

THE IMPACT OF LIBRARY CLOSURES ON STUDENT ACHIEVEMENT

by

Zak Mark Murphy Vinter

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Economics

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2024

©COPYRIGHT

by

Zak Mark Murphy Vinter

2024

All Rights Reserved

DEDICATION

Firstly, I would like to thank my girlfriend Ylva for supporting me throughout the thesis process. I could not have done it without your support and understanding. Additionally, thank you to my family for continually pushing me to fulfill my potential and for getting me to where I am today.

ACKNOWLEDGEMENTS

I would like to thank Dr. Gregory Gilpin, my thesis chair, for his unwavering support and invaluable advice throughout this thesis process. I could not have asked for better guidance during my research journey. Thank you as well to Dr. Wendy Stock and Dr. Christiana Stoddard for their feedback and input. You have both helped me become a better writer and researcher during my time in the program. Thank you to Samantha, Nick, Nate, Tavio, and Holly. Having the support of the cohort has enabled me to stay positive during tough periods, which I am greatly appreciative of.

TABLE OF CONTENTS

1. INTRODUCTION	1
2. BACKGROUND	3
Library Expansion.....	3
Library Impacts.....	3
School Achievement	4
Library Funding on Student Achievement.....	4
3. THEORETICAL FRAMEWORK	6
4. DATA	8
Libraries	8
Student Achievement	8
Supplementary Data Sources	9
Population	10
Metro Status.....	10
Education Spending	10
Public School Libraries.....	11
5. METHODOLOGY	12
Model	12
Identification	13
6. SUMMARY STATISTICS.....	15
Closures.....	15
Comparison of Means	16
Library Outcomes	17
7. RESULTS.....	22
Impact on Student Achievement	22
All Students.....	22
Metro vs Non-Metro	23
Elementary vs Middle School.....	25
Event Study	25
Impact on Libraries	26
8. IDENTIFICATION CONCERNS.....	38

TABLE OF CONTENTS CONTINUED

Student and Community Characteristics.....	38
Random Assignment	39
Public School Libraries.....	39
9. CONCLUSION.....	44
REFERENCES CITED.....	46
APPENDICES	49
SUPPLEMENTARY TABLES	50
ALTERNATIVE SPECIFICATION	54
Data	55
Summary Statistics.....	56
Model	57
Results.....	58

LIST OF TABLES

Table	Page
1. Table 1. Number of Library Outlets, by Structural Change.....	17
2. Table 2. Summary Statistics: Comparison of Means across Treatment Status	20
3. Table 3. Summary Statistics: County-by-Year Panel of Library Outcomes	21
4. Table 4. Impact of a Library Outlet Closure on Student Achievement.....	28
5. Table 5. Impact of a Library Outlet Closure on Student Achievement, by Race.....	29
6. Table 6. Impact of Library Outlet Closure on Student Achievement, by Gender and ECD	30
7. Table 7. Impact of a Library Outlet Closure on Student Achievement: Comparing Effects in Metro and Non-Metro Counties by Race	31
8. Table 8. Impact of a Library Outlet Closure on Student Achievement: Comparing Effects in Metro and Non-Metro Counties by Gender and ECD.....	32
9. Table 9. Impact of Library Outlet Closure on Student Achievement, by School Type.....	33
10. Table 10. Impact of a Library Outlet Closure on Public's Use of Library Services	36
11. Table 11. Impact of a Library Outlet Closure on the Supply of Library Services	37
12. Table A1. Impact of a Library Outlet Closure on Student Achievement, TWFE Model	51
13. Table A2. Impact of a Library Closure on Student Achievement, by Race (Comparable Model).....	52
14. Table A3. Impact of a Library Outlet Closure on Student Achievement: Comparing Effects in Metro and Non-Metro Counties by Race (Comparable Model).....	53

LIST OF TABLES CONTINUED

Table	Page
15. Table A4. Summary Statistics: Change in Public Distance to Library Outlet Following Closure, by County	60
16. Table A5. Impact of Distance to Nearest Library Outlet on Student Achievement	61
17. Table A6. Impact of Public Distance to Nearest Library Outlet on Student Achievement, by Race	62
18. Table A7. Impact of Public Distance to Nearest Library Outlet on Student Achievement, by Gender and ECD.....	63

LIST OF FIGURES

Figure	Page
1. Figure 1. Number of Library Outlet Closures by Year	18
2. Figure 2. Percentage of Library Closures in Rural Areas	19
3. Figure 3. Public Library Closures by State.....	20
4. Figure 4. Impact of a Library Outlet Closure on Mathematics Test Scores.....	34
5. Figure 5. Impact of a Library Outlet Closure on Reading Test Scores	35
6. Figure 6. Correlation Between a Library Outlet Closure and Median HH Income.....	40
7. Figure 7. Correlation Between a Library Outlet Closure and Unemployment Rate	40
8. Figure 8. Correlation Between a Library Outlet Closure and Percentage of Black Students	41
9. Figure 9. Correlation Between a Library Outlet Closure and Percentage of Hispanic Students.....	41
10. Figure 10. Correlation Between a Library Outlet Closure and School Expenditures	42
11. Figure 11. Correlation Between a Library Outlet Closure and Poverty Rate	42
12. Figure 12. Correlation Between a Library Outlet Closure and Student Population	43
13. Figure A1. Histogram of Average Weighted Distance in Treatment Counties	60

ABSTRACT

I examine the relationship between public library closures and student achievement. Previous studies have focused on improvements to library services; however, the reverse relationship has been overlooked. Using data from the Public Library Survey and Stanford Education Data Archive, I calculate the impact of closures on reading and mathematics test scores for grades 3-8. I estimate that average reading test scores decrease by 0.021 standard deviations while math test scores decrease by 0.034. Impacts vary by gender with males experiencing on average a 0.01 standard deviation larger impact than females across both reading and math. Elementary school students experience larger impacts to math scores than middle school students, however no differences in reading. Additionally, racial minorities only experience an effect in math scores. These results show that even with policy makers best efforts to limit the impacts of closures through library outlet selection and alternative services, closures are still having a profound impact on students. In addition, library spending is not falling after a closure, indicating that the costs of closures far out way any possible benefits. Further research should explore additional community areas impacted by closures to provide policymakers with a comprehensive understanding of the costs associated with library outlet closures.

INTRODUCTION

In light of the ongoing advancements in modern technology and the rise of online learning, it is imperative to critically examine the role and funding of public libraries. Despite their continued governmental support, there exists a noticeable gap in comprehensive research on the impact of public libraries in communities. In 2019 alone, local and state government sources allocated approximately 13 billion dollars (about \$40 per person) to support the operation of more than 17,000 public library outlets in the United States (IMLS, 2019a).

A major reason to support the continued funding of libraries, is the impact these services have on children. In 2019, over 750 million library items were checked out by children, representing 35 percent of all materials in circulation (IMLS, 2019b). Additionally, there were over 3 million child programs conducted by libraries with a total attendance of 81 million.

This paper addresses the question of whether the closure of library outlets affects student achievement. Negative effects of such closures could have enduring implications for individual lifetime earnings, underscoring the urgency of this research for policymakers. Furthermore, this study fills a notable research gap, as little is currently known despite the closure of over 600 library outlets in the United States over the past decade. While prior research by Gilpin et al. (2024) suggests a positive relationship between library capital spending shocks and reading scores, the possibility of asymmetric impacts between positive and negative shocks to library service availability remains largely unexplored. Given students' reliance on library services, particularly in the absence of immediate substitutes, the impact of closures may be considerable.

To investigate the relationship between library closures and student achievement, I utilize both library outlet data from the Public Libraries Survey and annual county-level reading and

math grade-level assessment data from the Stanford Education Data Archive (SEDA) for the 2009-2018 period. The results indicate that closing a library outlet negatively impacts student achievement. It is estimated that average math test scores decrease by 0.034 standard deviations and average reading scores decrease by 0.021. Impacts vary by gender with males experiencing on average a 0.01 standard deviation larger impact than females across both reading and math. Elementary school students experience larger impacts to math scores than middle school students, however no differences in reading. Additionally, racial minorities only experience an effect in math scores and no differences were found across metro status. These results suggest a complex interplay between gender, age, and race.

Local policymakers often prioritize closing the least impactful facilities, assuring voters of minimal service loss or potentially enhanced services at fewer locations. However, the findings of this research challenge such assumptions. These results suggest that even when precautions are taken, the impacts of closures are still felt, defying the notion of minimal disruption touted by policymakers. In addition, library spending is not falling after a closure, indicating that the costs of closures far out way any possible benefits.

These findings underscore the complexity of evaluating the consequences of library closures on student achievement. By illuminating the multifaceted nature of factors such as electronic resources and transportation accessibility, the study emphasizes the need for a comprehensive approach to policymaking. Understanding the nuanced dynamics at play is essential for informing effective decision-making processes aimed at mitigating the adverse effects of closures on society as a whole.

BACKGROUND

Library Expansion

Historically, libraries have undergone continuous expansion and enhancement, yet there is a lack of reevaluation regarding their contemporary relevance. Andrew Carnegie's financing of 1,618 public libraries in the United States between 1890 and 1921 marked a significant phase of expansion (Karger, 2021). Research by Karger indicates that grants for library construction during this period led to a notable increase of 0.10 years in child education attainment and a 4% rise in non-wage income.

Analyses of relevant research from this historical period demonstrate profound transformations resulting from the introduction of libraries in the United States. The expansion of libraries correlates with urbanization and institutional innovation (Kevane and Sundstrom, 2014). Furthermore, studies by Berkes and Nencka (2021) reveal that library funding positively influenced patenting rates in treated municipalities, while Kevane and Sundstrom (2016) argue that public library investments did not significantly impact political participation.

Library Impacts

Recent research has focused on the effects of library usage. Bhatt (2010) shows that proximity to a library increases children's reading time while reducing television consumption and encouraging parental reading involvement. Gilpin and Bekkerman (2020) demonstrate that during school closures, distance from a library negatively impacts its use, especially for children attending public schools. These findings lay the groundwork for exploring the impact of library closures on academic achievement.

Neto et al. (2023) reveal that library openings lead to significant reductions in crime rates, particularly in nearby areas. However, Neto (2019) finds no substantial effect of library programming on local labor market conditions. These studies highlight the multifaceted role of libraries in communities, indicating both direct and indirect influences worthy of further investigation.

School Achievement

A substantial body of literature helps to establish the impact of capital investment on school achievement. This research is relevant due to the nature in which both library and school capital investment affect a child's opportunity for learning. Nielson and Zimmerman (2014) find that after 6 years of a new school being constructed, reading test scores increase while math test scores stay the same. Jackson and Mackevicius (2021) support the finding of a positive relationship between student achievement and school spending through a meta-analysis of 31 causal studies. There is, however, mixed results with Cellini et al. (2010), Brunner et al. (2022), and Martorell et al. (2016) who find little to no positive impact from increases to capital investments.

Library Funding on Student Achievement

Gilpin et al. (2024) is the most relevant research to the work conducted in this paper combining both stages. The research estimates the relationship between capital spending shocks on both library usage and student achievement. Using a dynamic difference-in-difference approach, they find that total visits, children's attendance at library events, and library checkouts increase after capital investment. With regards to test scores, improvement is found in reading

scores, however, no tangible change in math scores is identified. This research builds on this finding by isolating a more abrupt change to library accessibility due to closures, arguably more concerning than an expansion in library services. Due to the abruptness, closures will provide substantial variation in the accessibility of library services to children. In addition, the importance of quantifying the impacts of a decrease to library accessibility cannot be understated.

THEORETICAL FRAMEWORK

Analyzing the expected public response to library closures is crucial for formulating compelling research questions and providing explanations for any observed results. The primary reason for anticipating a significant impact of library closures on student achievement and learning is accessibility. When a library outlet closes, students and families are compelled to travel greater distances to access the in-person benefits of a library. Research has shown that increased distance from a library reduces students' library usage (Gilpin & Bekkerman, 2020; Bhatt, 2010).

In rural towns, particularly those lacking reliable public transportation, closures can leave families with no practical means to access their local public library. Consequently, I anticipate that the effects of closures will be significantly more pronounced in non-metro counties. Moreover, non-metro counties typically have fewer libraries within the surrounding area. For instance, losing one of only two library outlets, compared to one of five, will have more substantial repercussions on students' accessibility.

Another reason for lasting, prominent effects is the limited availability of substitutes. While school libraries provide an alternative, their accessibility is restricted during weekends and particularly holidays. Additionally, a substantial number of public schools, 8,830, lack a school library as a viable substitute, and nearly 17,000 do not have a librarian on staff (ALA, 2020). Having access to educational support services has been found to be pivotal during the summer break. In a meta-analysis of 37 studies, Lynch et al. (2021) finds an average weighted increase of 0.09 standard deviations in mathematics test scores due to summer learning programs.

Online services can serve as substitutes for in-person libraries; however, many individuals still rely on libraries for internet access. This reliance is particularly pronounced in low-income neighborhoods, where 44 percent of individuals below the poverty line use libraries for computer and internet access. Additionally, half of all 14 to 18-year-olds report having used a library computer within the past year (Shepard & Bitter, 2010).

DATA

Libraries

Data on libraries is sourced from the Public Library Survey (PLS), an annual universe survey conducted by the Institute for Museum and Library Services (IMLS). In the most recent fiscal year (2021), 96.7 percent of the total 9,207 eligible public library administrative entities responded to the survey (IMLS, 2021). This provides variables pertaining to the demand and supply of library services.

Furthermore, data is provided at the library branch level, with over 17,000 reported. The key variable of interest utilized from this dataset is the structural change code. This variable records structural changes at the library outlet level in a given fiscal year, including actions such as adding, deleting, relocating, or merging library outlets.¹ When constructing the final dataset for analysis, this variable was used to signify treatment and control counties. In this analysis, I focus specifically on counties that have experienced a closure within the study period, but not openings. There are 222 treatment counties and 2,250 control counties that experienced no openings, closures, or relocations within the time period of the analysis.

Student Achievement

To measure the impact of library closures on student achievement, I use data from version 5.0 of the Stanford Education Data Archive (SEDA). This is created by the Educational Opportunity Project at Stanford University (Fahle et al. 2024). Aggregate test score data from

¹ This variable was established in the Public Library Survey in the 2005 fiscal year. However, closures were not recorded until the fiscal year of 2008.

each state is collected by the EdFacts data system. States are required to administrate annual standardized tests for public school students following the passage of the federal No Child Left Behind Act in 2002.² Each state however has control over the construction of these examinations and the thresholds set for performance levels. SEDA uses heteroskedastic ordered probit (HETOP) models to estimate cutscores by state, grade, subject, and year.³ These are then converted to the NAEP scale and standardized by cohort.⁴ Final estimates are reported in standard deviations for each grade, subject, county, and year.⁵

SEDA provides covariate data derived from two primary sources: the Common Core of Data (CCD) and the American Community Survey (ACS), the latter being compiled by the National Historical Geographic Information System (NHGIS). These covariates are used to address possible bias in the results, caused by time-varying changes that are correlated with both the closures and student achievement.

Supplementary Data Sources

I have some minor additional data sources surrounding population estimates, county metro status, library funding referenda, and education spending.

² They must implement both a Mathematics and Reading Language Arts exam for grades 3-8.

³ Cutscores are thresholds set out by states for specific educational proficiency levels.

⁴ The Nation's Report Card (NAEP) conducts a nation-wide standardized test with a sample of students giving SEDA the opportunity to compare cutscores.

⁵ More information on the creation of this dataset can be found in the technical documentation provided by Fahle et al. 2024.

Population

I collect population estimates from the National Cancer Institute's Surveillance, Epidemiology, and End Result Program. They provide single-year-of-age county-level bridged population estimates. I use this to control for population when conducting analysis on the impacts of library closures on library supply and demand.⁶

Metro Status

Data on metro status is collected from the Economic Research Service (ERS) which is part of the U.S. Department of Agriculture. The Urban Influence Codes that the department provides was established in 2013, splitting counties into 12 groups based on the population size of its metro area. The two groups with the highest populated metro areas are defined as metro which includes 1,167 counties.⁷ Metro classifications follow the delineation set by the Office of Management and Budget.

Education Spending

I gathered data on total education spending at the county level to understand whether there is a relationship between library closures and education spending. This was done because if the two are correlated, then estimates found are likely over-estimated. The data comes from the Common Core of Data (CCD) and is published in the Local Education Agency Finance Survey

⁶ A benefit to this data is that it enables controlling for children's populations when estimating the impact on circulation of materials and attendance at programs.

⁷ The two group classifications are "those in "large" areas have at least 1 million residents and those in "small" areas have fewer than 1 million residents" (ERS, 2013).

(F-33). Data is provided for all public schools in the US, for students from pre-kindergarten to grade 12.⁸

Public School Libraries

Lastly, data was gathered on the number of library staff within a county. This was done to determine whether after a public library closure, schools were compensating by investing in their own library services. The data used is the Local Education Agency Universe Survey from the Common Core of Data (CCD). Both the number of library specialists and support staff were measured.

⁸ When reported in summary statistics table 2, nominal education spending is reported. Data is in 2018 dollars.

METHODOLOGY

Model

To assess the impact of library closures on student test scores, a binary public library closure variable is used to signify treatment. A two-stage difference-in-difference method developed by Gardner (2022) was implemented, designed to address the inherent flaws of the basic difference-in-difference model when the treatment does not happen at the same time for all treated observations.⁹

The primary shortcoming is that the basic difference-in-difference model wrongfully assumes that outcomes are linear in group, period, and treatment status. The two-stage approach uses the assumption of untreated outcomes being linear in county and year under parallel trends to create an approach robust to treatment-effect heterogeneity.¹⁰

In the first stage, county, grade, and year fixed effects, as well as covariates are calculated using a sample of untreated and not-yet-treated counties under the assumption of parallel trends:

$$Y_{gct} = \delta_c + \alpha_g + \gamma_t + X'_{ct} \theta + v_{gct}$$

where Y_{gct} represents the average difference in standard deviations between test scores in a particular counties grade and the national reference cohort for that specific grade. Mathematics and Reading Language Arts are the two subjects in which tests have been taken. In this model δ_c represents county-level fixed effects, α_g grade-level, and γ_t year-level. X is a vector of

⁹ Additionally, results were estimated with a traditional two-way fixed effects model in table A1. Results were consistent with Gardner estimates.

¹⁰ Refer to Gardner (2021) for in-depth proof of the robustness of the estimator.

covariates, including unemployment rate, poverty rate, median income, school expenditures, percent of students receiving free or reduced lunches, and percent of Black and Hispanic student by county and year.

The second stage estimates:

$$Y_{gct} - \hat{Y}_{gct} = \beta(\text{Closure})_{ct} + u_{gct}$$

where $Y_{gct} - \hat{Y}_{gct}$ are the residual outcomes generated in the first stage. The established fixed effects and covariates calculated on the untreated and not-yet-treated are subtracted, and the adjusted outcomes are regressed on treatment status. Therefore, the coefficient of interest β , represents the average effect of a library outlet closure on student achievement for the treated. Standard errors are clustered at the county-level.¹¹

Identification

The key assumption relied upon in this model is parallel trends. For identification, in the absence of treatment, the trend of the treated groups must follow the trend of the untreated group. Therefore, in the absence of a library outlet closure, test scores in treated groups must follow untreated groups in the post treatment time period.

There are concerns that this assumption does not hold with respect to library closures. The major identification issue that has to be addressed is endogeneity with regards to when and

¹¹ I chose to analyze county-level outcomes instead of school district-level outcomes for two main reasons. First, it ensured consistency across library outcomes and student achievement measures, facilitating easier comparison. Second, it avoided the need to make subjective judgments about which school districts were treated, which could introduce bias and ambiguity into the analysis.

why a public library closure occurs. Economic intuition suggests that if a public library closes, it is likely to be part of a county-level funding issue that could be driving both increases in closures and declines in student achievement. Therefore, effects measured in this analysis could be driven by external factors, overestimating the impact of a closure, and validating the parallel trends assumption.

To address this concern, I take multiple steps to address it. Firstly, I compare the means of treatment and control counties for control variables to understand if there are community characteristics in which the two groups differ. In addition to the time varying controls included in the preferred model, state-time trends are added. Results are also calculated in an event-study framework to illustrate any possible differences in trend in the pre-period.¹² I also check that closures are having an impact on the demand and supply of library services. Lastly, as robustness checks, I establish any correlation between control variables and library closures through the event-study framework.

I also, explore an alternative approach using distance to library outlet as a continuous treatment variable. Using distance as a dependent variable instead of a binary closure variable provides an additional dimension of variation that is unlikely correlated with events that caused a library outlet to close. This reduces the risk of omitted variable bias, which could lead to spurious correlations between closures and academic outcomes.

¹² This is not the main specification, due to the small number of years being studied.

SUMMARY STATISTICS

Closures

To provide insight into how the number of library outlets are changing through time, I created summary statistics separating structural changes into opening, closures, and relocations.¹³ Table 1 provides an overview of outlet openings, closings, and relocations while Figure 1 graphs these closures through time. Notably, the number of closures exhibited fluctuations, ranging from 43 to 92 per year. A peak in closures occurred between 2010 and 2012, with subsequent years showing consistently lower closure numbers, with the exception of 2017.¹⁴ The overall trend of the number of openings and closures in the first years of the study likely point to the impacts of the 2008 financial crisis.

In Figure 2, I delve into the categorization of closures into metro and non-metro library outlets. As of 2012, 37% of library outlet closure's fall under the metro category, while 63% are classified as non-metro. Although this ratio fluctuates annually, it is noteworthy that in 9 out of 14 years, metro closures outnumber non-metro closures. It is important however that there is a representative number of closures in each category given the geographical distinction between areas. I also find that across the sample period, the percentage of metro and non-metro libraries

¹³ To differentiate between genuine openings and closures versus relocations, I utilize the three-digit FIPS county code. If both an opening and a closure happen within the same county, it's considered a relocation, and one opening and closing are deducted from the totals. For relocation classification, both events must occur within the same or adjacent years. If they occur in adjacent years, the relocation is recorded in the year when the relocation process concludes.

¹⁴ An increase in 2017 likely points to the change in presidential leadership and corresponding changes in funding strategies.

that close are similar. 11.87% of metro library outlets close between 2009-2018 while 13.45% close in non-metro counties.

Figure 3 illustrates the geographic distribution of closures across states. As anticipated, heavily populated states exhibit a higher frequency of closures, with Florida leading with 52 closures. Moreover, states within the Great Lakes region experienced substantial closures during the study period. Conversely, certain populous states experienced fewer closures, including Oklahoma with 3 closures, Wisconsin and Colorado with 6 each, and South Carolina with 9.

Comparison of Means

In table 2, a comparative analysis of both student and community characteristics between treatment and control groups offers insights into their comparability. Treated counties demonstrate a slightly higher proportion of Black students compared to non-treated counterparts, while no significant difference is observed in the percentage of Hispanic students. Notably, treated counties exhibit elevated poverty and unemployment rates, with a 0.55% disparity in unemployment and 0.23% in poverty compared to non-treated counties. Conversely, non-treated counties show a marginally higher percentage of economically disadvantaged students by 0.16%. However, there are no discernible discrepancies in terms of income across the groups. Overall, these findings suggest that treated counties tend to have a higher prevalence of economic disadvantage, although these differences are not of significant concern.

Moreover, school expenditures per student are notably higher in treated counties, reflecting a potentially greater investment in education. Additionally, the considerable disparity in grade-level student enrollment, with twice as many students in treated counties, is understandable, given their likelihood of hosting more library outlets. To address this

discrepancy, all models incorporate student population weights to ensure proper adjustment for this variation.

Library Outcomes

Table 3 provides county-by-year summary statistics on all library supply and demand outcomes. On average, counties receive upwards of 250,000 visits to their library outlets per year, with 400,000 items checked out.¹⁵ When considering children's outcomes, they account for over 25% of the total items in circulation. Additionally, on average, there are 220 available programs they can attend, with 14,000 attendees per year.

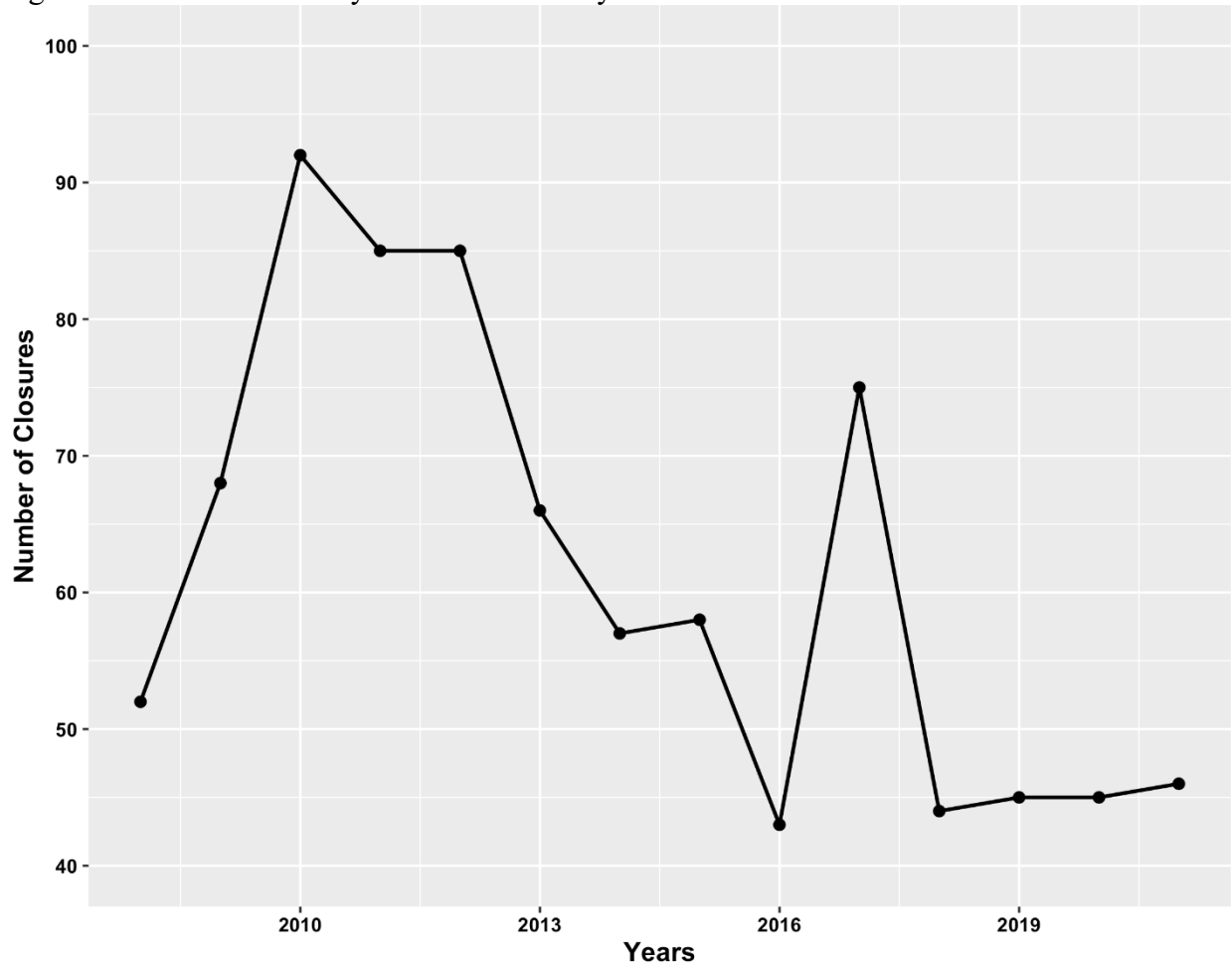
Table 1. Number of Library Outlets, by Structural Change

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total
Opening	109	80	84	49	75	55	63	39	67	59	44	724
Closing	52	68	92	85	85	66	57	58	43	75	44	725
Relocation	20	31	50	47	51	28	22	20	16	19	23	327
Total Branches	17510	17522	17514	17478	17468	17457	17463	17444	17468	17452	17452	

Source: IMLS PLS. Note: Structural change codes (STATSTRU) include 2, 3, 8, 23, and 24.

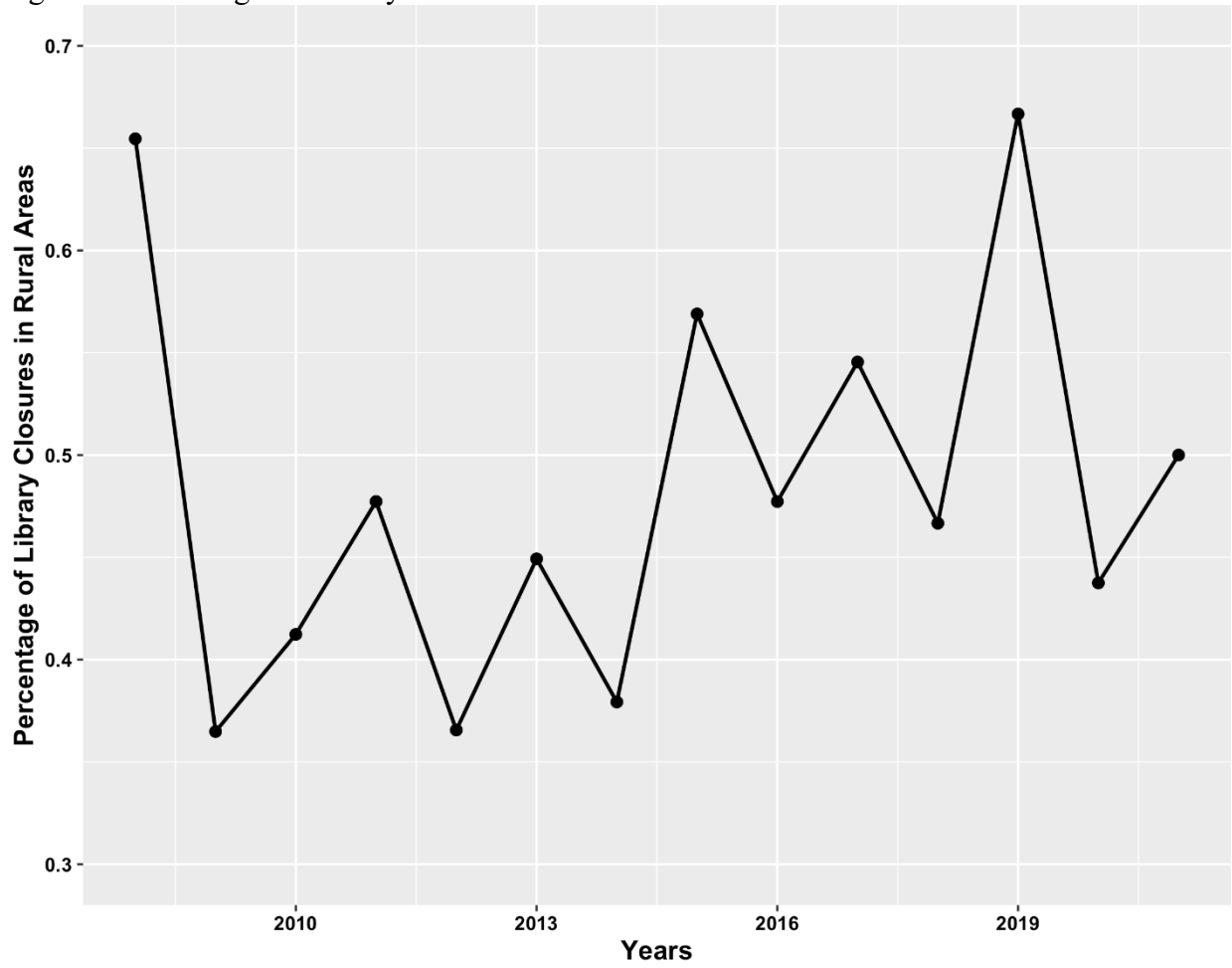
¹⁵ The minimum for all library outcomes is equal to zero due to some counties not having library outlets.

Figure 1. Number of Library Outlet Closures by Year



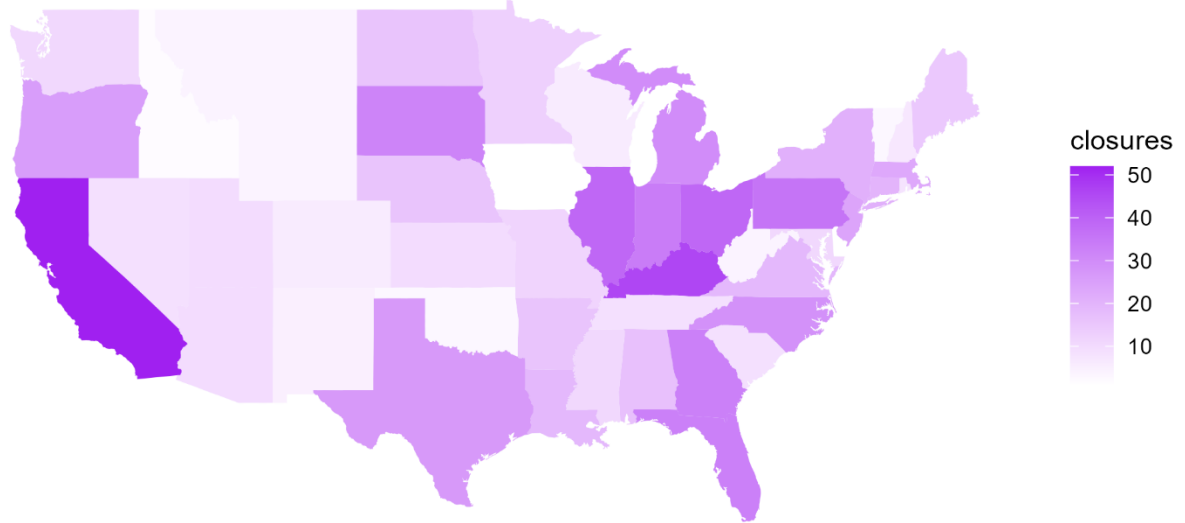
Source: IMLS PLS. Notes: Structural change codes (STATSTRU) include 2, 3, 8, 23, and 24.

Figure 2. Percentage of Library Closures in Rural Areas



Source: IMLS PLS. Notes: Structural change codes (STATSTRU) include 2, 3, 8, 23, and 24.

Figure 3. Public Library Closures by State



Source: IMLS PLS. Notes: Structural change codes (STATSTRU) include 2, 3, 8, 23, and 24. Years included are 2008-2021.

Table 2. Summary Statistics: Comparison of Means across Treatment Status

	Treated	Non-Treated	Difference
% Black Students	13.98%	11.48%	2.50%
% Hispanic Students	10.09%	10.15%	0.06%
% ECD Students	55.87%	56.03%	0.16%
Poverty Rate	16.60%	16.37%	0.23%
Unemployment Rate	8.25%	7.71%	0.54%
Log(HH Median Income)	10.69	10.68	0.01
Log(School Expenditures per Student)	12.09	12.00	0.09
Grade-Level Student Enrollment	1,156	530	626

Notes: Data Sourced from SEDA 5.0 dataset, spanning 2009 to 2018. Data on school expenditures is from the CCD Financial Survey (F-33) and data is nominalized to 2018 dollars using the GDP price index. County-by-Grade-by-Year Panel.

Table 3. Summary Statistics: County-by-Year Panel of Library Outcomes

	Obs.	Mean	Median	Min	Max
Total Visits (thousands)	26,400	254	91	0	16500
Total Circulation (thousands)	26,382	400	113	0	36400
Circulation of Children's Material (thousands)	26,385	133	34	0	15100
Attendance at Child Programs (thousands)	26,400	14	5	0	677
Public Internet Computer Use (thousands)	26,399	54	20	0	4909
Electronic Material (thousands)	14,718	36	6	0	8743
Ebook Collection (thousands)	26,335	72	8	0	5902
No. of Kid Programs (thousands)	26,398	0.22	0.55	0	34
Physical Materials (thousands)	26,383	157	75	0	7740
Hours Open (thousands)	26,400	9	6	0	216
No. of Total Staff	26,380	10	26	0	1752
No. of Librarians	26,365	4	10	0	448

Notes: Library data comes from the IMLS's PLS census of public libraries in the United States, spanning 2008-2019.

RESULTS

Impact on Student Achievement

In this section I estimate the effect of a library outlet closure on student achievement through SEDA test score data from 2009-2018 on students from 3rd to 8th grade.¹⁶ The two outcomes of interest are average math and reading test scores at the county-grade-year level. Test scores are normalized with mean zero and standard deviation one.

All Students

In table 4, analysis reveals a consistent decline in math and reading scores following outlet closures for all students.¹⁷ I employ three models: the standard Gardner model, one including covariates, and a third with covariates and state-time trends.¹⁸ Results consistently demonstrate a negative impact of closures on both math and reading scores across all specifications. Robustness to controls for community characteristics suggests that economic downturns do not drive these findings. Moreover, the resilience to state-time trends implies that closure effects are not solely due to temporal shifts within states, reinforcing the validity of the parallel trend assumption.¹⁹

It is estimated that average math test scores decrease by 0.034 standard deviations and average reading scores decrease by 0.021. Notably, math scores are more impacted than reading

¹⁶ Study period is shorter compared to library use analysis due to the availability of the SEDA data.

¹⁷ Results were also calculated isolating the Great Lakes region, due to the large number of closures in that area. However, results were consistent with baseline model.

¹⁸ All models are weighted by student populations.

¹⁹ Model 2 is the preferred model for further analysis. When analyzing subgroups with time trends, the limited number of observations limits statistical power.

scores, which contrasts with previous research suggesting that the impact of libraries on student achievement should be hardest felt in reading scores (Gilpin et al., 2024).²⁰

Table 5 additionally provides estimates separated by race. Strong impacts are found on Black and Hispanic students in math however, no impact in reading. Black students' average math test scores drop by 0.047 standard deviations while 0.031 for Hispanic. Impacts for White students are consistent across both subjects. There is no evidence of the gap between White and both racial minorities students growing.²¹

Table 6 provides estimates separated by both gender and whether a student is economically disadvantaged.²² When comparing estimates by gender, males experience larger impacts than females. The gap between the two genders experiences a notable change, with the gap in math growing by 0.009 and reading by 0.013 standard deviations, on average. Children that are more economically disadvantaged, are impacted in the same way by a closure as economically-advantaged children.

Metro vs Non-Metro

In both table 7 and 8, results are calculated separately for metro and non-metro counties. Economic intuition would suggest that students in non-metro areas would be more impacted,

²⁰ A statistically significant difference does exist between the mathematic and reading coefficients. Neither estimate falls within the others 95% confidence interval.

²¹ Comparisons across individual racial groups should be done with caution due to the differences in the number of observations. Many counties do not have students from racial minorities. Results are replicated in A2 where comparison can be made, with observations only included if there are test scores reported for White, Black, and Hispanic students.

²² The number of observations for subcategories are lower due to data entry issues. For example, if more that 5% of test scores did not have an assigned gender, then it was not recorded. This explains why both male and female estimates are above the estimate for all students.

however results calculated for all students suggest there is no difference across metro status. Coefficients are consistent across metro status for both subjects, however statistical power is lost.

An examination of individual racial groups within different metro statuses in table 7 reveals distinct impacts, particularly among Black students in non-metro areas.²³ On average, Black students experience a notable decline of 0.089 standard deviations in math test scores and 0.039 in reading scores following an outlet closure. Additionally, there is evidence indicating that White students are more adversely affected in non-metro areas compared to their counterparts in metro areas. However, no significant differential impacts were observed for Hispanic students.²⁴ These findings underscore a key aspect of this analysis by highlighting the differential impacts of closures within racial groups on educational outcomes in non-metro areas. Unfortunately, the PLS library data does not measure library use by race, so comparisons across usage cannot be made.

Insights into the disparity in both gender and ECD across metro status are provided in table 8. Interestingly, the gap between male and female grades only exists in metro areas. This is driven by females in metro areas not being impacted. This may point to females being more adaptable in their learning and able to use the substitutes a metro area provides more effectively. Additionally, males may be more reliant on library services, however unfortunately I do not have the data to show that. In terms of ECD status, there is no evidence of the gap changing between the rest of the student body, however impacts for ECD students themselves are more profound in non-metro area.

²³ Table A3 again provides comparable race estimates.

²⁴ No differences were found when comparing racial groups to each other, just discrepancies within individual racial groups.

Elementary vs Middle School

Table 9 shows the relationship between library outlet closures and student achievement by school level. The data is split into elementary and middle school with elementary representing grades 3-5 and middle school 6-8. There is no difference when comparing the two groups in reading outcomes, however there is a statistically significant difference in mathematics.

Elementary children are more impacted by the closure with an average decrease of 0.040 standard deviations compared to middle school children who experience a 0.024 decrease. This suggests that the role in which libraries play in children's mathematics development is heightened in their younger years. Children at a young age are still developing their relationship with education and libraries play a key role in facilitating that journey.

Event Study

In figure 4 and 5, the relationship between both math and reading test scores and library outlet closures, are calculated in an event study framework. This still uses the Gardner (2022) framework, however with event time treatment variables. The model includes all controls included in the preferred model from table 4. These results should be interpreted with caution due to the limited time-span of the analysis, however is a useful instrument to address the parallel trends assumption.

Examining math test scores in figure 4 reveals a notable decline following closure, with the most pronounced impact occurring in the year of closure, resulting in a 0.022 standard deviation decrease. Although these effects gradually diminish over time, they remain statistically significant and negative in the year following closure, suggesting an immediate but likely temporary impact, with scores returning close to baseline levels within two years post-closure.

Notably, statistical power decreases with time elapsed from closure. Pre-treatment trends exhibit consistency and stability, indicating comparability between the treated and untreated groups in the pre-treatment period.

A similar pattern is found when examining reading scores in figure 5. There is a sharp but temporarily impact from the closure, with statistically significant negative impacts found in the year of and post closure before effects dissipate. Again, the pre-treatment trend is flat, raising no cause for concern with respect to parallel trends.

Impact on Libraries

Having established a relationship between library closures and student achievement, I assess the effects of library outlet closures on the supply and demand of library services using PLS data from 2008-2019.²⁵ This is done to ensure that closing a library outlet actually has an impact on the use of the service. All outcomes are logged to allow for an elasticity interpretation.

Table 10 presents the impact of library closures on public library usage. Interestingly, no significant impact is observed on the circulation of materials for children or the general public. However, a noticeable decline is found in library visits post-closure. On average, total visits within the county decrease by 7.7%, with attendance at children's programs showing a more pronounced decline of 10.3%. This drop in attendance at children's programs, which offer interactive learning experiences beyond book services, likely contributes significantly to the observed decline in student achievement, particularly in math scores. The substantial impact on access to electronic materials, a service easily accessible from home, underscores the extent to

²⁵ Electronic Materials were not reported in the PLS dataset until 2013, therefore the analysis period for this variable is 2013-2019.

which individuals rely on library outlets for internet access, highlighting the broader community implications of library closures.

In Table 11, the supply of library services within a county after a closure are studied. Notably, there is a significant decrease in the availability of physical materials and the total number of opening hours offered to the public. This indicates that other library outlets within the county are not expanding their services to offset the impacts of closures. While there is no discernible impact on the availability of ebooks, a reduction in the number of child programs provided is evident. Interestingly, minimal disparities are observed across both supply and demand metrics when considering metro status.

In summary, the observed decreases in library visits, utilization of electronic resources, and attendance at library programs underscore the diminishing prominence of in-person services offered by public libraries. Notably, the decline in services like homework assistance correlates with significant impacts on math test scores. Moreover, the absence of these communal spaces represents a loss for children who rely on them. The evident reduction in scores shortly after the loss of such services, provides compelling evidence of the critical role played by valuable programming and homework assistance in educational outcomes.

Table 4. Impact of a Library Outlet Closure on Student Achievement

	(1)	(2)	(3)
Mathematics	-0.031*** (0.010)	-0.034*** (0.010)	-0.030*** (0.009)
Obs	125,069	125,069	125,069
Mean	-0.045	-0.045	-0.045
Reading	-0.017** (0.007)	-0.021*** (0.007)	-0.018** (0.008)
Obs	132,053	132,053	132,053
Mean	-0.042	-0.042	-0.042
Year Fixed-Effects	Y	Y	Y
County Fixed-Effects	Y	Y	Y
Grade Fixed-Effects	Y	Y	Y
Covariates	N	Y	Y
State-Time Trends	N	N	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Table 5. Impact of a Library Outlet Closure on Student Achievement, by Race

	Mathematics	Reading
Black	-0.047*** (0.014)	-0.009 (0.013)
Obs	45,839	49,239
Mean	-0.524	-0.477
Hispanic	-0.031** (0.014)	-0.015 (0.010)
Obs	48,504	49,192
Mean	-0.259	-0.307
White	-0.029*** (0.010)	-0.023*** (0.008)
Obs	118,488	125,064
Mean	0.082	0.090
White-Black Gap	-0.001 (0.010)	-0.013 (0.012)
Obs	43,055	46,310
Mean	0.596	0.583
White-Hispanic Gap	0.009 (0.011)	-0.008 (0.011)
Obs	47,058	47,865
Mean	0.340	0.450
Year Fixed-Effects	Y	Y
County Fixed-Effects	Y	Y
Grade Fixed-Effects	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level.

Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Table 6. Impact of Library Outlet Closure on Student Achievement, by Gender and ECD

	Mathematics	Reading
Male	-0.042*** (0.010)	-0.028*** (0.008)
Obs	116,387	122,600
Mean	-0.059	-0.162
Female	-0.033*** (0.010)	-0.015** (0.007)
Obs	116,387	122,600
Mean	-0.035	0.075
Male-Female Gap	-0.009** (0.004)	-0.013*** (0.005)
Obs	116,387	122,600
Mean	-0.024	-0.237
ECD	-0.034*** (0.010)	-0.016* (0.008)
Obs	107,768	114,025
Mean	-0.260	-0.272
Non-ECD-ECD Gap	0.007 (0.006)	0.003 (0.007)
Obs	107,768	114,025
Mean	0.515	0.536
Year Fixed-Effects	Y	Y
County Fixed-Effects	Y	Y
Grade Fixed-Effects	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. ECD stands for economically disadvantaged. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Table 7. Impact of a Library Outlet Closure on Student Achievement: Comparing Effects in Metro and Non-Metro Counties by Race

	Mathematics		Reading	
	Non-Metro	Metro	Non-Metro	Metro
All Students	-0.030*	-0.029**	-0.018	-0.016*
	(0.017)	(0.014)	(0.012)	(0.009)
Obs	86,598	38,471	90,070	41,983
Mean	-0.070	0.010	-0.070	0.016
Black	-0.089***	-0.025	-0.039*	0.001
	(0.029)	(0.016)	(0.024)	(0.015)
Obs	22,171	23,668	23,257	25,982
Mean	-0.574	-0.477	-0.536	-0.424
Hispanic	-0.033	-0.011	-0.031	0.004
	(0.037)	(0.014)	(0.023)	(0.012)
Obs	26,027	22,477	25,647	23,545
Mean	-0.289	-0.223	-0.362	-0.247
White	-0.038***	-0.020	-0.026**	-0.017*
	(0.014)	(0.014)	(0.011)	(0.010)
Obs	81,013	37,475	84,344	40,720
Mean	0.053	0.142	0.058	0.156
White-Black Gap	-0.012	-0.007	-0.001	-0.020
	(0.018)	(0.013)	(0.021)	(0.015)
Obs	19,756	23,299	20,788	25,522
Mean	0.547	0.637	0.556	0.604
White-Hispanic Gap	-0.025	0.002	-0.032	-0.017
	(0.026)	(0.011)	(0.024)	(0.012)
Obs	24,764	22,294	24,507	23,358
Mean	0.378	0.424	0.441	0.458
Year Fixed-Effects	Y	Y	Y	Y
County Fixed-Effects	Y	Y	Y	Y
Grade Fixed-Effects	Y	Y	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level.

Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Table 8. Impact of a Library Outlet Closure on Student Achievement: Comparing Effects in Metro and Non-Metro Counties by Gender and ECD

	Mathematics		Reading	
	Non-Metro	Metro	Non-Metro	Metro
All Students	-0.030*	-0.029**	-0.018	-0.016*
	(0.017)	(0.014)	(0.012)	(0.009)
Obs	86,598	38,471	90,070	41,983
Mean	-0.070	0.010	-0.070	0.016
Male	-0.048***	-0.035**	-0.033***	-0.025**
	(0.016)	(0.014)	(0.013)	(0.010)
Obs	37,614	78,773	40,757	81,843
Mean	0.005	-0.089	-0.103	-0.192
Female	-0.041***	-0.026*	-0.021**	-0.012
	(0.015)	(0.013)	(0.011)	(0.009)
Obs	37,614	78,773	40,757	81,843
Mean	0.023	-0.063	0.138	0.043
Male-Female Gap	-0.007	-0.009*	-0.010	-0.012**
	(0.007)	(0.005)	(0.008)	(0.006)
Obs	37,614	78,773	40,757	81,843
Mean	-0.018	-0.026	-0.241	-0.235
ECD	-0.044***	-0.024*	-0.029**	-0.006
	(0.002)	(0.014)	(0.013)	(0.010)
Obs	36,321	71,447	39,480	74,545
Mean	-0.265	-0.257	-0.274	-0.271
Non-ECD-ECD Gap	0.008	0.000	0.010	-0.009
	(0.009)	(0.009)	(0.010)	(0.010)
Obs	36,321	71,447	39,480	74,545
Mean	0.568	0.488	0.584	0.511
Year Fixed-Effects	Y	Y	Y	Y
County Fixed-Effects	Y	Y	Y	Y
Grade Fixed-Effects	Y	Y	Y	Y

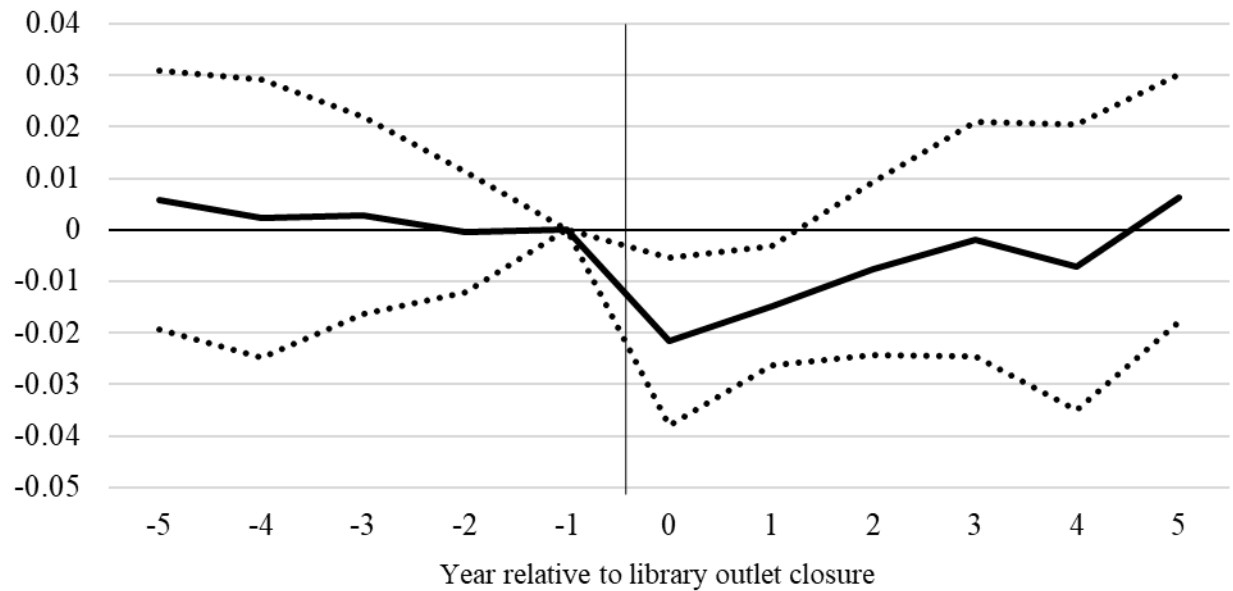
Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. ECD stands for economically disadvantaged. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Table 9. Impact of Library Outlet Closure on Student Achievement, by School Type

	Mathematics	Reading
Grade 3-5	-0.040*** (0.011)	-0.022*** (0.009)
Obs	75,424	74,355
Mean	-0.044	-0.035
Grade 6-8	-0.024** (0.011)	-0.019** (0.008)
Obs	49,645	57,698
Mean	-0.052	-0.065
Year Fixed-Effects	Y	Y
County Fixed-Effects	Y	Y
Grade Fixed-Effects	Y	Y

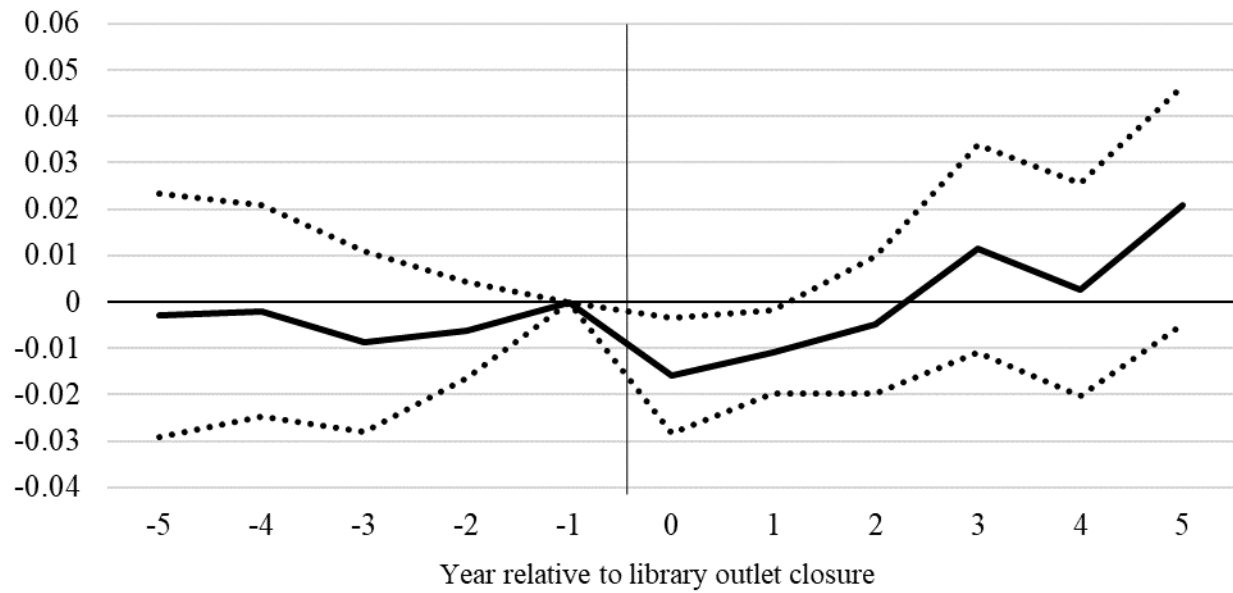
Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Figure 4. Impact of a Library Outlet Closure on Mathematics Test Scores



Notes: Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variable are reported in standard deviations. Dotted lines represent the 95% confidence interval. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, event-study model. Covariates included are discussed in table 2. Standard errors are clustered at the county level.

Figure 5. Impact of a Library Outlet Closure on Reading Test Scores



Notes: Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variable are reported in standard deviations. Dotted lines represent the 95% confidence interval. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, event-study model. Covariates included are discussed in table 2. Standard errors are clustered at the county level.

Table 10. Impact of a Library Outlet Closure on Public's Use of Library Services

		Non-Metro	Metro
Total Visits	-0.077** (0.035)	-0.093 (0.067)	-0.064** (0.028)
Total Circulation	-0.016 (0.037)	-0.042 (0.070)	-0.002 (0.028)
Circulation of Children's Material	0.006 (0.044)	-0.056 (0.082)	0.054 (0.034)
Attendance at Child Programs	-0.103** (0.047)	-0.130 (0.083)	-0.079 (0.048)
Public Internet Computer Use	-0.065 (0.046)	-0.096 (0.079)	-0.020 (0.049)
Electronic Material	-0.668** (0.270)	-0.662** (0.326)	-0.393*** (0.129)
Observations	26,404	18,114	8,290
Year Fixed-Effects	Y	Y	Y
County Fixed-Effects	Y	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Data comes from the IMLS's PLS census of public libraries in the United States, between 2008-2019. Children are defined as people aged 11 or under, as per The National Center for Education Statistics (NCES). Outcome variables are logged, using $\log(\text{var} + 1)$ due to the existence of zeros and numbers less than 1. Data on electronic materials is only available from 2013 onwards, therefore the number of observations is lower than stated. Model used is a Gardner (2022), population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Standard errors are provided in brackets and are clustered at the county level.

Table 11. Impact of a Library Outlet Closure on the Supply of Library Services

		Non-Metro	Metro
Physical Materials	-0.156*** (0.026)	-0.139*** (0.047)	-0.070*** (0.018)
Hours Open	-0.105*** (0.025)	-0.195*** (0.050)	-0.118*** (0.020)
Ebook Collection	-0.261 (0.269)	-0.410 (0.365)	0.034 (0.395)
No. of Kid Programs	-0.093* (0.049)	-0.135 (0.087)	-0.045 (0.048)
No. of Total Staff	-0.088*** (0.024)	-0.112** (0.044)	-0.064*** (0.020)
No. of Librarians	-0.080** (0.033)	-0.106* (0.062)	-0.054* (0.028)
Observations	26,404	18,114	8,290
Year Fixed-Effects	Y	Y	Y
County Fixed-Effects	Y	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Data comes from the IMLS's PLS census of public libraries in the United States, between 2008-2019. Children are defined as people aged 11 or under, as per The National Center for Education Statistics (NCES). Outcome variables are logged, using $\log(\text{var} + 1)$ due to the existence of zeros and numbers less than 1. Model used is a Gardner (2022), population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Standard errors are provided in brackets and are clustered at the county level.

IDENTIFICATION CONCERNS

Student and Community Characteristics

To assess the validity of the parallel trend's assumption, I employ an event study framework to examine the impact of library outlet closures on student and community characteristics. Figures 6-12 present the findings.

While there is a marginal decline in median household income around the time of closure, no statistically significant coefficients are detected. Similarly, there is no observable impact on unemployment rates or the racial composition of student bodies following library closures. Notably, Figure 10 reveals a positive correlation between school spending and library closures, suggesting that funds saved from closures may be redirected to schools to mitigate any negative effects on student achievement.

However, the rise in poverty rates following closures is concerning (figure 11). While the pre-treatment trend remains consistent, indicating no differential trends between treated and untreated groups, there is a statistically significant increase of 0.024% in poverty rates in the year of closure. Despite its statistical significance, the economic significance of this increase is limited. Given the short-lived nature of the impact and the absence of similar trends in other characteristics, it is unlikely to be the sole driver of the results.

Additionally, given the consistency of trends surrounding the closure period, this variation will be accounted for in the model. While it is conceivable that there may be another factor influencing both closures and poverty levels that has not been controlled for, this scenario is improbable, considering the inclusion of various socioeconomic variables in the model.

Random Assignment

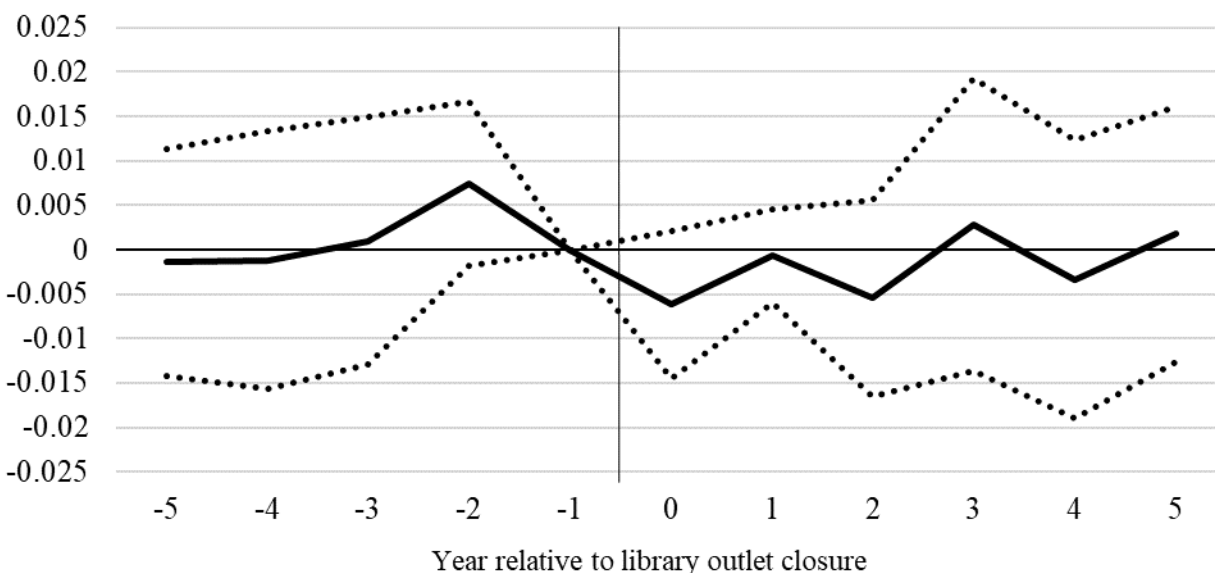
An issue to address, is that counties are able to decide which library outlet they close to cut costs. These decisions will be made with how that specific closure will affect the public at mind, especially when it comes to distance traveled.²⁶ However, with counties minimizing their impacts on the communities, this will dampen estimates. The closure of a randomly selected library outlet should have more profound effects on both library use and student achievement than the closures of the libraries being studied.

Public School Libraries

It is possible, in an effort to combat the closure of a local public library, that public schools within the county increase their library service offerings. This may offset any impacts found on student test scores, underestimating the findings. Using data from the Common Core of Data (CCD), I find that this is not the case. Following a public library closure, I find no effect on either the number of library specialists or assistants within the public-school network. Therefore, the results found on student achievement are not impacted by changes in school library services.

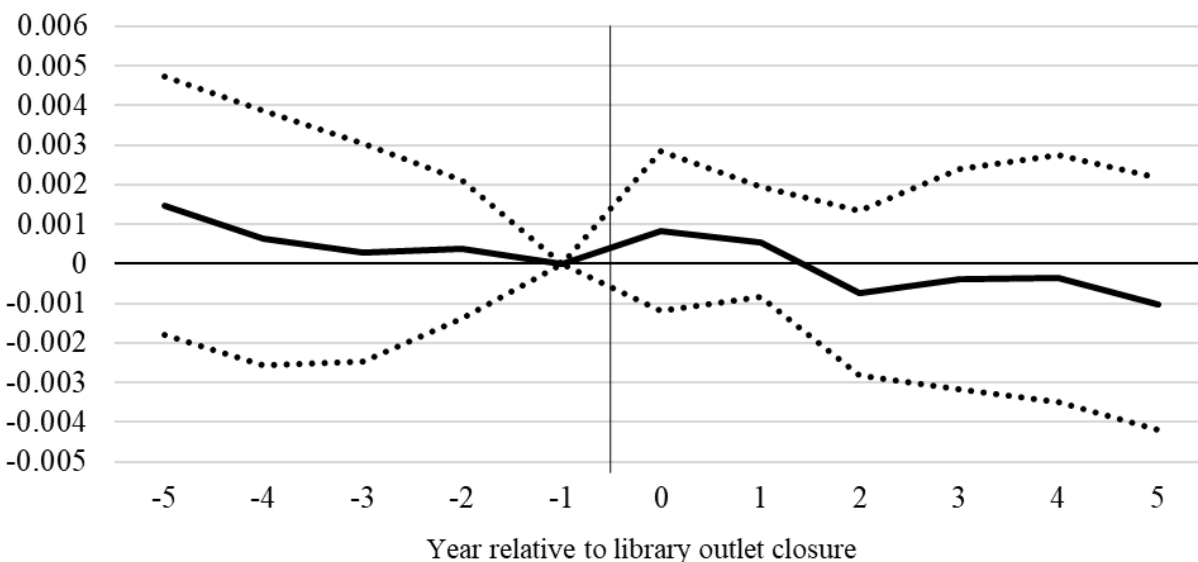
²⁶ The histogram in figure A1 illustrates this with a large right skew to the change in public travel distance to library outlets following a closure.

Figure 6. Correlation Between a Library Outlet Closure and Median HH Income



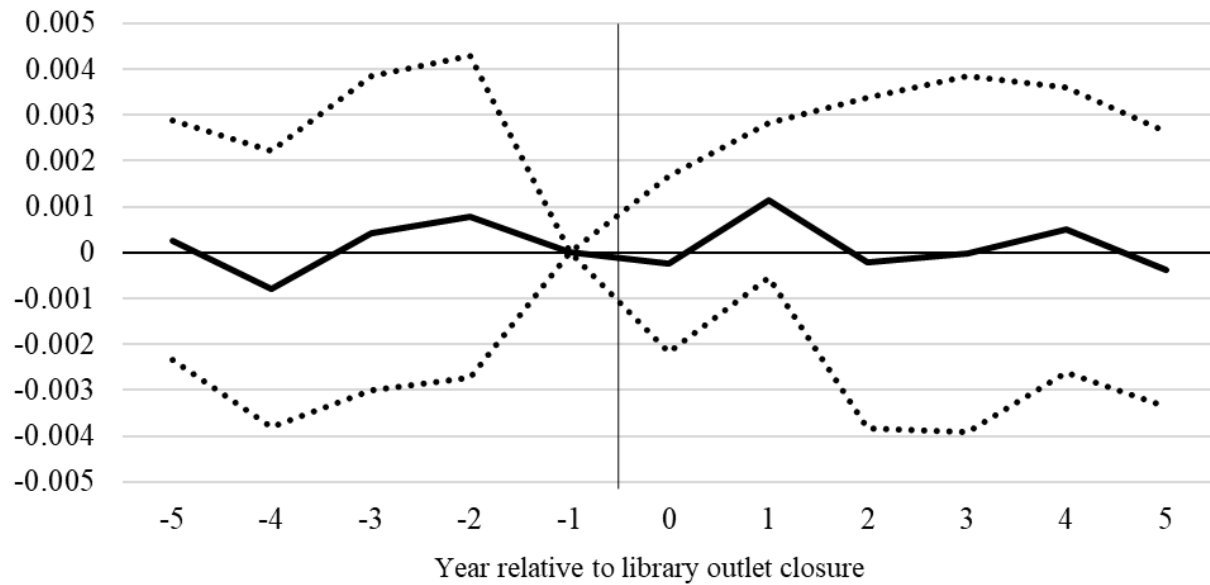
Notes: Data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variable is logged. Dotted lines represent the 95% confidence interval. Model used is a Gardner (2022), event-study model. Standard errors are clustered at the county level.

Figure 7. Correlation Between a Library Outlet Closure and Unemployment Rate



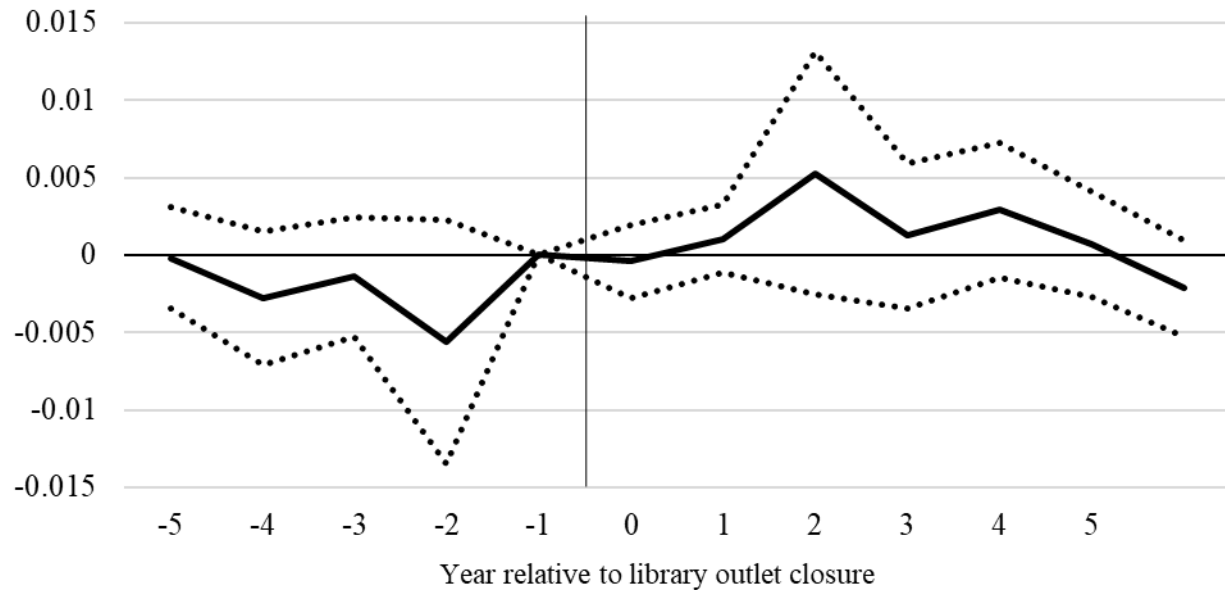
Notes: Data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Dotted lines represent the 95% confidence interval. Model used is a Gardner (2022), event-study model. Standard errors clustered at the county level.

Figure 8. Correlation Between a Library Outlet Closure and Percentage of Black Students



Notes: Data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Dotted lines represent the 95% confidence interval. Model used is a Gardner (2022), event-study model. Standard errors clustered at the county level.

Figure 9. Correlation Between a Library Outlet Closure and Percentage of Hispanic Students



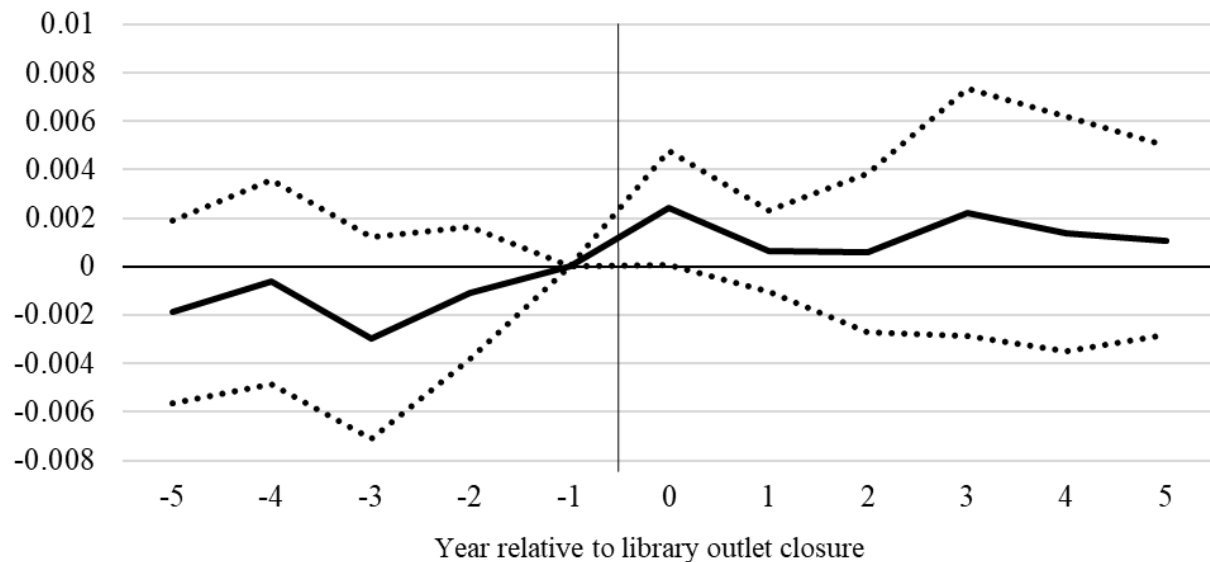
Notes: Data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Dotted lines represent the 95% confidence interval. Model used is a Gardner (2022), event-study model. Standard errors clustered at the county level.

Figure 10. Correlation Between a Library Outlet Closure and School Expenditures



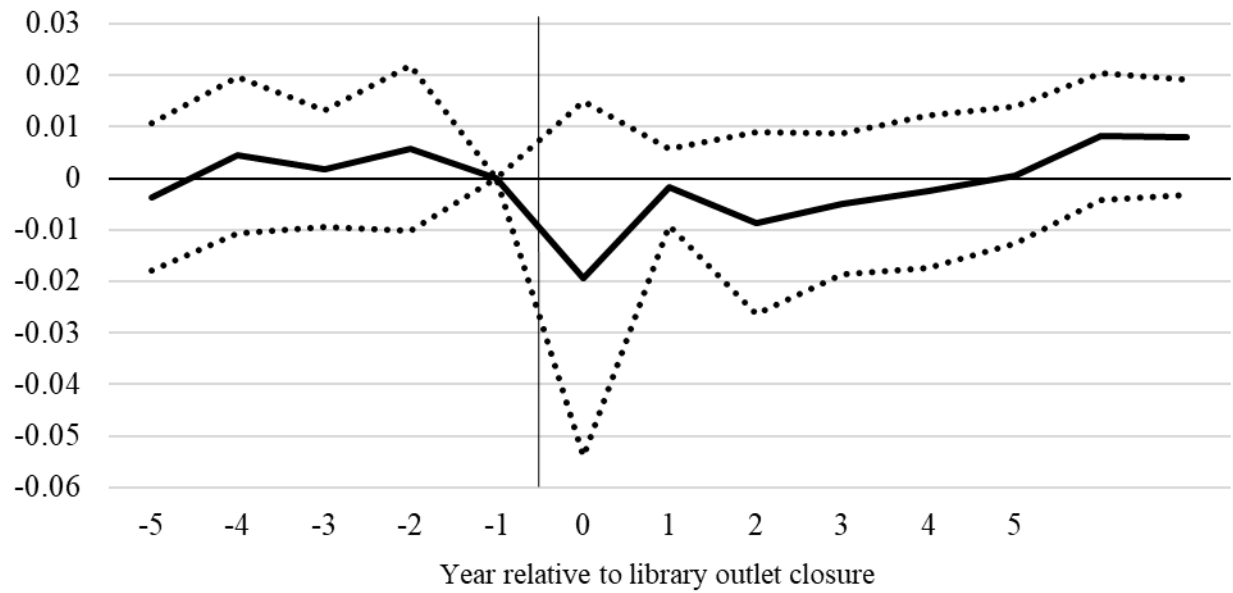
Notes: Data on school expenditures is from the CCD Financial Survey (F-33), between 2009-2018. Dotted lines represent the 95% confidence interval. Model used is a Gardner (2022), event-study model. Standard errors are clustered at the county level.

Figure 11. Correlation Between a Library Outlet Closure and Poverty Rate



Notes: Data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Dotted lines represent the 95% confidence interval. Model used is a Gardner (2022), event-study model. Standard errors clustered at the county level.

Figure 12. Correlation Between a Library Outlet Closure and Student Population



Notes: Data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Dotted lines represent the 95% confidence interval. Model used is a Gardner (2022), event-study model. Standard errors clustered at the county level.

CONCLUSION

This study delves into the multifaceted impacts of public library closures on communities and student achievement, highlighting differential effects across population subgroups. It reveals that closing a library outlet negatively influences both library service utilization and student achievement. Gender disparities are evident, with males experiencing, on average, a 0.01 standard deviation larger impact than females across both reading and math. Elementary school students show larger impacts on math scores compared to middle school students, although no differences are observed in reading. Notably, racial minorities only experience an effect on math scores. While little difference was observed across metro status, the effects for Black students in non-metro areas were notably amplified.

Throughout this analysis, various measures were taken to address potential biases associated with the location and timing of library closures. Despite inherent challenges and imperfections in studying non-random changes, these steps contribute to building confidence in the results. Notably, the limited relationships found between community characteristics and library closures help alleviate concerns that closures are merely a consequence of broader county-wide socioeconomic downturns.

This research fills a critical gap in understanding the effects of library closures, highlighting their societal significance and demonstrating the enduring relevance of libraries in contemporary society. Despite potential biases in closure decisions, the evidence suggests tangible impacts on society. Additionally, no statistically significant evidence was found regarding the impact of a library closure on operating, capital, or total library spending. Hence, the costs incurred by diminished student achievement are not offset by benefits from reductions

in library spending following closure. Furthermore, I find no observed impact on public school library staff levels, suggesting that schools are not compensating for the closure by increasing their own offerings. Therefore, based on this cost-benefit analysis, it's evident that closures are negatively impacting children and society as a whole.

Further investigation is needed to comprehensively grasp the full extent of these impacts. The relatively limited exploration of this topic within the field of student achievement suggests that the consequences of closures may extend far beyond current understanding, affecting society at a deeper level.

REFERENCES CITED

- American Library Association (2020). *School Libraries & Education*
<http://www.ala.org/advocacy/school-libraries>
- Berkes, E., & Nencka, P. (2021). *Knowledge Access: The Effects of Carnegie Libraries on Innovation* (SSRN Scholarly Paper 3629299). <https://doi.org/10.2139/ssrn.3629299>
- Bhatt, R. (2010). *The Impact of Public Library Use on Reading, Television, and Academic Outcomes*.
- Borusyak, K., Jaravel, X., & Spiess, J. (2023). *Revisiting Event Study Designs: Robust and Efficient Estimation* (arXiv:2108.12419). arXiv. <https://doi.org/10.48550/arXiv.2108.12419>
- Brunner, E., Hoen, B., & Hyman, J. (2022). School district revenue shocks, resource allocations, and student achievement: Evidence from the universe of U.S. wind energy installations. *Journal of Public Economics*, 206, 104586. <https://doi.org/10.1016/j.jpubeco.2021.104586>
- Callaway, Brantly, and Pedro HC Sant’Anna. “Difference-in-differences with multiple time periods.” *Journal of Econometrics* 225, no. 2 (2021): 200-230.
- Cellini, S. R., Ferreira, F., & Rothstein, J. (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. *The Quarterly Journal of Economics*, 125(1), 215–261.
- Everyday Reading (2023). *Where You Can Get A Non Resident Library Card*.
<https://everyday-reading.com/where-you-can-get-a-non-resident-library-card/>
- Ferreira Neto, A. B., Nowicki, J., & Shakya, S. (2023). Do public libraries help mitigate crime? Evidence from Kansas City, MO. *Journal of Cultural Economics*.
<https://doi.org/10.1007/s10824-023-09497-4>
- Gardner, J. (2022). *Two-stage differences in differences* (arXiv:2207.05943). arXiv.
<https://doi.org/10.48550/arXiv.2207.05943>
- Gilpin, G., & Bekkerman, A. (2020). Households’ public library use across the school calendar. *Library & Information Science Research*, 42(2), 101012.
- Gilpin, G., Karger, E., & Nencka, P. (2024). The Returns to Public Library Investment. *American Economic Journal: Economic Policy*. <https://doi.org/10.1257/pol.20210300>
- IMLS. 2019. Public Library Survey Fiscal Year 2021 Data File Documentation and User’s Guide [technical documentation]
- Institute of Museums & Libraries Services (2019a) FY 2019 PLS: supplementary tables–tables 7–13A. Institute of Museums & Libraries Services, Washington, DC

- Institute of Museums & Libraries Services (2019b) FY 2019 PLS: supplementary tables—tables 14–25A. Institute of Museums & Libraries Services, Washington, DC
- Jackson, C. K., & Mackevicius, C. (2021). *The Distribution of School Spending Impacts* (Working Paper 28517). National Bureau of Economic Research.
<https://doi.org/10.3386/w28517>
- Karger, E. (2021). *The Long-Run Effect of Public Libraries on Children: Evidence from the Early 1900s* [Preprint]. SocArXiv. <https://doi.org/10.31235/osf.io/e8k7p>
- Kevane, M. J., & Sundstrom, W. A. (2016). *Public libraries and political participation, 1870-1940*.
- Kevane, M., & Sundstrom, W. A. (2014). The Development of Public Libraries in the United States, 1870-1930: A Quantitative Assessment. *Information & Culture*, 49(2), 117–144.
- Lynch, Kathleen, Lily An, and Zid Mancenido. (2021). The Impact of Summer Learning Programs on Low-Income Children’s Mathematics Achievement: A Meta-Analysis. (EdWorkingPaper: 21-379). Retrieved from Annenberg Institute at Brown University:
<https://doi.org/10.26300/da7r-4z83>
- Martorell, P., Stange, K., & McFarlin, I. (2016). Investing in schools: Capital spending, facility conditions, and student achievement. *Journal of Public Economics*, 140, 13–29.
<https://doi.org/10.1016/j.jpubeco.2016.05.002>
- Neilson, C. A., & Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, 120, 18–31.
<https://doi.org/10.1016/j.jpubeco.2014.08.002>
- Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., Saliba, J. (2024). Stanford Education Data Archive (Version 5.0). Retrieved from
<https://purl.stanford.edu/cs829jn7849>
- Rubin, Donald B. 1981. “The Bayesian Bootstrap.” *The Annals of Statistics*: 130-134.
- Shepard, D., & Bitter, M. (2010). *First-Ever National Study: Millions of People Rely on Library Computers for Employment, Health, and Education*. Bill & Melinda Gates Foundation.
<https://www.gatesfoundation.org/ideas/media-center/press-releases/2010/03/millions-of-people-rely-on-library-computers-for-employment-health-and-education#:~:text=Low%2Dincome%20adults%20are%20more,Internet%20at%20their%20public%20libraries.>

APPENDICES

APPENDIX A

SUPPLEMENTARY TABLES

Table A1. Impact of a Library Outlet Closure on Student Achievement, TWFE Model

	(1)	(2)	(3)
Mathematics	-0.031*** (0.011)	-0.032*** (0.012)	-0.030*** (0.010)
Obs	125,069	125,069	125,069
Mean	-0.045	-0.045	-0.045
Reading	-0.017** (0.009)	-0.020*** (0.008)	-0.018** (0.009)
Obs	132,053	132,053	132,053
Mean	-0.042	-0.042	-0.042
Year Fixed-Effects	Y	Y	Y
County Fixed-Effects	Y	Y	Y
Grade Fixed-Effects	Y	Y	Y
Covariates	N	Y	Y
State-Time Trends	N	N	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Grades that are included are 3rd through 8th. Model used is a student population weighted, two-way fixed effects model. Each cell is a separate regression. Covariates included are discussed in table 2. Standard errors are provided in brackets and are clustered at the county level.

Table A2. Impact of a Library Closure on Student Achievement, by Race (Comparable Model)

	Mathematics	Reading
Black	-0.018 (0.015)	-0.003 (0.013)
Obs	25,409	26,289
Mean	-0.483	-0.443
Hispanic	-0.036** (0.015)	-0.009 (0.011)
Obs	25,409	26,289
Mean	-0.225	-0.270
White	-0.025* (0.013)	-0.025** (0.010)
Obs	25,409	26,289
Mean	0.165	0.172
Year Fixed-Effects	Y	Y
County Fixed-Effects	Y	Y
Grade Fixed-Effects	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level.

Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Table A3. Impact of a Library Outlet Closure on Student Achievement: Comparing Effects in Metro and Non-Metro Counties by Race (Comparable Model)

	Mathematics		Reading	
	Non-Metro	Metro	Non-Metro	Metro
Black	-0.042	-0.011	-0.030	0.003
	(0.032)	(0.017)	(0.025)	(0.015)
Obs	18,025	8,885	18,973	8,778
Mean	-0.467	-0.515	-0.412	-0.509
Hispanic	-0.070	-0.024	-0.028	-0.001
	(0.051)	(0.015)	(0.021)	(0.013)
Obs	18,025	8,885	18,973	8,778
Mean	-0.218	-0.241	-0.243	-0.328
White	-0.067**	-0.018	-0.065***	-0.017*
	(0.030)	(0.013)	(0.023)	(0.010)
Obs	18,025	8,885	18,973	8,778
Mean	0.212	0.068	0.221	0.068
Year Fixed-Effects	Y	Y	Y	Y
County Fixed-Effects	Y	Y	Y	Y
Grade Fixed-Effects	Y	Y	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level.

Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

APPENDIX B

ALTERNATIVE SPECIFICATION

Results were additionally estimated using distance to library outlet as a continuous treatment variable instead of a binary public library closure. This was done to understand how sensitive the relationship found is to model design. Additionally, by using distance, it provides an additional dimension of variation that is unlikely correlated with events that caused a library outlet to close. Distance accounts for the spatial distribution of library accessibility, which may be influenced by various county-level factors (e.g., population density, income levels, urbanization). By incorporating distance into the analysis, I can better control for these unobserved determinants of library closure decisions that may simultaneously affect student achievement. This reduces the risk of omitted variable bias, which could lead to spurious correlations between closures and academic outcomes.

Data

In order to calculate how far the public is travelling to library outlets, data from the 2020 United States Census was used. They provide Centers of Population estimates, offering mean center population calculations across various geographic levels. Notably, block groups - comprising populations between 600 and 3000 people - were employed with 239,780 block groups across all 50 states and the District of Columbia. Alaska was excluded due to its vast geographical nature.

The dataset comprises of longitude, latitude, and population data for each block group. To assess the impact of library closures on travel distance, the locations of these block groups were compared with the geographical coordinates of library outlets. The analysis was conducted

one state at a time, due to the complexity of the calculations being run.²⁷ Therefore, for each block group, I calculated which was the nearest library outlet within the state, and then recorded what the distance between the two points is in miles.²⁸

After determining the distance from each census block group to the nearest library outlet, the calculations were aggregated to the county level. Distances at the block group level were weighted based on population relative to the county's total population. Subsequently, these weighted distances were aggregated to derive the total weighted distance for each county. This process was repeated for all years within the sample period, with population estimates from the 2020 census consistently used. Variations in the distance variable are solely attributed to changes in the number or location of library outlets.²⁹

Summary Statistics

Focusing on treatment counties, I calculated the average distance increase after the closure of a library outlet. These summary statistics are provided in Table A4 and as a histogram in Figure A1. These statistics reveal that, on average, a library closure increased the distance

²⁷ I acknowledge that this could be problematic in border cities like Kansas City and St. Louis, however on the whole this is unlikely to majorly manipulate my findings. Additionally, attending an out-of-state library outlet is likely not possible as library card holders typically have to show proof of local residents. An article by Everyday Reading (2023) shows that even in states where attending an out-of-state is permitted, there is often a \$50-\$100 a year fee for doing so.

²⁸ This comparison was facilitated using the 'sf' package in R Studio, with a CRS code of 4326. This method aligns closely with previous literature, albeit implemented using a different coding language (Donnelly, 2015).

²⁹ At this point counties that experience an opening, whether that be temporary or permanent, are removed from the data. These counties were included in the distance calculations to allow for a block group's closest library outlet being in a different county.

required to travel for individuals to attend another library in their county by 0.54 miles.

However, the median estimate was more conservative at 0.14 miles due to the existence of some large counties with limited library services. Disparities were observed between metro and non-metro areas with non-metro areas experiencing higher changes. Notably, Corson County, South Dakota stood out as the county experiencing the largest change when it closed its only outlet, causing a subsequent 12.27-mile increase for residents. The existence of these outliers and the subsequent right-skew in the data, provides justification as to why travel distance will be logged throughout the analysis.³⁰

Model

To provide a different dimension to estimating the impact of library closures on student test scores, a two-way fixed effects empirical strategy using a continuous treatment variable is employed. Distance to the nearest library outlet is measured in logged meters³¹. This transformation is applied to account for the non-linear relationship between travel distance and its effect on library usage. For instance, a small increase in distance may have minimal impact on walking commuters, but a subsequent closure could necessitate a shift to alternative transportation methods, potentially discouraging visits altogether due to logistical challenges,

³⁰ Some counties experience no change in distance after a closure. This is partially due to counties strategically closing branches in areas which will have the least impact on residents and limitations to the census block group data used.

³¹ Meters are used rather than miles because when distances are reported in miles, many observations are below 1. Therefore, by using meters, it preserves the accuracy of the logged values. The $\log(\text{var}+1)$ transformation is still used for consistency with library outcome variables.

such as parking constraints. When calculating impacts on student achievement, a linear-log model is used due to student achievement data being reported in standard deviations.

The model for calculating the effects of distance travelled on student achievement is:

$$Y_{gct} = \beta D_{ct} + \delta_c + \gamma_t + \alpha_g + v_{gct}$$

where notion matches that used in the Gardner model. The independent variable D_{ct} represents the continuous treatment of distance. In control countries, this variable will stay constant throughout the study's sample period. The only change in the distance variable comes from the closing of library outlets in treatment counties. By including county fixed effects, all distance measures for the control counties are adjusted to zero, with no change in distance throughout the sample period. Due to this analysis utilizing a linear-log specification, the $\hat{\beta} \cdot \log\left(\frac{[100+p]}{100}\right)$ transformation is used to convert estimates to standard deviations. A 'p' of 2 will be used to illustrate the impact of a doubling in the average required travel distance throughout the results section.

Results

Results found in table A5 are consistent with results presented in the Gardner model, however statistical power is lost, therefore only a statistically significant result is found for reading test scores. For a 100% increase in distance travelled, it is estimated that average reading test scores decrease by 0.015 standard deviations.

Table A6 additionally provides estimates separated by race. Black students are largely impacted by library closures. Average math test scores for black students drops by 0.051 standard deviations while average reading drops by 0.040. Strong evidence also points to the gap

between White and Black students growing in math with on average a 0.023 standard deviation change. White students are impacted, however just not to the extent of Black students. Hispanic students see no impact from changes in distance travelled.

Table A7 provides estimates separated by both gender and whether a student is economically disadvantaged. No impacts are found across gender. In math the gap between Non-ECD and ECD students grew on average by 0.011 standard deviations following a 100% increase in distance travelled.

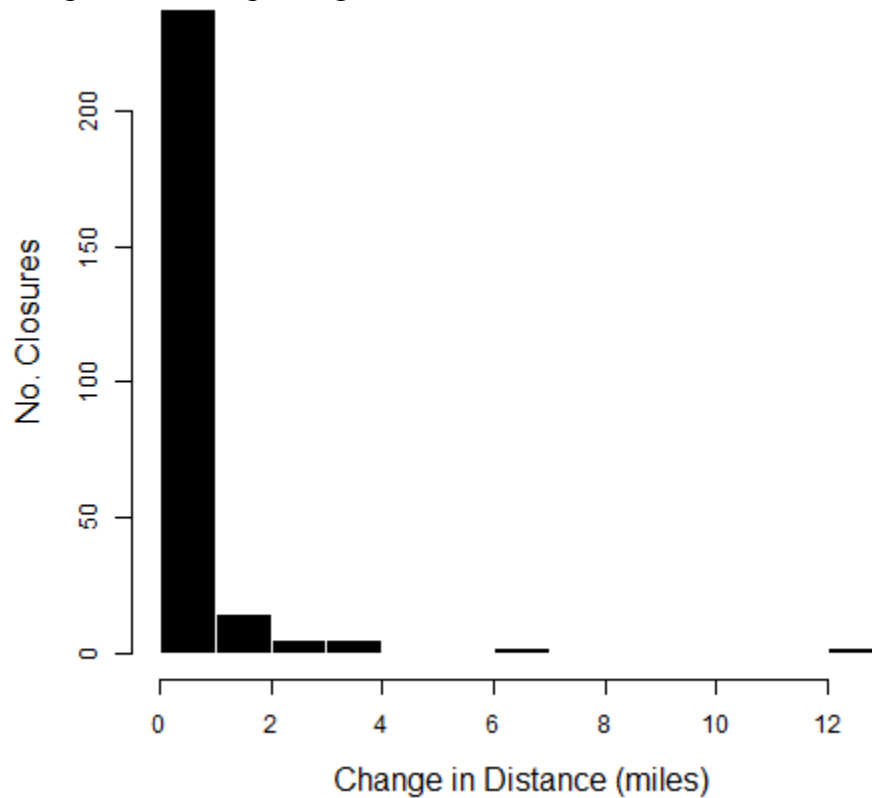
While these results may not be as robust as those obtained using a binary closure variable, they still offer valuable insights. Including this additional estimation allows policymakers to gain a more comprehensive understanding of the potential consequences of closing specific library outlets. It underscores the importance of considering distance as a crucial factor in mitigating the impacts of closures. This information can inform policymakers' decision-making processes and aid in implementing strategies to minimize adverse effects on the public.

Table A4. Summary Statistics: Change in Public Distance to Library Outlet Following Closure, by County

	Obs.	Mean	Median	Minimum	Maximum
All Counties	222	0.54	0.14	0	12.27
Metro	130	0.35	0.08	0	6.43
Non-metro	92	0.69	0.20	0	12.27

Notes: Library closure data comes from the IMLS's PLS census of public libraries in the United States. Distance is calculated by determining the distance between library outlets and Census Block Groups. These calculations are then aggregated at the county level. Data is reported in miles. This analysis was conducted between 2008-2019.

Figure A1. Histogram of Average Weighted Distance in Treatment Counties



Notes: Library closure data comes from the IMLS's PLS census of public libraries in the United States. Distance is calculated by determining the distance between library outlets and Census Block Groups. These calculations are then aggregated at the county level. Data is reported in miles. This analysis was conducted between 2008-2019.

Table A5. Impact of Distance to Nearest Library Outlet on Student Achievement

	(1)	(2)	(3)
Mathematics	-0.062	-0.051	-0.059
	(0.049)	(0.047)	(0.050)
Obs	125,069	125,069	125,069
Mean	-0.045	-0.045	-0.045
Reading	-0.054*	-0.035	-0.051*
	(0.029)	(0.027)	(0.029)
Obs	132,053	132,053	132,053
Mean	-0.042	-0.042	-0.042
Year Fixed-Effects	Y	Y	Y
County Fixed-Effects	Y	Y	Y
Grade Fixed-Effects	Y	Y	Y
Covariates	N	Y	Y
State-Time Trends	N	N	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Distance is logged with the $\log(\text{var} + 1)$ specification. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Standard errors are provided in brackets and are clustered at the county level.

Table A6. Impact of Public Distance to Nearest Library Outlet on Student Achievement, by Race

	Mathematics	Reading
Black	-0.170** (0.068)	-0.132** (0.062)
Obs	45,839	49,239
Mean	-0.524	-0.477
Hispanic	0.019 (0.024)	-0.017 (0.016)
Obs	48,504	49,192
Mean	-0.259	-0.307
White	-0.120** (0.048)	-0.094** (0.040)
Obs	118,488	125,064
Mean	0.082	0.090
White-Black Gap	-0.075* (0.039)	-0.014 (0.045)
Obs	43,055	46,310
Mean	0.596	0.583
White-Hispanic Gap	-0.004 (0.045)	0.027 (0.056)
Obs	47,058	47,865
Mean	0.340	0.450
Year Fixed-Effects	Y	Y
County Fixed-Effects	Y	Y
Grade Fixed-Effects	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Distance is logged with the $\log(\text{var} + 1)$ specification. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.

Table A7. Impact of Public Distance to Nearest Library Outlet on Student Achievement, by Gender and ECD

	Mathematics	Reading
Male	-0.057	-0.048
	(0.058)	(0.037)
Obs	118,754	125,096
Mean	-0.059	-0.158
Female	-0.067	-0.056
	(0.062)	(0.037)
Obs	118,466	124,888
Mean	-0.035	0.077
Male-Female Gap	0.009	0.002
	(0.018)	(0.018)
Obs	116,663	122,979
Mean	-0.024	-0.023
ECD	-0.068	-0.054
	(0.055)	(0.034)
Obs	118,518	124,931
Mean	-0.272	-0.279
Non-ECD-ECD Gap	0.037**	0.017
	(0.019)	(0.020)
Obs	108,044	114,397
Mean	0.511	0.535
Year Fixed Effects	Y	Y
County Fixed Effects	Y	Y
Grade Fixed Effects	Y	Y

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Student test score data comes from version 5 of the Standard Education Data Archive (SEDA), between 2009-2018. Distance is logged with the $\log(\text{var} + 1)$ specification. Outcome variables are reported in standard deviations. Grades that are included are 3rd through 8th. ECD stands for economically disadvantaged. Model used is a Gardner (2022), student population weighted, difference-in-difference model. Each cell is a separate regression. Covariates included are discussed in table 2. Mean represents the average difference, in standard deviations, between the sample and the national reference cohort for the outcome variable. Standard errors are provided in brackets and are clustered at the county level.