Literature Review

The benefits of libraries have primarily been measured in two respects, the first pertaining to education quality. Rodriguez, Paul, and Valderrama (2015) examines this question through a difference-in-differences approach in Colombia to calculate the effect of newly created libraries on test scores. Using library distance from schools as the treatment/control decider, with the treatment group being widely defined as those schools "near" libraries and the control group as those "far" from libraries, the study ultimately finds a non-significant impact on test scores from the new libraries. While helpful, this does not necessarily tell the full story as to why libraries did not significantly impact education in Colombia. For example, it could be the case that individuals in the country don't have a thorough understanding of the proper utilization of these resources, thus leading to no impact. Bhatt (2010), however, potentially clarifies this through an empirical model looking into the impact of public library use on television use, homework, and reading in the United States. Employing data from several national surveys including the Current Population Survey and National Household Education Survey, Bhatt's analysis also includes the instrumental variable of household distance to nearest public library to account for potential endogeneity. Overall, the results show that those who used public libraries not only spent more time reading and doing homework, but also had decreased television use. While reading and doing homework more are not necessarily beneficial on their own, they do help to build skills that are necessary for future education and employment opportunities.

Additionally, another area in which the benefits of public libraries have been explored is innovation. Berkes and Nencka (2019) uses a difference-in-differences model to examine how the construction of Carnegie libraries affected patents.[2] In cities where Carnegie libraries had been built, the paper did find significantly higher amounts of patents, thus suggesting libraries have a positive impact on innovation. While the literature has examined both education and innovation, there are so many other possible benefits yet to

[2]Carnegie Libraries refer to the more than 1500 libraries built between 1886 and 1919 from the $1.2 billion worth of grants donated by Andrew Carnegie (Berkes & Nencka, 2019).

be explored. These include things such as financial success, civic engagement, and even political participation.

Political participation in the literature has mainly been examined through two lenses similar to libraries, education and organizational affiliation. Regarding education, Parinduri (2019) examines the impact that being a high school graduate has on being an informed voter in Indonesia. While the author does not find a significant impact, these results may not directly translate to the US due to the two countries having different voting laws and education systems. Milligan, Moretti, and Oreopoulos (2004), however, looks into education on voting in the US and UK, finding a positive, significant relationship in the US data. Ajilore and Alberda (2017) also examines education, but instead uses peer effects. Peer effects can be loosely defined as the impact that the surrounding students or "peers" have on some aspect of an individual. In this case, the authors find that those who sat near peers with high grades in social studies classes were more likely to participate in civic engagement.

Focusing on an individual's organizational affiliation, Gordon (2003) analyzes the effect of common interest developments in California on political participation, since these institutions act as "private governments."[3] While these institutions have grown in both size and popularity, there has been a simultaneous decrease in civic engagement. Despite this, the authors find no significant difference in measures of political participation including registration, voter turnout, and party registration. Additionally, Schulz and Bailer (2012) look at what attributes of an organization encourage political participation. Overall, they find that expressive forms of political participation—such as activism and protest—require institutions that reach out to politicians through phone, email, or in-person contact. Less expressive forms such as voting, however, were shown to merely require a place for the political socialization process to occur rather than a specific behavior.

Similar to Schulz & Bailer with regard to fostering political participation, Binder (2021) examines the validity of Tocqueville's Treatise of Democracy in America, which

[3]Common interest developments refer to institutions such as condominiums, cooperatives, planned developments, and similar community structures (Gordon, 2003).

believes volunteering and voluntary associations create democratic virtues, ultimately enhancing political participation. To do this, Binder utilizes the British Household Panel Survey from 1991-2008 and analyzes civic engagement's impact through a fixed effects methodology. While finding a positive effect on partisan support from volunteering and voluntary associations, the author notes that these are less economically and statistically significant than previously conducted studies, especially when association types are broken down into more specific groups. Nonetheless, Binder's study still provides evidence for the truth of the Tocqueville hypothesis and further justifies the impact that voluntary associations have on political participation.

This study is the first to directly examine the impact of libraries on political participation. While libraries and education are highly connected and the impact of education on political participation has been studied previously, one of the major factors that separates libraries from education is the element of choice. In most cases, children are being required to be educated from kindergarten through high school. Libraries, however, are almost completely voluntary institutions. Children and adults alike must make the choice to attend a library and use its services. This element of choice is quite similar to that of the political participation process. Unlike education, both libraries and political participation are optional; people can choose to participate as much or as little as they would like. Libraries also are not necessarily functioning in the role of a private government, but rather provide a space for communication and conversation. This dismisses non-significance concerns from Gordon (2003) and provides potential hope from Schulz and Bailer (2012). Thus, examining libraries can be easily differentiated from examining education in the research, while still being a legitimate institution to analyze due to their political socialization and voluntary association aspects.

III Research Questions

My analysis seeks to answer three key questions. My primary research question focuses on whether or not libraries have a significant impact on a region's political participation rates. Due to the conversation, connection, and community that results from

libraries, it is likely that this is a place where the political socialization process is occurring. Thus, determining whether or not these institutions are actually impacting political engagement is a question of high importance.

Second, I will ask—if libraries do have a significant impact—which indicator of political participation is impacted the most. With regard to political participation, it is important to remember that politics are a part of so many aspects of daily life. As such, the decision of whether or not to participate in the political process is constantly playing out in both measurable and immeasurable ways. Something as simple as sending a postcard in the mail to a relative or sharing a post on Facebook could potentially be an act of political participation. With this in mind, my research will examine political participation from multiple contexts, as discussed in the *Model & Methodology* section.

Third, I want to know whether or not the impact of a library affects an individual's political affiliation. In theory, libraries should be serving as places that create dialogue, encourage communication, and push individuals to pursue their ideas through research. Thus, this provision of conversation allowed by libraries should lead to a rise in overall political participation. Additionally, it is possible that one group may be more conducive to this communication, thereby causing more individuals to register with them as a result of libraries. As such, it is critical to understand if libraries have different impacts on political party affiliation in order to understand the role these institutions play in the political process. By answering these three questions, I plan to fill the current gap in the research that connects libraries and political participation, as well as gain a better understanding of how social change can be enacted in this seemingly indirect manner.

IV Model & Methodology

IV.A Data

This paper utilizes a panel dataset of California counties from 2000 to 2019. California, being composed of 58 counties and 482 municipalities, provides a large enough sample to thoroughly test my research questions. The data set employs information on

both central and branch libraries in the state of California to measure the impact on political participation. Data regarding libraries and their location was obtained from the California State Library Branch Directory. In order to obtain the years that branches opened, both county and individual library websites were examined for opening, closing, and renovation years. In cases where this information was not readily available, individual libraries were either contacted by phone and email, or local news articles were used. Once this information was gathered, it was aggregated at the county level and used to create the *libraryquantity* variable that tracks changes in the number of libraries by county per year. Similarly, *libqtypersqmile* was calculated to track change in the number of libraries per square mile by county per year to account for library density.

For political participation, the California Secretary of State website has both eligible and registered voters by county available, as well as the number or registered Democrats and Republicans. These numbers for registered Democrats and Republicans were then converted to percentages using the number of eligible voters as a divisor. Additionally, voter participation for presidential level elections was collected for each county using information from the MIT Election Data + Science Lab. Similar to the registration variables, voting in election variables were created as follows: *totalreg_pct*, the percent that voted out of those who had registered; *totalelig_pct*, the percent that voted out of those who were eligible; and *presdempct* & *presrepubpct*, the percent who voted as Democrats and Republicans, respectively.

Additionally, my analysis will rely on the inclusion of four key demographic variables to act as controls. To adjust for the relative number of individuals within a given county, I used population per square mile or *popbysqmile*. Population data per year was collected from the US Census Bureau, with information from the years 2000 and 2010 being actual census data and all other years being approximate population estimates. This information was then divided by the land size in square miles using data from the California State Association of Counties. To control for financial stability and wealth, my analysis uses the Local Area Unemployment Statistics from the State of California's Employment Development Department. This dataset had information regarding the number

of those who were in the labor force and unemployed, which I then utilized to create the unemployment variable, *unemp_rate*, that is measured as a percent. Finally, I include demographic information for race and gender. Thus, *poc_pct* and *fem_pct* were created using census data as measures of the percent of people of color and percent of women, respectively. Finally, since these data examine voter registration and registration occurs at its highest frequency around elections, the dummy variable *presyear* has been included. This variable takes on the value of 1 for presidential elections years (in the case of this data set, the years 2000, 2004, 2008, 2012, and 2016) and 0 for all other years.

After the data was collected and cleaned, the summary statistics in Table 1 were obtained.

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Library Quantity	1,160	18.52	34.04	1	248
Libraries Per Square Mile	1160	0.020	0.0777	0.0002	0.596
Percent of Registered Voters	1,160	0.741	0.0732	0.511	0.952
Percent of Registered Democrats	1,160	0.393	0.0822	0.190	0.582
Percent of Registered Republicans	1,160	0.360	0.0998	0.0654	0.524
Population Per Square Mile	1,160	665.7	2,317	1.419	18,756
Percent of POC Individuals	1,160	0.171	0.100	0.0394	0.507
Unemployment Rates	1,160	0.0840	0.0402	0.0208	0.295
Percent of Females	1,160	0.494	0.0230	0.351	0.521
Presidential Year	1,160	0.250	0.433	0	1

Table 1: Data from the California State Library Branch Directory, California Secretary of State, California State Association of Counties, California Employment Development Department, MIT Election Data + Science Lab, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. All variables except Library Quantity, Population Per Square Mile, and Presidential Year are percentages. Presidential Year is the only binary variable included.

As shown above, the observation count of 1,160 represents data for 58 counties between the years 2000 and 2019. In terms of variables of importance, the mean value of *libraryquantity*, suggests that the average county has 18.52 libraries. Additionally, there appears to be a slight majority of Democrats, with 39.3% of individuals in these counties being registered to the Democratic Party and 36% being registered to the Republican Party.

Additionally, a similar set of summary statistics for information on presidential elec-

tion voting is provided in Table 2. Similar to before, the observation count of 290 repre-

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Library Quantity	290	18.41	34.00	1	248
Libraries Per Square Mile	290	0.020	0.078	0.0002	0.596
Democratic Presidential Voters	290	0.491	0.139	0.211	0.850
Republican Presidential Voters	290	0.468	0.140	0.0929	0.726
Population Per Square Mile	290	658.8	2,298	1.419	18,543
Percent of POC Individuals	290	0.167	0.0989	0.0409	0.492
Unemployment Rate	290	0.0832	0.0372	0.0282	0.277
Percent of Females	290	0.494	0.0232	0.356	0.521
Percent of Registered Voters	290	0.755	0.0644	0.473	0.924
Percent of Eligible Voters	290	0.578	0.0866	0.329	0.792

Table 2: Data from the California State Library Branch Directory, California Secretary of State, California State Association of Counties, California Employment Development Department, MIT Election Data + Science Lab, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. All variables except Library Quantity and Population Per Square Mile are percentages.

sents data for 58 counties, but only during the years 2000, 2004, 2008, 2012, and 2016. Thus, while the mean for *libraryquantity* is not the same due to the removal of all other years, the two are still close, with these observations having an average of 18.41 libraries per county.

IV.B The Removal of Los Angeles County

While creating and examining the dataset, it became clear that there was a substantially large number of libraries in Los Angeles County relative to the other California counties being examined. Additionally, Los Angeles County has an extremely large population in comparison to most other counties in California, with the 2010 U.S. Census listing nearly ten million inhabitants. Thus, there was some uncertainty on whether Los Angeles County should be included in the analysis. In order to look into this issue further, I recompiled the original summary statistics without Los Angeles County included.

As shown in Table 3, removing Los Angeles County drastically decreases the average quantity of libraries within a county, with the number falling from 18.52 to just 14.53 libraries. Similarly, population per square mile also experiences a significant change,

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Library Quantity	1,140	14.53	15.98	1	83
Percent of Registered Voters	1,140	0.741	0.0733	0.511	0.952
Percent of Registered Democrats	1,140	0.391	0.0814	0.190	0.582
Percent of Registered Republicans	1,140	0.363	0.0991	0.0654	0.524
Population Per Square Mile	1,140	634.8	2,326	1.419	18,756
Percent of POC Individuals	1,140	0.169	0.100	0.0394	0.507
Unemployment Rates	1,140	0.0842	0.0404	0.0208	0.295
Percent of Females	1,140	0.494	0.0231	0.351	0.521
Presidential Year	1,140	0.250	0.433	0	1

Table 3: Data from the California State Library Branch Directory, California Secretary of State, California State Association of Counties, California Employment Development Department, MIT Election Data + Science Lab, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Los Angeles County has been removed, as described in Section IV.B. All variables except Library Quantity, Population Per Square Mile, and Presidential Year are percentages. Presidential Year is the only binary variable included.

dropping by over twenty persons when removed. For further confirmation, the summary statistics for presidential election years without Los Angeles County were also obtained and can be found in Table 4.

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Library Quantity	285	14.42	15.87	1	82
Democratic Presidential Voters	285	0.488	0.138	0.211	0.850
Republican Presidential Voters	285	0.472	0.139	0.0929	0.726
Population Per Square Mile	285	627.9	2,306	1.419	18,543
Percent of POC Individuals	285	0.165	0.0987	0.0409	0.492
Unemployment Rate	285	0.0834	0.0373	0.0282	0.277
Percent of Females	285	0.494	0.0234	0.356	0.521
Percent of Registered Voters	285	0.756	0.0642	0.473	0.924
Percent of Eligible Voters	285	0.578	0.0871	0.329	0.792

Table 4: Data from the California State Library Branch Directory, California Secretary of State, California State Association of Counties, California Employment Development Department, MIT Election Data + Science Lab, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Los Angeles County has been removed, as described in Section IV.B. All variables except Library Quantity and Population Per Square Mile are percentages.

Like the registration summary statistics, library quantity and population per square mile both decrease by similar amounts.

While these changes do indicate a potential need to remove Los Angeles County from the data set, there was still a possibility that this wouldn't impact the coefficients in the regression analysis. To test this claim, I performed a Chow test in order to determine whether the parameters between Los Angeles County and all other California counties were different. With an F-statistic of 45.11 and critical value of 3.00, the test determined that these parameters were not equal at the 5% level. As such, my analysis of the data and results does not include Los Angeles County.

IV.C Ordinary Least Squares Regression

To test for the impact of libraries on political participation, my analysis utilizes two main regression frameworks. First, I will use the following standard Ordinary Least Squares regression:

$$PoliticalParticipation_{it} = \beta_0 + \beta_1 libraryquantity_{it} + \beta_2 \mathbf{X}_{it} + \epsilon$$

In the above regression, "political participation" refers to one of seven measures at the percent level: registered voters, registered Democrats, registered Republicans, total eligible voters voting during a presidential election, total registered voters voting during a presidential election, total Democrats voting during a presidential election, and total Republicans voting during a presidential election. Other than the presidential election information that is only available every four years, all other registration and eligibility data is annual. The regression will be run seven times to determine whether any, if not all, of these measures are significantly impacted by the independent variable of focus, *libraryquantity*.

The *libraryquantity* variable represents the number of libraries in a county and changes when a new library is added to the county. In the case of a library opening in order to replace an existing library being taken out of use, this variable would remain

constant. This will allow me to determine what percentage point impact the addition of a new library has on the political participation of citizens. Finally, X is a vector of county characteristics including factors such as population per square mile, the unemployment rate, percent of women living in the county, and percent of people of color in the county. Additionally, a dummy variable for whether the form of political participation occurs during an election year (1 for yes, 0 for no) is also represented under this X vector.

IV.D Fixed Effects Regression

My second regression framework utilizes a fixed effects regression methodology. A fixed effects analysis controls for county and time characteristics that are constant, thereby helping to eliminate potential omitted variable bias. In my analysis, fixed effects controls for the average differences in any observable or unobservable predictors across counties each year. The regression is set up as follows:

$$PoliticalParticipation_{it} = \beta_0 + \beta_1 libraryquantity_{it} + \beta_2 \mathbf{X}_{it} + \theta_i + \theta_t + \epsilon$$

Similar to the previous regression, this fixed effect model will be run a total of seven times to account for each of the different measures of the dependent "political participation" variable. The *libraryquantity* independent variable of focus still denotes the number of libraries in a given county during a certain year. Additionally, the θ_i and θ_t represent the fixed effects being controlled for by county and year, respectively. Despite this control of variables that remain constant, X is still included as a vector of characteristics in order to control for time and location variant factors that are occurring alongside the library addition that cannot be netted out through the fixed effects. These include population per square mile, the unemployment rate, percent of women living in the county, and percent of people of color in the county. The same dummy variable for presidential election year used previously has also been included in this model under the X vector.

V Results

V.A Ordinary Least Squares model

Using traditional OLS regression analysis for registration data, the results in Table 5 were obtained in Stata. In the regression output on the percent of registered voters,

VARIABLES	(1) Registered Voters	(2) Registered Democrats	(3) Registered Republicans
Library Quantity	-0.000324**	-0.000465***	0.000690***
	(0.000140)	(0.000123)	(0.000156)
Population Per Square Mile	4.47e-06***	4.32e-06***	-7.79e-06***
	(8.51e-07)	(4.20e-07)	(5.18e-07)
Unemployment Rate	-0.635***	0.0389	0.212**
	(0.0478)	(0.0671)	(0.0832)
Percent of POC Individuals	-0.138***	0.365***	-0.464***
	(0.0263)	(0.0203)	(0.0241)
Percent of Females	-0.184*	0.840***	-0.820***
	(0.101)	(0.0717)	(0.0641)
Presidential Election Year	0.0275***	0.00938**	0.00219
	(0.00447)	(0.00463)	(0.00534)
Constant	0.904***	-0.0870**	0.822***
	(0.0507)	(0.0348)	(0.0307)
Observations	1,140	1,140	1,140
R-squared	0.174	0.302	0.368

Table 5: Data from the California State Library Branch Directory, California Secretary of State, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Marginal effects are reported. Coefficient estimates are taken from an ordinary least squares (OLS) regression per Section IV.C. Los Angeles County is not included in the analysis to prevent potential problems from the unusual extremities from the region, as stated in Section IV.B. Robust standard errors are reported in parentheses.
$* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$

registered Democrats, and registered Republicans, there are a few key points of importance. First, most of the variables are statistically significant at the 5% level, including the independent variable of interest, *libraryquantity*. Additionally, this suggests a causal effect between the number of libraries in a county and the registration variables, assuming the regression is correctly specified.

Second, alongside this relationship, there is an interesting difference between the

effect of the number of libraries on the percent of registered Democrats and Republicans. Instead of both trending in the same direction, an increase in libraries actually appears to have an opposing effect on the two affiliations. While an increase in the number of libraries correlates with an increase in the percent of registered Republicans, the library increase also occurs with a decrease in the percent of registered Democrats. This would suggest that the atmosphere and dialogue created by libraries may be more conducive to that of Republican ideologies than Democratic ideologies.

Despite the statistical significance of these variables, a somewhat more important factor to consider is their economic significance. Focusing first on the *libraryquantity* variable for the percent of registered voters, each new library corresponds to a 0.044 percentage point decrease in the percent of registered voters in the average county. Although elections have become much closer in terms of voting in recent years—with some candidates even winning by less than one percent of the vote—this decrease is unlikely to make a difference. Simply put, it is too small to be economically significant when considering election voting. However, a one standard deviation change in the number of libraries would produce a 1.5 percentage point decrease in the percent of registered voters. This number is economically significant but would require 34 additional libraries to be built. Additionally, one library would decrease the percent of Democrats and increase the percent of Republicans by 0.118 and 0.192 percentage points, respectively. These are equivalent to a respective 4.012 decrease and 6.528 percentage point increase for Democrats and Republicans when the standard deviation of libraries is taken into consideration.

Similarly, the Stata results in Table 6 were collected in order to examine library quantity increases per square mile. While the values are still statistically significant, the coefficients for registers voters, Democrats, and Republicans have all changed sign. This is likely due to the fact that this measure of library density requires a large change in the number of libraries within a county. For example, a one unit increase per square mile in the number of libraries within the average California county would require a county to

VARIABLES	(1) Registered Voters	(2) Registered Democrats	(3) Registered Republicans
Libraries Per Square Mile	0.585**	1.594***	-2.285***
	(0.241)	(0.172)	(0.194)
Population Per Square Mile	-1.54e-05*	-5.01e-05***	7.02e-05***
	(8.31e-06)	(5.88e-06)	(6.67e-06)
Unemployment Rate	-0.634***	0.0348	0.215***
	(0.0479)	(0.0659)	(0.0814)
Percent of POC Individuals	-0.143***	0.392***	-0.503***
	(0.0249)	(0.0184)	(0.0213)
Percent of Females	-0.192*	0.867***	-0.856***
	(0.104)	(0.0698)	(0.0603)
Constant	0.911***	-0.107***	0.854***
	(0.0523)	(0.0343)	(0.0295)
Observations	1,140	1,140	1,140
R-squared	0.150	0.326	0.403

Table 6: Data from the California State Library Branch Directory, California Secretary of State, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Marginal effects are reported. Coefficient estimates are taken from an ordinary least squares (OLS) regression per Section IV.C. Los Angeles County is not included in the analysis to prevent potential problems from the unusual extremities from the region, as stated in Section IV.B. Robust standard errors are reported in parentheses.
$* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$

construct over 2600 libraries (a number twice the size of the current libraries in the state). With this expectation, the previous measure of library quantity appears to be a more realistic approach to examining library construction.

Additionally, my OLS results from the presidential election information are listed in Table 7. Unlike with the registration data, many of the variables in this model are not statistically significant. Focusing on the independent variable of interest, libraryquantity is only significant at the 5% level for the percent of eligible voters that vote in the presidential election. Regardless of this however, similar trends are seen between the two datasets. In both the registration and presidential voting data, the total percent and percent Democrat variables experience decreases when a new library is added. Republican percentages, however, increase in the model for voting data and experiences a similar lack of economic significance, except when examined from the perspective of a one standard deviation increase.

VARIABLES	(1) Percent Registered	(2) Percent Eligible	(3) Percent Democrat	(4) Percent Republican
Library Quantity	-0.000404	-0.000675**	-2.99e-05	0.000290
	(0.000245)	(0.000302)	(0.000374)	(0.000400)
Population Per Square Mile	2.82e-07	5.37e-06***	9.15e-06***	-9.01e-06***
	(1.69e-06)	(1.28e-06)	(1.44e-06)	(1.38e-06)
Unemployment Rate	-0.693***	-1.111***	-0.173	0.415
	(0.0971)	(0.112)	(0.242)	(0.263)
Percent of POC Individuals	-0.116***	-0.221***	0.632***	-0.619***
	(0.0432)	(0.0591)	(0.0630)	(0.0647)
Percent of Females	0.135	-0.0721	1.376***	-1.401***
	(0.146)	(0.211)	(0.187)	(0.183)
Constant	0.772***	0.750***	-0.286***	1.232***
	(0.0717)	(0.105)	(0.0917)	(0.0897)
Observations	285	285	285	285
R-squared	0.179	0.260	0.392	0.383

Table 7: Data from the MIT Election Data + Science Lab, California State Branch Library Directory, California Secretary of State, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Marginal effects are reported. Coefficient estimates are taken from an ordinary least squares (OLS) regression per Section IV.C. Los Angeles County is not included in the analysis to prevent potential problems from the unusual extremities from the region, as stated in Section IV.B. Robust standard errors are reported in parentheses.
$* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$

V.B Fixed Effects Model

To get a more accurate understanding of the data and provide a more thorough analysis, output results from the fixed effects model were calculated in Table 8. Similar to the traditional OLS model for registration data, most of the coefficients in the fixed effects model are statistically significant at the 5% level. However, this is one of the only similarities between the two analyses.

With the inclusion of fixed effects for year and county, the trend direction of library quantity on the registration percentages changes sign. As such, a new library coincides with an increase in both registered voters and registered Democrats, as well as a decrease in the percent of registered Republicans. The likely reason for this change in sign is due to the removal of unobserved heterogeneity that does not change over time or by county. Due to the nature of an OLS analysis, this was unable to be accounted for in the previous regression. As such, these fixed effects estimates should be viewed as being more accurate than their OLS counterparts.

VARIABLES	(1) Registered Voters	(2) Registered Democrats	(3) Registered Republicans
Library Quantity	0.00546***	0.00349***	-0.00188**
	(0.00154)	(0.000856)	(0.000830)
Population Per Square Mile	-9.49e-06	1.81e-05***	-6.78e-06
	(9.27e-06)	(5.15e-06)	(5.00e-06)
Unemployment Rates	-0.150***	0.0435**	0.104***
	(0.0393)	(0.0219)	(0.0212)
Percent of POC Individuals	0.168**	-0.425***	-1.375***
	(0.0792)	(0.0440)	(0.0427)
Percent of Females	1.252***	-0.852***	-0.659***
	(0.366)	(0.204)	(0.198)
Presidential Election Year	0.0308***	0.00530***	-0.00345**
	(0.00260)	(0.00145)	(0.00140)
Constant	0.0258	0.817***	0.944***
	(0.180)	(0.100)	(0.0973)
Observations	1,140	1,140	1,140
R-squared	0.152	0.125	0.627
Number of Counties	57	57	57
County FE	YES	YES	YES
Year FE	YES	YES	YES

Table 8: Data from the California State Library Branch Directory, California Secretary of State, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Marginal effects are reported. Coefficient estimates are taken from a fixed-effects regression per Section IV.D. Los Angeles County is not included in the analysis to prevent potential problems from the unusual extremities from the region, as stated in Section IV.B. Robust standard errors are reported in parentheses.
$* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$

Another important difference between the two regressions is that the magnitude of the coefficients on *libraryquantity* are larger by a factor of ten in the fixed effects regression. As such, these coefficients are much more economically significant than they were previously. To put this into perspective, we can consider the addition of one library in the average county. Based on the model, this one library would cause the total percent of registered voters to increase by 0.737 percentage points, the percent of registered Democrats to increase by 0.471 percentage points, and the percent of registered Republicans to decrease by 0.254 percentage points in the average county. At first glance, these may still seem like relatively small changes. However, elections—as previously stated—are

much closer in recent years than in the past, with some races being decided by less than a percentage point. When factored together, this increase in Democrats and decrease in Republicans equates to a 0.725 percentage point total change in voting registration, favoring the Democratic Party. Ultimately, this could be the difference between a win or loss in an election cycle. Additionally, these number become even more economically significant when considering a one standard deviation change rather than a one library change. As such, the fixed effects model suggests promising results for Democrats, and voter participation more generally, when a new library is built.

Results are similarly provided for examining libraries per square mile in Table 9. In

VARIABLES	(1) Registered Voters	(2) Registered Democrats	(3) Registered Republicans
Libraries Per Square Mile	0.250	1.518***	-0.113
	(1.051)	(0.552)	(0.533)
Population Per Square Mile	-6.93e-06	9.50e-06	-7.41e-06
	(1.26e-05)	(6.61e-06)	(6.38e-06)
Unemployment Rate	-0.156***	0.0408*	0.105***
	(0.0420)	(0.0221)	(0.0213)
Percent of POC Individuals	0.210***	-0.374***	-1.413***
	(0.0731)	(0.0384)	(0.0371)
Percent of Females	1.270***	-0.845***	-0.672***
	(0.391)	(0.206)	(0.199)
Constant	0.0907	0.831***	0.931***
	(0.193)	(0.101)	(0.0977)
Observations	1,140	1,140	1,140
R-squared	0.032	0.106	0.623
Number of Counties	57	57	57
County FE	YES	YES	YES
Year FE	YES	YES	YES

Table 9: Data from the California State Library Branch Directory, California Secretary of State, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Marginal effects are reported. Coefficient estimates are taken from a fixed-effects regression per Section IV.D. Los Angeles County is not included in the analysis to prevent potential problems from the unusual extremities from the region, as stated in Section IV.B. Robust standard errors are reported in parentheses.
$* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$

these results, the only coefficient of importance with statistical significance is registered

Democrats. Additionally, as was the case for the OLS results, this measure is also fairly unrealistic as a one-unit increase, on average, requires the construction of more than twice the number of current libraries in California. Thus, the main focus of a library's impact should be placed on the previous non-square mile results.

For comparison, the fixed effects results for the presidential voting data were also calculated in Table 10. As was the case with the traditional OLS model, most of the

VARIABLES	(1) Percent Registered	(2) Percent Eligible	(3) Percent Democrat	(4) Percent Republican
Library Quantity	0.00263	0.00405	0.00232	-0.00154
	(0.00314)	(0.00288)	(0.00273)	(0.00263)
Population Per Square Mile	1.80e-05	-1.60e-06	9.12e-06	-1.04e-05
	(2.06e-05)	(1.89e-05)	(1.79e-05)	(1.72e-05)
Unemployment Rate	0.0719	-0.172*	0.640***	-0.0875
	(0.107)	(0.0981)	(0.0931)	(0.0896)
Percent of POC Individuals	0.308*	0.292*	0.961***	-1.517***
	(0.169)	(0.155)	(0.148)	(0.142)
Percent of Females	-0.246	0.498	-0.406	-0.905
	(0.777)	(0.713)	(0.677)	(0.651)
Constant	0.771**	0.241	0.437	1.205***
	(0.381)	(0.350)	(0.332)	(0.319)
Observations	285	285	285	285
R-squared	0.059	0.059	0.460	0.509
Number of Counties	57	57	57	57
County FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Table 10: Data from the MIT Election Data + Science Lab, California State Library Branch Directory, California Secretary of State, and individual county/district library branch websites, as described in Section IV.A. for the years 2000 to 2019. Marginal effects are reported. Coefficient estimates are taken from a fixed-effects regression per Section IV.D. Los Angeles County is not included in the analysis to prevent potential problems from the unusual extremities from the region, as stated in Section IV.B. Robust standard errors are reported in parentheses.
$* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$

coefficients in the presidential election voting regressions are not statistically significant. However, similar to the registration fixed effects, their relative size does indicate economic significance with regard to election voting percentages. Additionally, the change in trend direction carries over to this model as well, with Democrats and total voters increasing, and Republican percentages decreasing.

VI Directions for Future Research

The fixed effects estimates previously provided suggest promising results for the impact of libraries on political participation. Nonetheless, there are still some potential limitations with my analysis that need to be considered. First, since libraries are built according to the available budgets of counties/cities, as well as funds from private donors, it may be possible that high-income areas are receiving more funding for new libraries compared to low-income areas. These high-income areas, because they are wealthier, may be better informed than low-income areas, ultimately making it easier for them to engage politically. Thus, libraries increasing political participation could potentially be the result of the area simply being wealthier. While this may be the case, the inclusion of variables accounting for income (i.e. the unemployment rate) and use of a fixed-effects analysis helps to remove a majority of the potential omitted variable bias. However, it may also be helpful to look for other potential controls to better account for this variation and produce a more predictive model. One control in particular that I would have liked to include in my analysis is specifically regarding migration from other states. State migration could be playing a significant role if more individuals of a certain political affiliation are choosing to move to states like California to increase their representation and relative power of their vote.

Second, due to the nature of the regression framework being employed, the OLS analysis is only able to show causation between an increase in the number of libraries and a change in political participatory behaviors when correctly specified. Even with the fixed-effects model, there are still potential variables not included in the regression that have an impact on political participation. One potential way to solve this issue is by finding an instrumental variable to use in the analysis for the number of libraries in a county. In this case, an instrumental variable would be some variable that is both associated with *libraryquantity* and only connected to the political participation variables through its relationship with *libraryquantity*.

One instrumental variable I considered for my analysis was book circulation. With book circulation in a county effectively being directly connected to how many libraries

there are, this variable does have some validity. Additionally, since book circulation solely focuses on quantity and does not differentiate between different types of books, there should be little connection to political participation, other than through the number of libraries.[4] However, as stated earlier, books are just a small piece of the services provided by libraries. As such, book circulation is still an imperfect possible instrumental variable. However, by finding a variable that better fits the described criteria, I would be able to tackle the potential endogeneity problem.

With regard to expanding the current model, adding another independent variable of interest that examines library renovations may also provide new insight. While libraries are not being built frequently, they are much more often being renovated. Additionally, having this renovation may encourage non-users to access library services, thus encouraging more individuals to engage in political socialization. To do this, it would be necessary to define what exactly counts as a renovation to a library. For example, tearing down and rebuilding an existing portion of a library would certainly be regarded as a renovation to almost all individuals. However, updating the lighting system may not fit the necessary criteria to some, despite it still involving a change to the building. Thus, determining renovation criteria is a necessary point to consider if employing this variable in future research.

Geographic information system mapping or GIS may also be useful to incorporate. Through GIS, one could potentially examine the impact that distance to a library has on civic behaviors. With libraries typically being located in central and/or walkable distances, understanding the role of library proximity in a person's decision to engage politically may be critical. Based on the way my data is structured with the use of library quantity, a spatial analysis did not make sense. However, if future research chooses to use a different variable for examining library impact on political participation, GIS may be extremely helpful in understanding the relative effects.

Examining the decision to construct a new library may also be an important addition

[4]In theory, while book circulation includes books that encourage political participation, it also includes others such as Dolly Parton's *I Am a Rainbow* and Ina Garten's *Modern Comfort Food* that should have no impact by their subject matter. Thus, the variable should be minimally connected to political participation, if at all.

to the research. While some counties have gone decades without an update, others have built multiple within the past five years. Of course, the ability to build a library is partially based on the financial capabilities of a community. Still, there are so many more potential reasons for why a community may or may not choose to build a library, including who actually makes the decision, grant availability, locations available, etc. As such, trying to understand why certain counties get a new library may help to address potential error, and should be considered in future research.

VII Conclusion

Despite previous attempts to measure the value and impact that libraries provide, there is currently a lack of research regarding the impact they have on political engagement. As something that is a part of each aspect of our lives—when watching the news, scrolling through social media, and even leaving the house—being able to understand some of the factors that affect political participation is crucial. Especially when considering the opportunities provided to low-income communities by libraries, understanding their impact could potentially provide a way to increase civic participation. By knowing how to increase political participation, we can gain valuable insight into how to effect change with regard to electing representatives who fit a certain agenda and/or establishing policy that meets the needs of a particular group. Thus, an analysis like this is crucial to making change that impacts individuals in their everyday lives.

Using the fixed effects methodology, the data revealed that libraries have both a statistically and economically significant impact on the percent of registered Democrats, Republicans, and voters overall. This provides an extremely interesting manner of examining libraries through that has not been examined by current research in the field. Rather than being merely used as places to rent a book from, it suggests that libraries may be able to be used to effect actual change. It gives us hope that libraries are meeting one of their fundamental goals: providing a space to communicate, research, and freely express one's ideas. Additionally, it suggests that continuing to fund, build, and improve these facilities is a critical part of our society.

Ultimately, this paper serves as a jumping off point for future research regarding libraries and political participation. As suggested from the results, there is a clear and significant connection between libraries and political participation. Understanding and testing this relationship more is crucial for counties in their decision to build a library. Considering the resources required to build a library, and the typically limited overflow of both county and city budgets, knowing the full extent of their benefits is a key aspect of this decision. Further, continuing this research is necessary to allow for a society that more accurately represents the views of those living in it. Based on the amount of uninformed low-income individuals who choose not to participate and the resources made available to them through libraries, libraries provide a potential avenue to not only make the public better informed and civically engaged, but also prepared to create change at local, national, and global levels.

References

About us. (2022). The Digital Public Library of America. Retrieved from https://dp.la/about

Ajilore, O., & Alberda, G. (2017, Winter). Peer Effects and Political Participation: What is the Role of Coursework Clusters? *The Review of Regional Studies*, *47*(1), 47-62. Retrieved from https://ideas.repec.org/a/rre/publsh/v47y2017i1p47-62.html

Berkes, E., & Nencka, P. (2019). *'Novel' Ideas: The Effects of Carnegie Libraries on Innovative Activities* (Tech. Rep.).

Bhatt, R. (2010). The impact of public library use on reading, television, and academic outcomes. *Journal of Urban Economics*, *68*(2), 148-166. Retrieved from https://EconPapers.repec.org/RePEc:eee:juecon:v:68:y:2010:i:2:p:148-166

Binder, M. (2021, January). Enhancing Democracy: Can Civic Engagement Foster Political Participation? *Social Science Quarterly*, *102*(1), 47-68. Retrieved from https://ideas.repec.org/a/bla/socsci/v102y2021i1p47-68.html DOI: 10.1111/ssqu.12882

Child illiteracy in america: Statistics, facts, and resources. (2021, Sep).

Desktop ownership among adults in the us. (2020, Apr). Statista. Retrieved from https://www.statista.com/statistics/756054/united-states-adults-desktop-laptop-ownership/

Eshelman, W. (1987). Serving the Public Good: Coalitions for Free Library Services. *Coalitions and the Public Good*, 103-104.

Gordon, T. M. (2003). Crowd out or crowd in?: The effects of common interest developments on political participation in california. *The Annals of Regional Science*, *37*, 203-233.

Hartley, R. P. (2020). Unleashing the power of poor and low-income americans. *The Poor People's Campaign Report*, 1-20.

Healthier, more equitable america. (2021, Oct). Institute of Museum and Library Services. Retrieved from https://www.imls.gov/news/new-research-underscores-role

-museums-libraries-play-create-healthier-more-equitable-america

Libraries are america's most democratic institutions. (2022, Mar). Libraries Transform. Retrieved from https://ilovelibraries.org/librariestransform/libraries-are-americas-most-democratic-institutions/

Libraries foster community engagement. (2022). American Library Association. Retrieved from https://connect.ala.org/communities/community-home?CommunityKey=1976b40b-edf8-4252-a976-d3a2655d3df3

Library services in the digital age. (2020, May). Pew Research Center. Retrieved from https://www.pewresearch.org/internet/2013/01/22/library-services/

Milligan, K., Moretti, E., & Oreopoulos, P. (2004). Does education improve citizenship? evidence from the united states and the united kingdom. *Journal of Public Economics*, *88*(9-10), 1667-1695. Retrieved from https://EconPapers.repec.org/RePEc:eee:pubeco:v:88:y:2004:i:9-10:p:1667-1695

National poverty in america awareness month. (2022).

Parinduri, R. A. (2019). Does education increase political participation? evidence from indonesia. *Education Economics*, *27*(6), 645-657. DOI: 10.1080/09645292.2019.1668914

Political socialization. (2016, Nov). University of Minnesota. Retrieved from https://open.lib.umn.edu/americangovernment/chapter/6-2-political-socialization/

Rodriguez, L., Paul, J. D. T., & Valderrama, D. (2015). Are public libraries improving quality of education?: When the provision of public goods is not enough. *World Bank Policy Research Working Paper*(7429), 1-27.

Schulz, T., & Bailer, S. (2012). The impact of organisational attributes on political participation: results of a multi-level survey from switzerland. *Swiss Political Science Review*, *18*(1), 1-27. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1662-6370.2012.02052.x DOI: https://doi.org/10.1111/j.1662-6370.2012.02052.x

U.s. households with internet subscriptions. (2021, Jan). Statista. Retrieved

from https://www.statista.com/statistics/189349/us-households-home-internet-connection-subscription/

White, B. (2012). Guaranteeing access to knowledge: The role of libraries. *World Intellectual Property Organization Magazine*, *4*.