

THREE ESSAYS IN LABOR ECONOMICS

By

Salem Rogers

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2023

ABSTRACT

CHAPTER 1: Do Mid-Career Teacher Trainees Enter and Persist Like Their Younger Peers? (with Jane Arnold Lincove)

In the context of an ongoing national conversation about teacher shortages, we build on prior literature on the efficacy of teacher certification pathways by comparing entry and exit patterns based on age at the time of initial certification. All trainees who complete a state certification process have invested substantial time and resources into entering teaching, competing employment opportunities and expectations might vary by age. We use both linear regression and discrete-time hazard models to examine employment and subsequent exit of newly certified teacher trainees in Michigan from 2011 to 2022. We find that while mid-career entrants in their 30s and 40s compose a small share of new certificates, they are more likely to enter a public-school teaching position and no more likely to subsequently exit than counterparts who were certified in their early 20s. Mid-career pathways also contribute to teacher diversity by attracting more Black and male teachers who enter and persist.

CHAPTER 2: Diminishing returns across the day: evidence from school schedules

Cognitive fatigue – the decline in cognitive performance over time during sustained cognitive demand – is thought to be an important determinant of productivity. I analyze how cognitive fatigue and time of instruction affect student performance as measured by class marks and state standardized test scores. Further, I examine if there exist heterogeneous effects across courses and student subpopulations such that a readjustment of the school schedule may result in efficiency gains. I use panel data from 386 high schools in North Carolina containing over 1 million student level observations from 2000 – 2019 within a fixed effects framework. I find that having an English or math class in the last block of the day decreases a student's GPA by 0.062

(0.007) and 0.064 (0.007), respectively. I further find having a math class in the first block of the day instead of the fourth improves standardized math test scores by as much as increasing teacher quality by one half of a standard deviation.

CHAPTER 3: Spending & achievement effects of increased funding to rural school districts:

Evidence from Wisconsin (with Riley Acton & Cody Orr)

We study the spending and achievement effects of increased state funding to rural American school districts by leveraging the introduction and subsequent expansion of Wisconsin's Sparsity Aid Program. We find that the program, which provides additional state funding to small and isolated school districts, increased spending in eligible districts by 2% annually. Districts mostly allocate the funds toward non-instructional areas, such as hiring additional administrative staff and increasing spending on general operations and food service. As a result, we do not find consistent evidence that the increased funding improved standardized test scores or changed postsecondary enrollment and completion patterns. However, our confidence intervals do not exclude positive effects for rural schools as large as those found elsewhere in the literature.

ACKNOWLEDGEMENTS

This dissertation would not have been possible with the support of many. A village, some might say.

TABLE OF CONTENTS

CHAPTER 1: DO MID-CAREER TEACHER TRAINEES ENTER AND PERSIST LIKE THEIR YOUNGER PEERS?.....	1
BIBLIOGRAPHY.....	31
APPENDIX.....	34
 CHAPTER 2: DIMINISHING RETURNS ACROSS THE DAY: EVIDENCE FROM SCHOOL SCHEDULES.....	39
BIBLIOGRAPHY.....	65
 CHAPTER 3: SPENDING & ACHIEVEMENT EFFECTS OF INCREASED FUNDING TO RURAL SCHOOL DISTRICTS: EVIDENCE FROM WISCONSIN	68
BIBLIOGRAPHY.....	116
APPENDIX.....	119

CHAPTER 1: DO MID-CAREER TEACHER TRAINEES ENTER AND PERSIST LIKE THEIR YOUNGER PEERS?

Introduction

Nationwide teacher shortages have highlighted gaps in the teacher preparation pipeline and the need to recruit more future teachers. While much discussion has centered on the need to build more robust educator preparation programs (EPPs), states can choose from many policy options when investing in the pipeline. Ideally, public investments will result in both long and productive teaching careers for the beneficiaries of these investments and substantial returns to public education. While most teachers still come through undergraduate education programs at universities, evidence suggests that traditional bachelor's degree programs might not be the most efficient target for further expansion. In many states, a large share of undergraduate education majors never enter teaching (Cowan et al., 2016; Goldhaber et al., 2022), and universities are seeing declining interest in education majors and teacher career preparation among first-time college students (Bartanen & Kwok, 2022). Policy innovations in teacher training often include efforts to bring mid-career professionals into teaching through alternative pathways such as graduate-level university programs, fellowship programs, and other accelerated routes for those who already have a bachelor's degree. Prior research suggests that students of alternative route teachers fare as well as students of graduates of undergraduate pathways (e.g., Goldhaber, 2000; Glazerman et al., 2006), but less attention has been paid to the entry and exit behavior of post-baccalaureate program completers (see Zhang & Zeller, 2016 for a notable exception). Job entry and persistence rates of EPP completers are critical factors in determining if investments in pathways are an efficient strategy to bolster the teacher workforce in the long run.

In this study, we focus on the relatively understudied influence of teacher age at career

entry. Using the framework for EPP typologies in Lincove et al. (2015), we differentiate university-based programs for undergraduates ("traditional" certificate pathways) from programs for individuals who have already completed a bachelor's degree ("alternative" certificate pathways), which may be offered by many types of organizations including universities, community colleges, school districts, and private for-profit and non-for-profit organizations. A key characteristic of many alternative certification pathways is that they attract mid-career teacher candidates who already have a bachelor's degree and some work experience, and thus are older than undergraduate trainees, who are typically certified in their early 20s. While mid-career professionals might provide an attractive new pool of potential teachers, it is unclear whether older teacher trainees will be more or less likely than the typical undergraduate to complete the training, enter a teaching job, and persist in a long-term teaching career. It is possible that trainees who enter older have a stronger commitment to teaching and are therefore more likely to be employed as teachers in the long run. Alternately, mid-career trainees might have higher opportunity costs in other industries that make them less likely to eventually enter and persist in a teaching position.

Using data from Michigan, we examine the relationship between age at initial teaching certification and both entry into public school teaching and subsequent exit from teaching. Certification is a critical step to teacher employment in all public schools in Michigan (both traditional public and charter schools) that requires completion of multiple years of pre-service coursework and student teaching. Thus, certification signals a substantial commitment to future teacher employment. However, over the eight-year period of this study, only 45% of those who completed EPP training and received a Michigan teacher certificate entered teaching in the year following certification, and 32% never taught in a Michigan public school across the length of

our panel. This suggests that substantial public and private investments in teacher training do not deliver the intended benefits.

To investigate the role of age in entry and persistence, we follow six cohorts of newly certified EPP completers and track them forward for up to eight years from initial certification through employment in the Michigan public school system. To understand the role of macro labor market conditions, the time-period studied ranges from 2011 to 2022, which includes both periods of over-supply of teachers and periods of teacher shortages. We use linear probability models and hazard/survival analysis to estimate differences in entry/exit probabilities across teacher age groups, controlling for other characteristics such as demographics and performance on pre-service exams. We also investigate differential age effects for high-demand teaching specializations such as STEM, special education, and English as a Second Language (ESL) endorsement. Finally, we examine how age interacts with race and gender to inform efforts to diversify the teacher workforce.

The results provide several insights into the efficacy of public investments in recruiting mid-career professionals into teaching. While most teachers were certified in their early 20s, rates of entry into teaching jobs were higher for those who entered in their 30s or 40s, both in unconditioned regressions and when conditioning on teacher characteristics associated with sorting into both EPPs and teaching positions. After entry, we find that older entrants are no more likely to exit teaching within the first five years of teaching than younger entrants. Gaps in entry probabilities by age are larger for male teachers, and Black teachers certified in their 30s and 40s were significantly less likely to exit than those certified in their early 20s. This suggests that investments in recruiting older teacher trainees might both increase and diversify the pool of potential teachers with greater efficacy in terms of long-term persistence in the profession.

Framework and Prior Literature

In this study, we build on three overlapping streams in the literature on teacher training and employment. First, we add to the large and broad literature on factors that influence entry into and exit from public school teaching. The extant literature documents both low rates of entry into the profession post-certification and high rates of exit post-entry (Goldhaber et al., 2021). Past studies highlight several factors that influence entry and exit behavior. Novice teachers are more likely to exit, with 14 percent leaving within one year and 33 percent within three years (Ingersoll, 2003). High ability college students are both less likely to enter the profession and more likely to exit (Podursky et al., 2004). Once employed, more effective teachers are more likely to remain in the classroom, while additional training, such as National Board Certifications, can predict exit instead of persistence or quality improvement (Goldhaber & Hansen, 2009; Goldhaber et al., 2010). We add to this literature by investigating the effects of teacher age, highlighting that age can serve as a proxy for the unexamined effects of entering teaching as a second career with opportunity costs in other employment sectors.

Second, we contribute to literature that investigates the implementation and effects of alternative pathways into teaching. While the overall objective is to understand effects of alternative pathways in general on the teacher supply and teacher quality, the literature highlights that alternative routes differ substantially both across and within states. Alternative routes generically refer to programs that enroll individuals who already have a bachelor's degree, most likely in a non-education field, and provide only the coursework and student teaching experiences directly required for teacher certification. Importantly for this study, alternative EPPs generically encompass both programs that target recent college graduates with no work experience (such as Teach for America) and programs that target older workers who may have

had long careers in another sector (such as mid-career teaching fellowships). Like alternative EPPs themselves, findings regarding the efficacy of alternative EPPs are heterogeneous, with some finding that alternative route options decrease teacher qualifications, while others find that the least restrictive alternative routes attract the most qualified prospective teachers (Shen, 1997; Boyd et al, 2011; Sass, 2015). Looking at the value-added effects of different program types in Texas on the standardized test performance of students of program graduates, Lincove et al. (2015) find that alternative programs produce teachers of similar quality to traditional, university programs and can be a critical source of teachers in communities that are not located near a large university. There is little information in the extant literature about whether alternative training programs that target older applicants have similar success to those that target new college graduates. Although our focus is on age rather than EPP pathways, we add to this literature by examining age as a mechanism. Further, because we are able to observe all certified trainees and not just employed teachers, we can focus on the relatively neglected step of initial entry for certified EPP completers, as well the decision to persist or exit once employed.

Finally, we draw on occupational change literature within the broader study of labor markets in general. These studies look at the propensity of workers to switch occupations and their reasons for doing so, noting that workers are drawn to a second profession in pursuit of both economic and intrinsic rewards (Chambers, 2002; Serow and Forrest, 1994; Zimmerman et al., 2020). Higher commitment suggests that mid-career trainees might be more likely to enter and persist in teaching than younger peers. However, low entry salaries and the steep learning curve for novice teachers might reduce economic and intrinsic benefits for mid-career teachers with outside options. The literature provides little evidence on the outcomes of those who do choose to switch careers. We begin to resolve this ambiguity by examining whether older individuals

who complete certification for an intended new career in teaching enter and persist in that career at a differential rate than their first-career peers.

Michigan Context and Data

Michigan offers a compelling context in which to study the entry and exit behavior of career-switchers. EPP providers are approved and regulated by the Michigan State Board of Education (SBE), which since 2005, has restricted the entry of new EPP providers to only those that target specific needs in the state. Currently, nine non-university providers EPPs train less than 3% of newly certified teacher candidates each year, and the majority come from 31 university based EPPs operated by education programs at public and private universities. These include 31 programs that target undergraduates, and 29 alternative programs that offer post-baccalaureate training. Thus, most new teachers of all ages in Michigan are trained by universities but on both traditional undergraduate and alternative post-graduate pathways. Historically, the number of initial certificates in Michigan exceeded the number of newly hired teachers, but after 2013-2014, the state began to experience a shortage (Stackhouse, 2017). In addition to the low rates of entry for newly certified trainees reported above, the pool of applicants had declines. From 2012 to 2020, the number of new certificates declined by 48%.

Like many states, Michigan also experiences a mismatch between teacher and student demographics. While 64% of students are white, 18% Black, and 9% Hispanic, our sample of Michigan teachers is 92% white and 78% female. Given the growing literature that makes clear the importance for student success of race and ethnicity-matched students and teachers (Todd and Wolpin, 2007; Reardon and Galindo, 2009; Fryer and Levitt, 2013; Lindsay and Hart, 2017; Harbatkin, 2021), it is important to understand which pathways are most accessible and attractive for males and trainees of color. Since 2020, Michigan has invested over \$1 billion in new teacher recruitment efforts including expanding alternative pathways, such as “grow your own”

district-run EPPs and mid-career fellowships, and increasing undergraduate EPP enrollment (Ackley, 2023). It is against this backdrop that we examine the historic entry and exit behavior of newly certified teachers by age group, with a special focus on differential effects among Black and male teacher trainees.

Our analysis employs data from administrative certification, testing, and employment records provided by the Michigan Department of Education (MDE) and the state's Center for Educational Performance Information (CEPI). Certification records are available from 2011-2012 school year to 2021-2022, and contain information on EPP, certificate type, and special teaching endorsements. We merged this information with teacher employment records from 2012-2013 through 2021-2022 to observe post-certification employment in Michigan public schools. These records contain employee demographic information, employer information, job assignments, and scores on the Michigan Test for Teacher Certification (MTTC), which is required for Michigan teacher certification.¹

To create our sample, we begin with all in-state-prepared individuals who were issued an initial Michigan teaching certificate between 2011 and 2017.² To focus on newly certified trainees, we exclude those who had a teaching certificate in a prior period, those with a first certificate that required prior teaching experience, and those with work history in a teaching role in a Michigan public school in a prior period (dating back to 2003). We also exclude those who are missing reading or math MTTC scores.

Importantly, the job market for teachers in Michigan changed substantially during the

¹ Until 2013, the MTTC was based on the Basic Skills Test (BST) in reading and math. Starting in 2014, the BST was replaced by the Professional Readiness Exam (PRE). To address the change in tests, we standardized scores within test and year to a mean of 0 and standard deviation of 1. Starting in 2018, candidates could also substitute recent SAT or ACT scores, and in these cases, we are missing data for MTTC. The impact of missing test data is discussed in Results.

² We include those issued a Standard, Standard CTE, or Interim Certificate. This excludes those issued Temporary Teaching Certificates, which are issued to teacher who were prepared at out of state EPPs.

period studied, and macro labor market conditions might interact with age in determining entry and exit. Some experienced a relatively tight teacher labor market in the early 2010s, when public school enrollment was dropping and the number of teacher candidates was growing. Others entered during subsequent periods of unmet demand, and some also faced classroom disruptions related to COVID beginning in 2019-20. For this reason, we group trainees into cohorts based on the first full school year for which they were eligible for teacher employment. For example, trainees issued an initial certificate from September 1, 2012 to Sept 1, 2013, would be assigned to the 2013-14 certificate-year cohort. We then follow each cohort for at least five years (up to eight years for earlier cohorts) to observe: 1) if and when they enter a teaching job in Michigan, and 2) conditioned on entry, if and when they exit. We report both aggregate results and results by cohort to observe how broader labor market conditions are a moderating influence on entry age.

Our primary independent variable of interest in age at certification, we identify through the individual's birth year and initial certificate date. There is no set age for completing undergraduate education, and we don't observe employment history outside of Michigan public school, so to better delineate between those who are likely entering straight from undergrad and those who are likely changing careers, we place those aged 22–25 years into the “20s” age group and those aged 30-49 in the “30s and 40s” age group.³ Anyone outside of these two age groups at initial certification is not included in our analysis. We also use administrative data to create additional independent variables for individual characteristics that might be associated both with employment outcomes and age at certification. From MTTC records, we create within-cohort z-

³ In section 5, we show that our results are robust to and not driven by different age classifications. See Figure A1 for the full distribution of certificates by age. The vast majority of those certified in their 20s are certified by age 25.

scores (mean=0, sd=1) in reading and math. From demographic records, we create indicators for Black, Latino/Hispanic, white, other races, and male. Finally, from certification records, we identify whether teachers have additional certificates in hard-to-staff areas (STEM, Special Education, and ESL) that might increase the probability of employment and persistence.

Summary Statistics

Table 1.1 describes characteristics and job entry rates for newly certified teacher trainees by certificate-year and age group for our analytic sample, which include those initially certified at 22-25 years old (Panel A), and those certified at 30-49 years old (Panel B). Across certificate-year cohorts, 11-13% of the analytic sample were in the older group. Older teachers across all cohorts have higher MTTC reading scores by approximately 0.25 standard deviations but slightly lower or comparable MTTC math scores. Older teachers also reflect substantially greater diversity by race and gender. For example, 2012-13 cohort members in their 20s were only 2% Black and 23% male, while older members were 9% Black and 34% male. The two age groups have similar rates of endorsement in STEM, special education, and ESL. We also note important differences between cohorts over time. In 2012-13, Michigan had an over-supply of trainees. Mean MTTC scores, as a proxy for academic aptitude, are highest in that year and decline substantially through 2017-2018 when teacher hiring was a greater challenge. Similarly, the number of newly-certified trainees is substantially higher in the early 2010s, as is the racial and gender diversity within both age groups. Table 1.1 also reports rates of entry into teacher jobs within three and five years by certificate-year and age group. Overtime, 3-year entry rates increase substantially from approximately 55% to over 70% as the teacher labor market shifted from an over-supply to the beginning of the current shortage. For every certificate-year, rates of both 3- and 5-year entry are higher for older teachers by as much as 8 percentage points.

We further illustrate the first five years of employment outcomes over time by age group in Sankey plots displayed in Figure 1.1. To illustrate five full years of outcomes, we include members of the 2012 to 2017 cohorts, with those certified in their early 20s displayed in Figure 1.1A and those certified in their 30s and 40s in Figure 1.1B. We divide employment outcomes into three categories: teaching, other, and not observed. The teaching category contains those employed in a teaching position on either the fall or spring count day in a Michigan K-12 public school, including charter schools. Other refers to employment in a public school or district office but in a non-teaching position. This category includes primarily substitute teachers, aides, and support staff.⁴ The not observed category contains those not employed in a Michigan public school who might be employed in another state or another sector. We construct these categories to be mutually exclusive, giving preference to counting an individual as a teacher. For example, someone who is employed as a teacher as well as a substitute or athletic director in the same year is counted in the teaching category and not in the “other” category. Data labels show the percent of the age group in each Sankey bar.

Several interesting trends emerge in these figures. During the first post-certification year, the proportion of newly certified trainees in teaching jobs is about 45% in both age groups. After year one, the proportion who in teaching jobs rises in both groups but is always higher for those certified in their 30s or 40s. By year 5, 49% of the younger group is teaching, compared to 54% of the older group. Over time, 4-5% of those employed as teachers exit each year in both age groups. We also observe continued entry each year that comes mostly through trainees who were employed in the “other” category in the prior year. This suggests that many newly certified

⁴ Due to small numbers, we also count administrative and principal positions as “other.” Moving to a principal position might be considered a positive outcome for a teacher, compared to, for example, working as an aide because a teacher position is not available. However, only 0.31% of our sample is observed in the role of principal during the length of the panel.

teachers accept other school district jobs while they wait for a teaching position and may use non-teaching roles to increase the likelihood of being hired for future teaching openings. There is very little inflow from the “not observed” category to teaching, suggesting that if newly certified teachers are not connected with public school employment within their first year, they are unlikely to ever enter the teaching profession. Finally, we observe that continued inflow from other employment is more common in the older age group than those certified in their early 20s.

The Sankey plots and summary statistics provide several insights into the teacher pipeline for policymakers. First, a large portion of Michigan trainees who complete certification never teach in the state’s public schools. Second, entry into public school employment soon after certification is critical for future teacher employment. And third, both initial and subsequent entry rates are higher among those who enter older.

Regression Analysis and Results

Entry into Teaching

Our first empirical objective is to estimate age group differences in the probability of entering a teaching position after initial certification. We do this first by sequential linear estimation for being employed as a teacher in the first, second, and third years after certification. We estimate for trainee i who was initially certified with cohort j :

$$Entry_{ijt} = \alpha + \beta_1 X_i + \beta_2 Older_i + \beta_3 (Older_i \cdot Black_i) + \beta_4 (Older_i \cdot Male_i) + \delta_j + \varepsilon \quad (1)$$

where $Entry$ is a binary variable equal to 1 if individual i has entered the teaching a teaching job in a Michigan public school prior to year t , and equal to 0 otherwise. X is a matrix of time-invariant individual characteristics including race/ethnicity, gender, MTTC scores, and extra pre-service endorsements in STEM, special education, or ESL that indicate preparation for harder-to-staff positions. $Older$ is equal to one if the teacher was 30-49 years old at initial certification and

equal to zero otherwise (i.e. age 22-25 years). We include both the average difference by age group (β_2) and interactions between age group and race and gender (β_3 and β_4 , respectively) to measure both the age difference overall and age differences relative to critical areas of teacher diversity.⁵ δ_j controls for macro labor market conditions as a certification cohort fixed-effect, and ε is random error. For ease of interpretation of coefficients, we estimate (1) as a linear probability model through OLS.

Estimates of equation (1) are presented in Table 2, where column (1) displays the probability of entering by year 1, column (2) by year 2, and column (3) by year 3. The omitted group in these estimates is a white, non-Latina female who is 22-25 years old and has no extra endorsements at initial certification. Overall, both age groups are equally likely to enter in year 1. Teachers who are Black and with special endorsements are more likely to enter. Males are slightly less likely to enter than females, but the interaction term suggests that older males are significantly more likely enter. Combining the relevant coefficients and interactions, white males in their 30-40's are more likely to enter the Michigan public teacher workforce than younger white females by about 4 percentage points by the first year after certification.

Over time, we see a growing gap in entry rates where older teachers are more likely to enter by 2.0 percentage points by year 2, and by 4.9 percentage points by year 3. Black trainees across all three years are more likely to have entered than whites. Male trainees are always less likely to enter than females in the younger age group, but the interaction term between males and age suggests that males in their 30-40s are either more likely or similarly likely to enter than females in that age group.

We acknowledge that this is descriptive evidence that might be influenced by sample

⁵ Other racial subgroups in the sample are too small to test interaction terms.

selection decisions. One concern is that the somewhat arbitrary cut-points that we selected for age groups influence our results. In appendix Table A.1.1, we show that these results are robust to incrementally expanding the 20s age group all the way to including those certified ages 21-29. A second concern is, as noted above, that we excluded those without pre-service MTTC scores from our sample, this eliminated 2.5% of those certified in their 30s and 40s and 5% of those certified in their 20s. As trainees who do not go on to be employed are less likely to be associated with their test scores in the data set, those eliminated from the sample were less likely to enter teaching. The higher rate of missingness among the 20s groups likely biases our entry results downward, meaning the true effect of being older when certified on entry is likely higher than 4.9 percentage points.

We next use a discrete Cox proportionate hazard model to better model the role age in *time-to-entry*. Cox proportionate hazard models predict the likelihood of an event, in this case becoming an actively employed public school teacher, occurring across multiple observations of an individual. The considers not only whether a certified individual enters teaching each year, but also the influence of time on entry probabilities. In this case, it is likely that the probability of entering teaching diminishes over time for individuals not already employed as teachers, who might be employed in other sectors. The hazard model requires multiple observations of individuals over time, with each teacher remaining in the longitudinal data until they are first employed as a teacher. We follow trainees four years to eight years (depending on the certificate year relative to the end of the panel). The discrete Cox proportionate hazard is modeled as:

$$h(entry|X, Age\ Group) = k(t) + \exp(\gamma_2 Older) + \varepsilon \quad (2)$$

where $k(t)$ is the discrete baseline hazard function, which accounts for time. Results from three different specifications of equation (2) are presented in Table 1.3. The simple bivariate

relationship between age groups is presented in column (1), we include certificate-year cohort fixed effects in column (2), and in column (3) we add full controls for sex, race/ethnicity, specialized endorsements on initial certificate, and MTTC scores and interactions for age x Black and age x male. Across specifications, we estimate that older trainees have a significantly higher log entry hazard of 0.091-0.107, or approximately an 11% higher probability of entry than an otherwise similar younger trainee. The estimated difference due to age gets larger not smaller when adding control variables, suggesting that it is not differences in other characteristics that are driving the findings regarding age. The hazard model further demonstrates a positive relationship between Black trainees in their 30-40s and entry between age x male, though neither is statistically significant.

The tabled results mask cohort differences in the macro labor market, so we explore disaggregated cohort effects in more detail in Figures 1.2 and 1.3. Figure 1.2 shows estimates including only cohort fixed effects (Table 1.3, column 2), and Figure 1.3 includes full controls (Table 1.3, column 3). Here we display the estimated difference in entry probabilities by age across certificate-year cohorts. The baseline hazard is for a white female in her 20s from the 2012 certificate-year cohort, and all values on the y-axis are relative to that value. We see consistently higher entry rates across cohorts, as the Michigan labor market provided more open positions for fewer candidates over time. Across cohorts, we also see a consistently higher probability of entry for older trainees with larger gaps by age in the years when teaching jobs were more readily available. Figures 1.2 and 1.3 are quite similar suggesting again that very little of these differences would be attributable to observable differences in characteristics other than age.

Exit

We now turn our attention to estimating the effects of age at certification on time to career exit. Because a trainee cannot exit without entering, this analysis is conditional on ever having entered the teaching profession. Selection tendencies described in the previous analysis apply to the subsample of trainees who can be included in analysis of exit. Specifically, trainees who entered are more likely to be Black or Latina, more likely to have extra endorsements, and more likely to be older, older and black, and older and male, than the full analytic sample described in Table 1.1.

After conditioning the sample on having been employed as a teacher, we estimate age group differences in post-employment persistence using the same methods described above. We first estimate the probability that trainee has exited by years 1, 2, and 3 after initial employment as a linear probability model similar to equation (1), and then as a Cox hazard model similar to equation (2). The outcome variable is exit from teaching and the time variable is measured from initial employment rather than initial certification. We continue to control for the timing of first certification in models with certificate-year cohort fixed effects.

We present results for the exit version of equation (1) in Table 1.4. As above, the specifications include full controls, cohort fixed effects, and interactions for age x Black and age x male and follow employed trainees for three years after employment. We estimated that being certified in one's 30s or 40s instead of one's 20s is associated with no difference in exit one year after employment, but a growing difference in years 2 and 3. By year 3, older teachers are 2.3 percentage points less likely to have exited teaching than a similar younger teacher. Black teachers and male teachers are more likely to exit overall. In the case of Black teachers, a large significant and negative coefficient on the interaction term suggests that older Black teachers are

substantially less likely to exit than younger Black teachers and also, (summing up the coefficient on Black and the interaction term) less likely to exit than older white teachers. The opposite is true of older male teachers, who are more likely to exit than all other demographic groups.

Turning to the hazard results for exit, being older is associated with a 0.144 decrease in the log exit hazard. However, this is not statistically significant at the 10% level, and we cannot rule out that the likelihood of exit is the same for both age groups. Figures 4 and 5 illustrate exit probabilities as marginal differences due to age across cohorts. Unlike entry probabilities, exit probabilities decline as cohorts entered in friendlier labor market conditions. In the specifications with only cohort controls (Figure 1.4) we see only small differences in exit probabilities by age. Adding additional controls (Figure 1.5) substantially increased the difference by age for each cohort. Although these differences are not statistically significant, they suggest that the differential sorting into employment of trainees by age, gender, and race might also influence differential exits.

Discussion and Policy Implications

With nationwide concern about teacher shortages, it is critical for research to inform public investments in expanding and diversifying the teacher pipeline. States and districts have many choices to shore up the supply of teachers ranging from various pipeline expansions to raising salaries to improving working conditions. Part of making good investments in training is ensuring that EPP graduates enter and persist in successful teaching careers. Recognizing that teacher training might be useful in other settings as well, public investments in the teacher pipeline are intended to produce a stable supply of teachers for public schools. State-level investments might be viewed as less effective if a state's teacher trainees exit for other states,

private schools, or other sectors. Losing nearly 50% of fully trained candidates each year from the teacher labor market is a substantial loss of public and private investment that might have benefited public schools.

Substantial prior research has investigated differences between alternative and traditional EPP pathways, but often with greater focus on quality and persistence of those who enter teaching jobs than on the rates of entry themselves. We learn from prior research that alternative pathways typically produce teachers of similar quality to traditional pathways on average (Goldhaber and Brewer, 2000). Further, Sass (2015) finds that EPPs vary substantially in selectivity and quality, and that the best candidates often prefer the most flexible programs. As undergraduate interest in teacher training and employment declines, older college graduates who might pursue a career change might provide an untapped pool of future teachers. While prior work recognizes that alternative EPPs vary in the age of candidates they target and attract, prior research has not explored the direct influence of age on teacher employment outcomes. This study begins to fill this gap.

First, our descriptive analysis identifies that trainees often take several years to gain a teaching position, but an early connection to public school employment predicts future employment, consistent with Goldhaber et al. (2022). With a coinciding shortage of substitute teachers, districts might improve employment and retention through programs that onboard teachers through substitute teaching or other positions designed to feed into permanent employment. Second, the large pool of certified but not employed trainees is a potential source for outreach during times of teacher shortages. Qualitative research in Michigan points to low salaries and a lack of career growth opportunities as the most common reasons why certificated

candidates leave teacher employment (REL Midwest, 2021). Further research is needed to see what if anything might induce these qualified individuals to enter teaching.

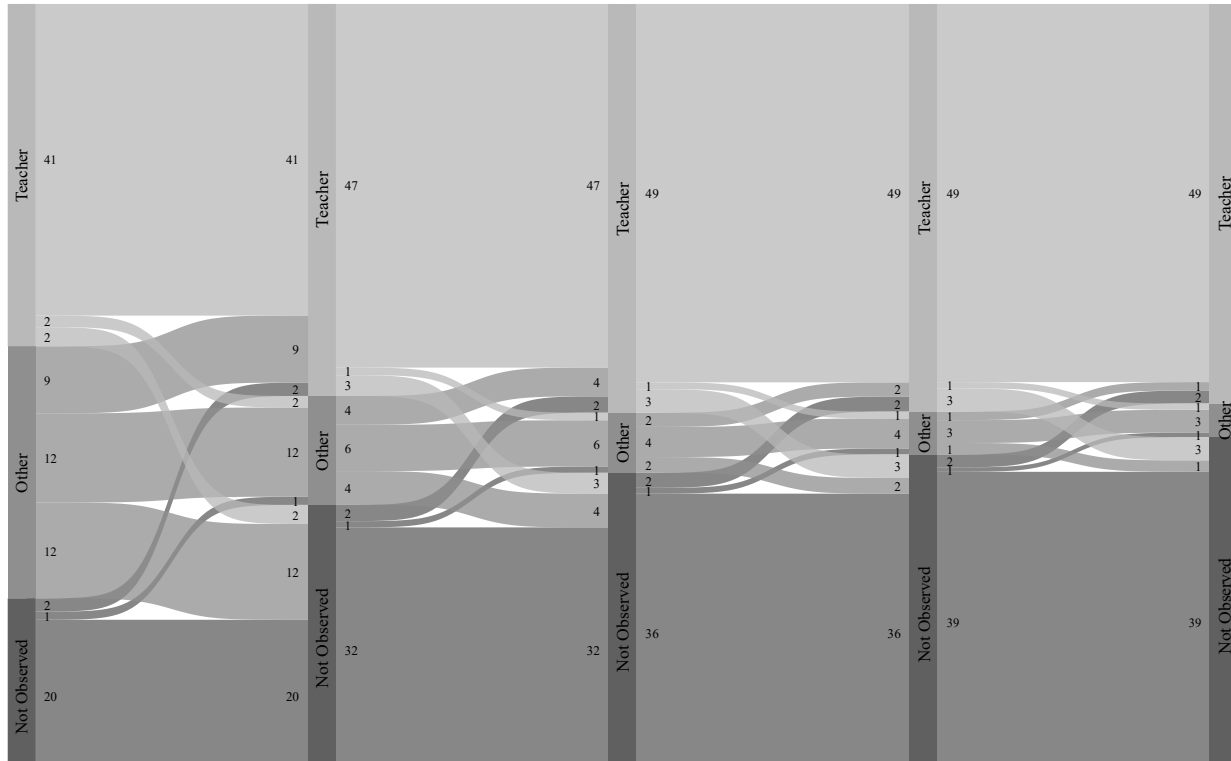
Our findings regarding trainee age suggest that while the pool of older teacher candidates in Michigan has been small, among those who show a strong enough commitment to complete certification requirements, older candidates are more likely to become public school teachers and potentially to persist in longer careers. Our findings further suggest that investments in alternative pathways that are attractive to older candidates might also meet the goal of increasing diversity in states like Michigan where the number of Black and male teachers leave many students without same-race or same-gender teacher role models. It is possible that Black males entering college in Michigan do not conceive of themselves as future teachers, because they have been exposed to few Black male teachers in the K-12 system (Goings and Bianco, 2016). We note that Michigan has since 2005 restricted the entry of new EPP providers and thus the market of alternative programs is smaller than other states. In this setting, our sample of teachers who enter mid-career might be smaller and more motivated than in other states. As more states expand alternative pathways to overcome shortages, more research is needed on the relationship between age and outcomes as more older teachers are recruited (Moseley, 2023).

Acknowledging a potential benefit of recruiting older teachers, particularly older males and older Black males and females, suggests that it might be productive to focus on alternative EPPs that cater to mid-career switchers more than recent college graduates. Changes might range from recruitment strategies that target job searchers instead of undergrads to structural changes in the way coursework is delivered. For example, models that offer flexible schedules and nighttime or weekend coursework might be more effective than those that require full-time enrollment. Finally, we note that school districts can also make adjustments that would attract

and retain older teachers who, in theory, are attracted both through commitment and rewards. For example, peer mentoring and professional development from other mid-career switchers might be more appropriate for older entrants, and signing bonuses or credit for non-teaching work in setting initial salaries might overcome the higher opportunity costs for those with other work experience. Overall, our results suggest further investigation and investment in mid-career switchers could improve the teacher pipeline in the future.

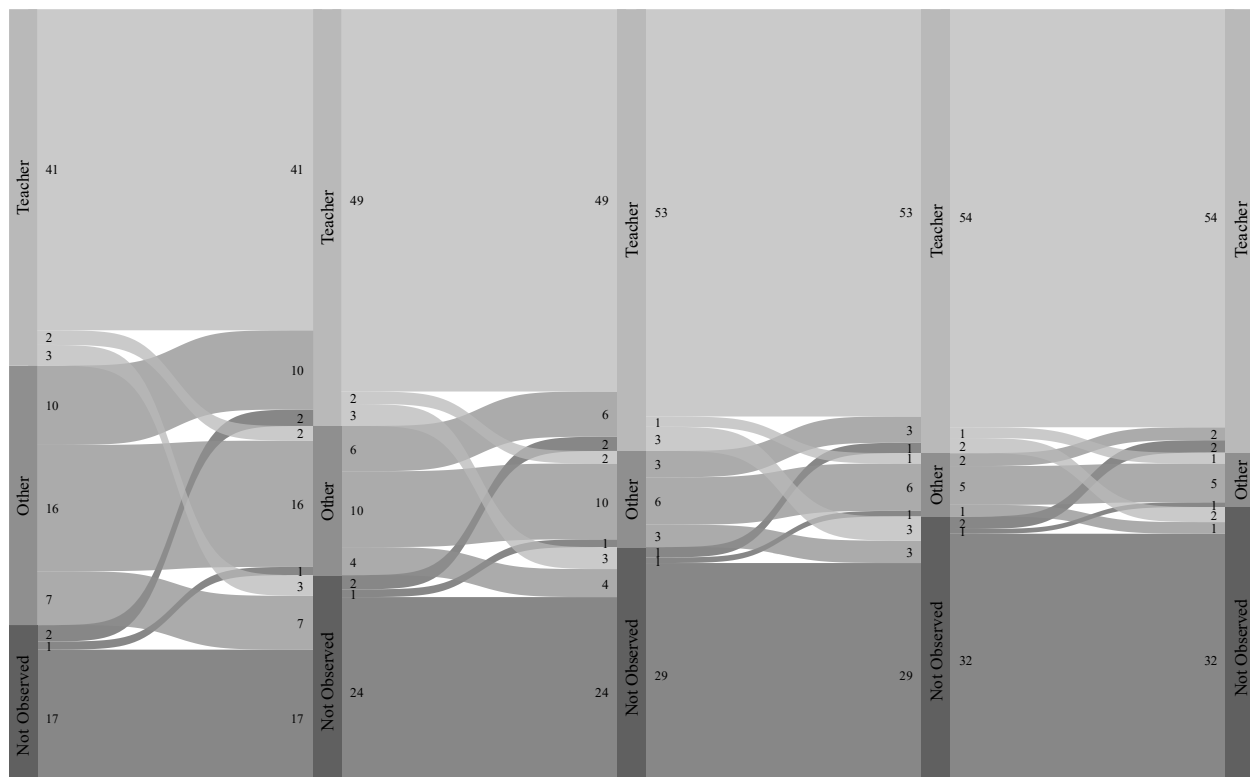
TABLES AND FIGURES

Figure 1.1.A Five Year Employment Outcomes for 2012 – 2017 Cohorts: Trainees Certified in their Twenties



Note: Teaching = Employed in a teaching position on either the fall or spring count day in a Michigan K-12 public school. Other = Employed in a non-teaching position on either the fall or spring count day in a Michigan K-12 public school. While this category does include principals and other administrators, we note that only 0.31% of our sample is observed in the role of principal during the length of the panel. Not Observed = Not employed in a Michigan K-12 public school on either the fall or spring count day.

**Figure 1.1.B Five Year Employment Outcomes for 2012 – 2017 Cohorts: Trainees
Certified in their Thirties or Forties**



Note: Teaching = Employed in a teaching position on either the fall or spring count day in a Michigan K-12 public school. Other = Employed in a non-teaching position on either the fall or spring count day in a Michigan K-12 public school. While this category does include principals and other administrators, we note that only 0.31% of our sample is observed in the role of principal during the length of the panel. Not Observed = Not employed in a Michigan K-12 public school on either the fall or spring count day.

Figure 1.2. Hazard Model Estimates of Age Effect by Cohort (no additional controls)

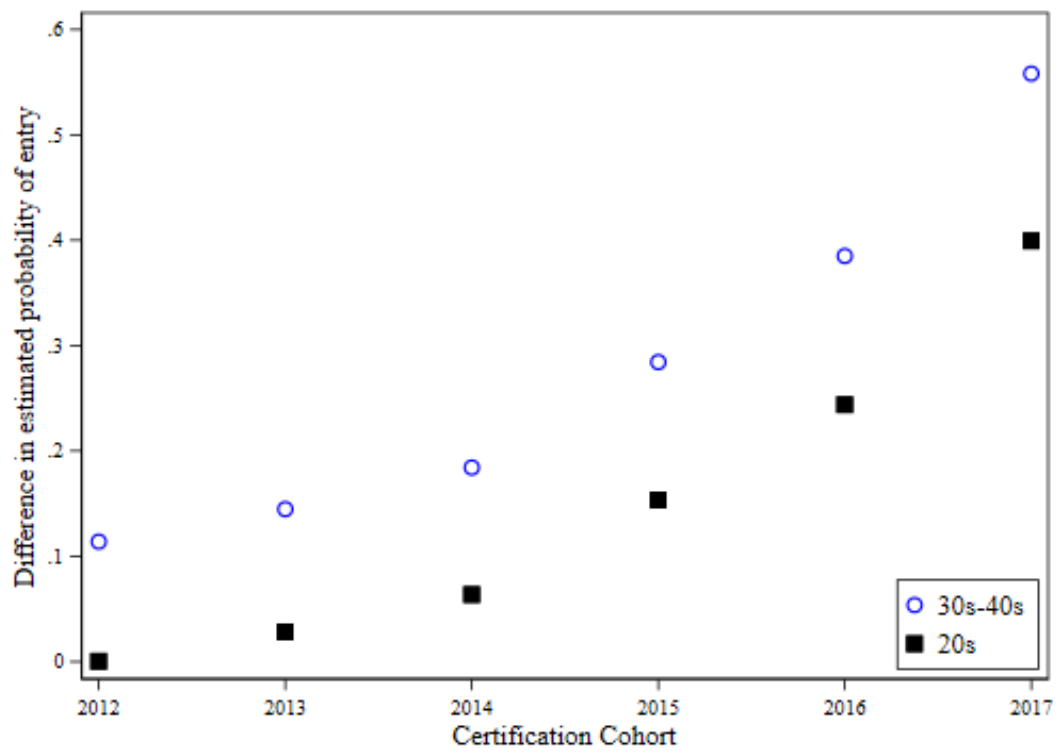


Figure 1.3. Hazard Model Estimates of Age Effects on Entry by Cohort (full controls)

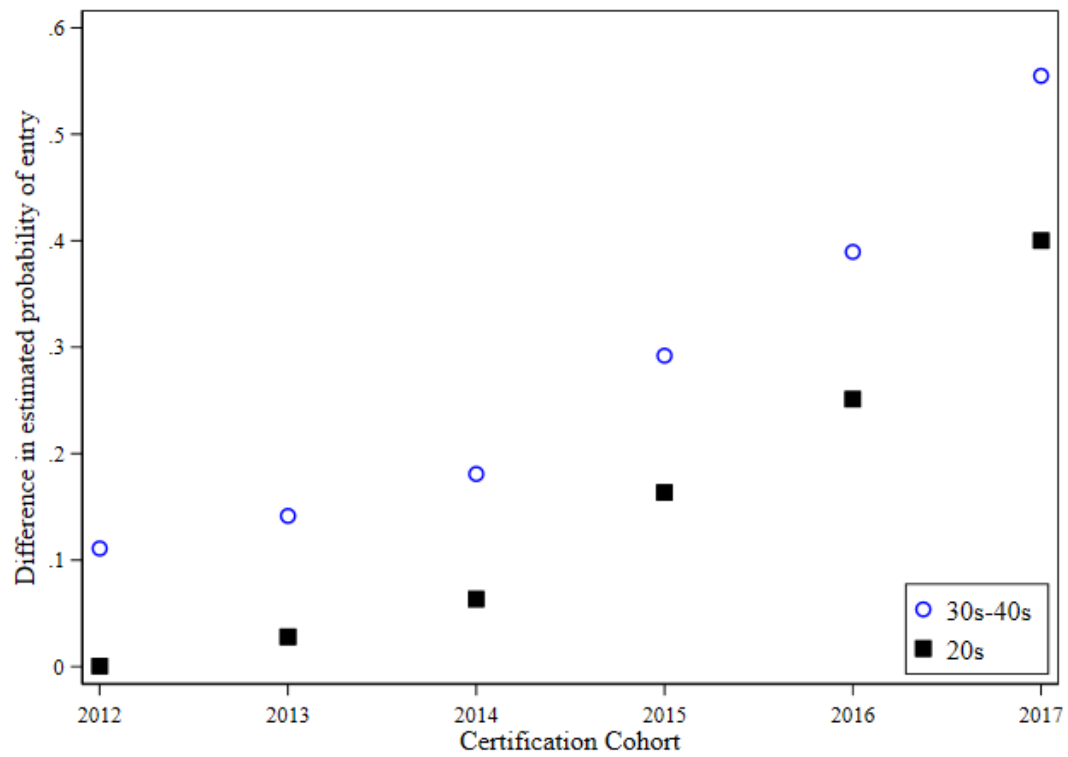


Figure 4. Hazard Model Estimates of Age Effect on Exit by Certificate-Year Cohort (no additional controls)

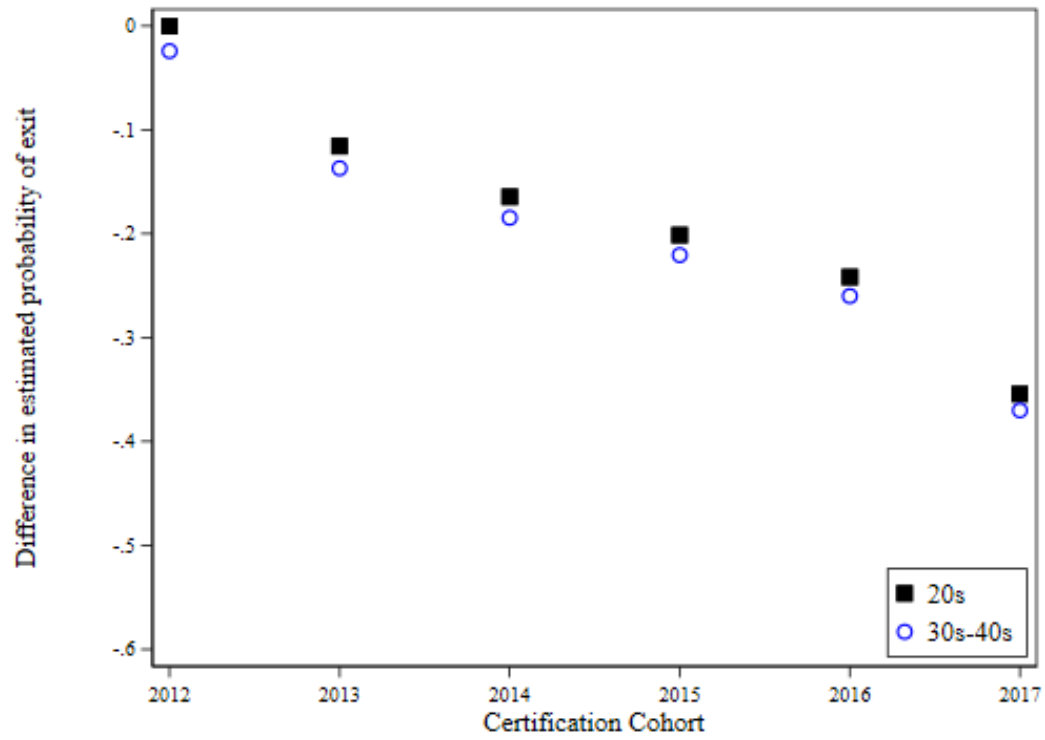


Figure 5. Hazard Model Estimates of Age Effect on Exit by Cohort (full controls)

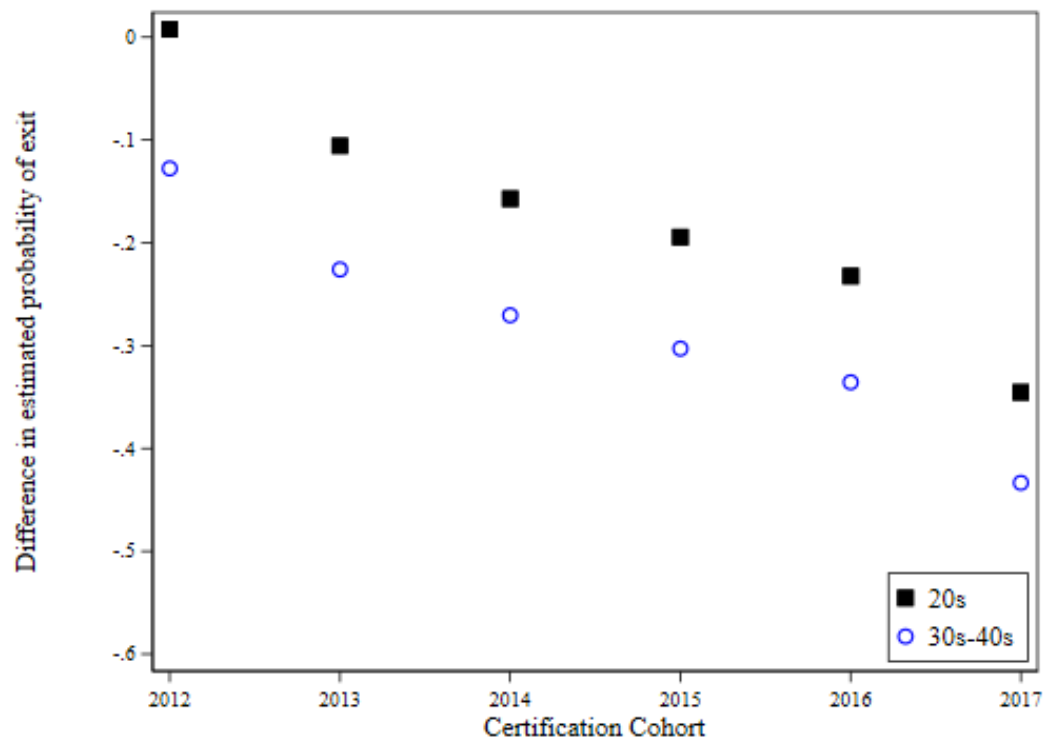


Table 1.1. Characteristics of Newly Certified Teachers in Michigan

Year	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18
N	2,584	2,895	2,708	2,529	2,129	1,542
Panel A – Teachers Initially Certified in their 20s						
Percent of new certificates	86.2	88.7	87.7	88.7	88.6	86.6
MTTC reading z-score	.11	.09	.05	-.00	-.15	-.36
MTTC math z-score	.38	.35	.30	.20	.03	-.56
Percent Black	1.9	2.3	2.2	2.1	1.9	1.8
Percent Latino	1.4	1.2	1.4	1.2	1.5	1.6
Percent white	93.4	93.2	93.2	92.6	92.2	93.3
Percent other race	3.2	3.2	3.2	4.1	4.4	3.2
Percent male	22.7	21.5	19.5	20.8	18.9	16.2
Percent STEM certified	18.0	16.8	17.7	17.2	15.2	14.3
Percent SPECIAL EDUCATION certified	5.3	4.6	5.5	6.2	7.6	8.3
Percent ESL	0.8	1.1	1.1	1.3	2.1	2.5
Percent who enter within 5 years	60.0	60.8	63.2	66.2	68.7	74.1
Percent who enter within 3 years	55.3	55.3	57.3	61.1	64.9	71.0
Percent who exit within 3 years	17.3	17.2	17.0	15.3	14.0	10.5
Panel B – Teachers Initially Certified in their 30s or 40s						
Percent of newly certified	13.78%	11.26%	12.30%	11.27%	11.37%	13.10%
MTTC reading z-score	.37	.28	.33	.23	.30	.03
MTTC math z-score	.23	.24	.20	.04	-.05	-.39
Percent Black	8.99%	8.59%	6.91%	7.72%	4.13%	8.91%
Percent Latino	1.69%	1.53%	2.10%	3.16%	0.83%	3.96%
Percent white	85.11%	84.97%	84.68%	82.81%	85.95%	79.70%
Percent other race	4.21%	4.91%	6.31%	6.32%	9.09%	7.43%
Percent male	33.99%	32.82%	39.34%	30.18%	34.71%	36.14%
Percent STEM certified	18.26%	18.71%	20.12%	15.44%	18.60%	18.32%
Percent SPECIAL EDUCATION certified	4.49%	4.29%	5.41%	5.96%	4.55%	0.99%
Percent ESL	0.84%	0.31%	0.90%	0.70%	0.83%	1.98%
Percent who enter within 5 years	62.08%	70.25%	66.97%	74.74%	73.14%	77.72%
Percent who enter within 3 years	53.93%	64.72%	60.96%	71.93%	69.42%	74.26%
Percent who exit within 3 years	22.03%	14.22%	15.70%	16.11%	19.05%	9.15%

Table 1.2. Linear Probability Model Estimates for Teacher Entry

	Year 1	Year 2	Year 3
Older	-0.008 (0.003)	0.020*** (0.000)	0.049*** (0.000)
Black	0.173*** (0.003)	0.114** (0.002)	0.077** (0.002)
Latino	0.029 (0.006)	-0.004 (0.008)	-0.010* (0.000)
Other Race	-0.067 (0.045)	-0.126 (0.022)	-0.140* (0.018)
Male	-0.014*** (0.000)	-0.040** (0.001)	-0.038** (0.001)
Endorsement	0.100** (0.005)	0.078** (0.004)	0.059** (0.004)
30s & 40s x Black	-0.007 (0.002)	0.022* (0.002)	0.003 (0.002)
30s & 40s x Male	0.048*** (0.001)	0.053** (0.001)	0.028** (0.001)
MTTC Controls	X	X	X
Cohort Controls	X	X	X
Observations	14,387	14,387	14,387
R-squared	0.032	0.025	0.021

Note: Models include all certified teacher trainees in defined age group who were eligible to first enter teaching from 2012-13 through 2017-18. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, math, or science. *** p<0.01, ** p<0.05, * p<0.1

Table 1.3. Cox Proportional Hazard Estimates for Teacher Entry

	(1)	(2)	(3)
Older	0.091*** (0.030)	0.093*** (0.030)	0.107*** (0.038)
Black			0.154** (0.073)
Latino			-0.051 (0.086)
Other Race			-0.313*** (0.059)
Male			-0.088*** (0.028)
Endorsement			0.119*** (0.024)
30s & 40s x Black			0.011 (0.126)
30s & 40s x Male			0.017 (0.067)
MTTC Controls			X
Cohort Controls		X	X
Observations	53,104	53,104	53,104

Notes: Hazard is defined as initial entry into teaching in a Michigan public school teacher. Cohorts able to initially enter from 2012-13 through 2017-18 are tracked through 2021-22. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, or STEM. *** p<0.01, ** p<0.05, * p<0.1

Table 1.4. Linear Probability Model Estimates for Exit after Employment as a Teacher

	Year 1	Year 2	Year 3
Older	0.0004 (0.000)	-0.008 (0.002)	-0.023* (0.003)
Black	0.004 (0.001)	-0.006 (0.001)	0.066** (0.002)
Latino	-0.020** (0.000636)	0.005 (0.0141)	0.00460 (0.016)
Other Race	0.053 (0.013)	0.092 (0.042)	0.102 (0.051)
Male	0.006* (0.000)	0.015** (0.000)	0.031** (0.000)
Endorsement	-0.022 (0.008)	-0.028 (0.012)	-0.027 (0.017)
AgexBlack	-0.056*** (0.001)	-0.032** (0.002)	-0.052** (0.003)
AgexMale	0.022** (0.001)	0.030** (0.001)	0.036** (0.001)
MTTC Controls	X	X	X
Cohort Controls	X	X	X
Observations	9,240	9,240	9,240
R-squared	0.006	0.008	0.010

Note: Models include all certified teacher trainees in defined age group who were eligible to first enter teaching from 2012-13 through 2017-18. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, math, or science. *** p<0.01, ** p<0.05, * p<0.1

Table 1.5. Cox Proportional Hazard Estimates for Exit after Teacher Employment

	(1)	(2)	(3)
Older	-0.026 (0.065)	-0.024 (0.065)	-0.144* (0.088)
Black			0.299** (0.120)
Latino			-0.100 (0.184)
Other Race			0.377*** (0.108)
Male			0.043 (0.056)
Endorsement			-0.134*** (0.049)
30s & 40s x Black			-0.149 (0.227)
30s & 40s x Male			0.143 (0.138)
MTTC Controls			X
Cohort Controls		X	X
Observations	32,951	32,951	32,951

Notes: Hazard is defined as exit from teacher in a Michigan public school teacher. Cohorts able to initially enter from 2012-13 through 2017-18 are tracked through 2021-22. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, or STEM. *** p<0.01, ** p<0.05, * p<0.1

BIBLIOGRAPHY

- Ackley, M. (2023, July 27). *Michigan among nation leaders in addressing teacher shortage*. Michigan Department of Education. <https://www.michigan.gov/mde/news-and-information/press-releases/2023/07/27/michigan-among-nation-leaders-in-addressing-teacher-shortage>
- Bartanen, B., and Kwok, A. (2022). From Interest to Entry: The Teacher Pipeline from College Application to Initial Employment. (EdWorkingPaper: 22-535). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/hqn6-k452>
- Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., O'Brien, R., and Wyckoff, J. (2011). The effectiveness and retention of teachers with prior career experience. *Economics of Education Review*, 30 (2011), 1229– 1241.
- Chambers, D. (2002). The Real World and the Classroom: Second-Career Teachers. *The Clearing House*, 75(4), 212-217.
- Cowen, J., Goldhaber, D., Hayes, K., & Theobald, R. (2016). Missing elements in the discussion of teacher shortages. *Educational Researcher*, 45(8), 460–462.
- Fryer, R. and Levitt, S. (2013) Testing for racial differences in the mental ability of young children. *American Economic Review*, 103(2), 981–1005.
- Goings, R., and Bianco, M. (2016). It's Hard to be who You Don't See: An Exploration of Black Male High School Students' Perspectives on Becoming Teachers. *Urban Review*, 48, 628-646.
- Glazerman, S., Mayer, D., and Decker, P. (2006). Alternative Routes to Teaching: The Impacts of Teach for America on Student Achievement and Other Outcomes. *Journal of Policy Analysis and Management*, 25:75 96.
- Goldhaber, D. and Brewer, D. (2000). Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement. *Educational Evaluation and Policy Analysis*, Vol. 22, No. 2, pp. 129-145.
- Goldhaber, D., Gross, B., and Player, D. (2011) Teacher Career Paths, Teacher Quality, and Persistence in the Classroom: Are Public Schools Keeping Their Best? *Journal of Policy Analysis and Management*, 30(1), 57-87.
- Goldhaber, D., and Hansen, M., (2009). National Board Certification and Teachers' Career Paths: Does NBPTS Certification Influence How Long Teachers Remain in the Profession and Where they Teach? *Education Finance and Policy*, 4 (3), 229-262.
- Goldhaber, D., Krieg, J., Liddle, S, and Theobald, R. (2022). Lost to the System? A Descriptive Exploration of Teacher Candidates' Career Paths. *Educational Researcher*, 51(4), 255– 264.

- Harbatkin, H. (2021). Does student-teacher race match affect course grades? *Economics of Education Review*, 81:102081.
- Ingersoll, R. (2003). Is There Really a Teacher Shortage? Center for the Study of Teaching and Policy, R-03-4, 1-30.
- Lincove, J.A., Osborne, C., Mills, N, and Bellows, L. (2015). Teacher Preparation for Profit or Prestige: Analysis of a Diverse Market for Teacher Preparation. *Journal of Teacher Education*, 66(5), 415-434.
- Lindsay, C., and Hart, C. (2017) Exposure to Same-Race Teachers and Student Disciplinary Outcomes for Black Students in North Carolina. *Educational Evaluation and Policy Analysis*, 39(3), 1–9.
- Moseley, B. (2023, June 13). Gov. Kay Ivey signs alternative teacher certification legislation. *Alabama Today*. <https://altoday.com/archives/51988-gov-kay-ivey-signs-alternative-teacher-certification-legislation>
- Podgursky, M., Monroe, R., and Watson, D. (2004). The academic quality of public school teachers: an analysis of entry and exit behavior. *Economics of Education Review*, 23 (2004), 507–518.
- Reardon, S.F., and Galindo, C. (2009) The Hispanic-white Achievement Gap in Math and Reading in the Elementary Grades. *American Educational Research Journal*, 46(3), 853–891.
- Regional Education Laboratory Midwest (2021). Michigan Teachers who are not Teaching: Who are they, and what would motivate them to teach? Institute of Education Science: US Department of Education. Accessed at: https://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_2021076.pdf
- Sass, T. R. (2015). Licensure and Worker Quality: Comparison of Alternative Routes to Teaching. *Journal of Law & Economics*, 58(1), 1-36.
- Serow, R.C., and Forrest, K.D. (1994). Motives and Circumstances: Occupational-Change Experiences of Prospective Late-Entry Teachers. *Teaching & Teacher Education*, 10(5), 555-563.
- Shen, J. (1997). Has the Alternative Certification Policy Materialized Its Promise? A Comparison between Traditionally and Alternately Certified Teachers in Public Schools. *Educational Evaluation and Policy Analysis*, 19(3), 276-283.
- Stackhouse, S.A. (2017). Trends in Michigan Teacher Certification Initial Certificates Issued 1996- 2016 [White paper]. Michigan Department of Education. https://www.michigan.gov//media/Project/Websites/mde/educator_services/research/5year_certificate_trend_with_endorsement_code_appendix.pdf
- Todd, P. and Wolpin, K. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, 1(1), 91–136.

- Yin, J., & Partelow, L. (2020, December 7). An Overview of the Teacher Alternative Certification Sector Outside of Higher Education [Review of An Overview of the Teacher Alternative Certification Sector Outside of Higher Education]. Center for American Progress. <https://www.americanprogress.org/article/overview-teacher-alternative-certification-sector-outside-higher-education/>
- Zimmerman, R, Swider, B.W., and Arthur, J.B. (2020) Does Turnover Destination Matter? Differentiating Antecedents of Occupational Change Versus Organizational Change. *Journal of Vocational Behavior*, 121 (2020) 103470.
- Zhang, G. and Zeller, N. (2016). A Longitudinal Investigation of the Relationship between Teacher Preparation and Teacher Retention. *Teacher Education Quarterly*, Vol 43, No. 2, pp. 73-92.

APPENDIX

Figure A.1.1. The Age Distribution of Newly Certified Individuals

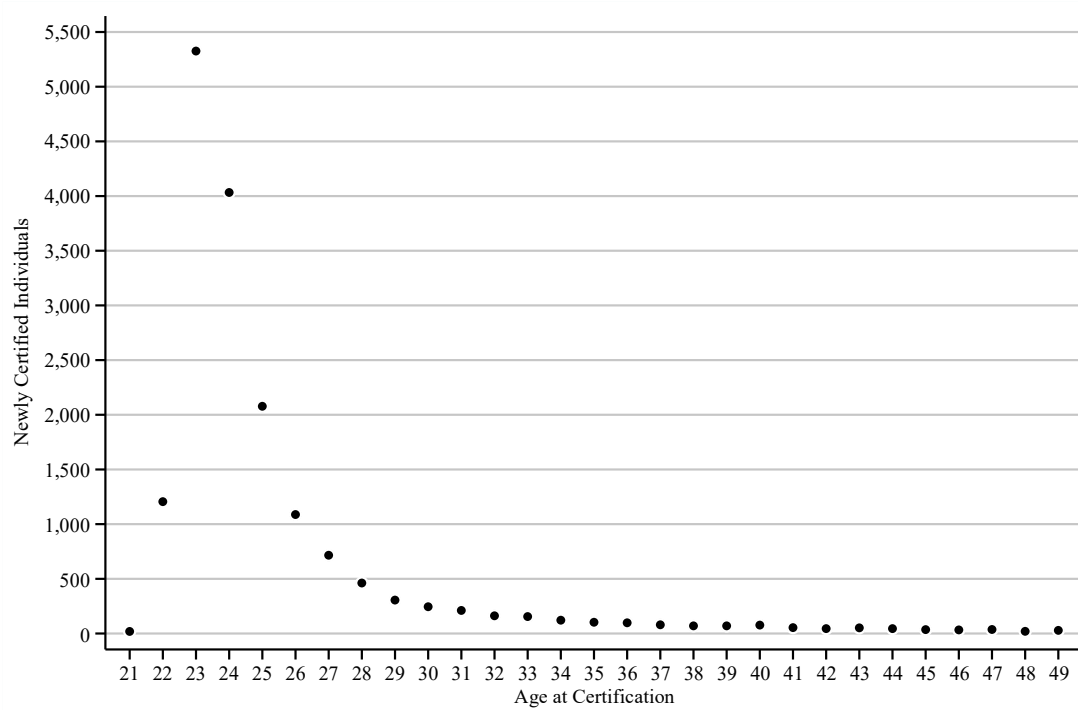


Table A.1.1. Linear Probability Model Estimates for Teacher Entry - Defining the 20s as 21-29

	Year 1	Year 2	Year 3
Older	-0.005 (0.002)	0.021*** (0.000)	0.049*** (0.000)
Black	0.187*** (0.002)	0.121*** (0.002)	0.083** (0.002)
Latino	0.015** (0.001)	-0.019** (0.000)	-0.036 (0.007)
Other Race	-0.060 (0.034)	-0.112* (0.014)	-0.123** (0.009)
Male	-0.017*** (0.000)	-0.038*** (0.001)	-0.039** (0.001)
Endorsement	0.107** (0.004)	0.082** (0.003)	0.061** (0.003)
30s & 40s x Black	-0.020* (0.002)	0.016* (0.002)	-0.002 (0.002)
30s & 40s x Male	0.051*** (0.000)	0.051*** (0.001)	0.029** (0.001)
MTTC Controls	X	X	X
Cohort Controls	X	X	X
Observations	16,981	16,981	16,981
R-squared	0.034	0.026	0.022

Note: Models include all certified teacher trainees in defined age group who were eligible to first enter teaching from 2012-13 through 2017-18. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, math, or science. *** p<0.01, ** p<0.05, * p<0.1

Table A.1.2. Cox Proportional Hazard Estimates for Teacher Entry - Defining the 20s as 21-29

	(1)	(2)	(3)
Older	0.089*** (0.030)	0.0918*** (0.030)	0.105*** (0.038)
Black			0.176*** (0.062)
Latino			-0.073 (0.075)
Other Race			-0.277*** (0.052)
Male			-0.088*** (0.024)
Endorsement			0.124*** (0.022)
30s & 40s x Black			-0.007 (0.120)
30s & 40s x Male			0.017 (0.065)
MTTC Controls			X
Cohort Controls		X	X
Observations	62,609	62,609	62,609

Notes: Hazard is defined as initial entry into teaching in a Michigan public school teacher. Cohorts able to initially enter from 2012-13 through 2017-18 are tracked through 2021-22. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, or STEM. *** p<0.01, ** p<0.05, * p<0.1

**Table A.1.3. Linear Probability Model Estimates for Exit after Employment as a Teacher -
Defining the 20s as 21-29**

	Year 1	Year 2	Year 3
30s and 40s	-0.001 (0.000)	-0.009* (0.001)	-0.026* (0.002)
Black	0.010* (0.001)	0.010* (0.001)	0.077** (0.002)
Latino	-0.024*** (0.000)	-0.007 (0.008)	-0.002 (0.011)
Other Race	0.064 (0.014)	0.092 (0.034)	0.101 (0.040)
Male	-0.001 (0.000)	0.010** (0.000)	0.021** (0.000)
Endorsement	-0.021 (0.007)	-0.026 (0.010)	-0.025 (0.014)
AgexBlack	-0.061*** (0.001)	-0.048** (0.002)	-0.061** (0.002)
AgexMale	0.029** (0.001)	0.035** (0.001)	0.046** (0.001)
MTTC Controls	X	X	X
Cohort Controls	X	X	X
Observations	10,898	10,898	10,898
R-squared	0.005	0.007	0.009

Note: Models include all certified teacher trainees in defined age group who were eligible to first enter teaching from 2012-13 through 2017-18. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, math, or science. *** p<0.01, ** p<0.05, * p<0.1

Table A.1.4. Cox Proportional Hazard Estimates for Exit after Teacher Employment -
Defining the 20s as 21-29

	(1)	(2)	(3)
30s and 40s	-0.026 (0.064)	-0.026 (0.064)	-0.144* (0.087)
Black			0.328*** (0.102)
Latino			-0.034 (0.160)
Other Race			0.364*** (0.096)
Male			0.013 (0.050)
Endorsement			-0.130*** (0.045)
30s & 40s x Black			-0.168 (0.218)
30s & 40s x Male			0.173 (0.136)
MTTC Controls			X
Cohort Controls		X	X
Observations	38,550	38,550	38,550

Notes: Hazard is defined as exit from teacher in a Michigan public school teacher. Cohorts able to initially enter from 2012-13 through 2017-18 are tracked through 2021-22. White is the reference category for race while the “other” category includes Asian, Native American, Native Hawaiian or Pacific Islander, and two or more races. We do not report these separately due to low Ns. Endorsement denotes a teaching certificate with a subject area endorsement in either English as a second language, special education, or STEM. *** p<0.01, ** p<0.05, * p<0.1

CHAPTER 2: DIMINISHING RETURNS ACROSS THE DAY: EVIDENCE FROM SCHOOL SCHEDULES

Introduction

The relationship between schooling and increased earnings is well documented in the literature (Mincer, 1984; Schultz, 1988; Becker, 2009; Goldin and Katz, 2010). Previous research demonstrates that student achievement is lower when they are cognitively fatigued or when they are at certain points in their circadian rhythm when their bodies' melatonin production is higher, such as the early morning (Persson et al, 2007; Schmidt et al, 2007). The aim of this paper is to map the associate between a student's daily schedule and her academic achievement. Specifically, I seek to examine and disentangle the effects of two facets of a student's daily schedule: (1) the start time of their school day and (2) the cognitive load she has experienced before the course is held. Through isolating the effect of these two elements of a student's daily schedule on her achievement, I can provide policymakers and school administrators recommendations on how school schedules can be realigned both at the system level as well as at the student level to improve the achievement of all students.

I overcome typical identification challenges by using panel data from North Carolina public high schools whose institutional characteristics allow for a causal assessment of schedule on achievement. I exploit the fact that within North Carolina public schools, students register for courses prior to those courses being assigned to instructors or time slots. Additionally, a computer program assigns teacher-course pairings to time slots and students to those teacher-course-time pairings in order to minimize scheduling conflicts.

I find that having an English or math class in the last block of the day decreases a student's GPA by 0.062 (0.007) and 0.064 (0.007), respectively. I further find having a math

during the first class of the day increases math EOC scores by as much as improving teacher quality by one fourth of a standard deviation while having math during the last class of the day decreases math test scores by as much as decreasing teacher quality by one half of a standard deviation (Rockoff, 2004). The first class of the day effect disappears if that class occurs before 7:30am, likely due to the effects of circadian rhythm on adolescents' early morning alertness. As the last class penalty is most pronounced for math classes, school administrators could see test score gains by shifting math courses away from block 4 as much as possible within their given staffing constraints.

There are several mechanisms that could drive this block 4 decline in productivity. This could be the result of student or instructor fatigue. Though I note previous research examines instructor schedules and finds that they suffer from minimal fatigue and gain "practice efficiency" if teaching multiple sections of the same course, therefore student fatigue is the more likely driver of these results (Williams and Shapiro, 2018). It is also possible that this effect is driven by differential attendance whereby students are disproportionately absent during the last block of the day. However, if students are disproportionately absent during the first half of the day, then attendance may be muting the result I find. These results may also in part be a function of student and instructor circadian rhythm. Though as circadian rhythm suggests instructors (adults) are most alert during mid-morning and students (adolescents) are most alert in the afternoon, circadian rhythm alone cannot be the only driver of the block 4 decline in productivity (Crowley, Acebo, & Carskadon, 2007; Wolfson and Carskadon, 1998). At this phase, my work is unable to clearly define the mechanism behind this last class penalty. These results do, however, suggest that the interplay between start time and the fatigue generated by a student's

particular course choices are critical to understanding the efficiency gains possible by reorganizing school schedules. This may prove a fruitful avenue of enquiry for future work.

Literature Review

This paper builds on two distinct streams in the literature: workplace productivity and school day schedules. The workplace productivity literature contains two different streams of interest, the first compares the productivity of labor across different shift times, most often comparing day shifts to night shifts, often across industries. Levin, Oler, and Whiteside (1985) find decreased productivity and an increased rate of accidents on the night shift. Taking evidence from manufacturing facilities in which employees work on a rotating shift, they compare the same group of employees at different shift times. The majority of papers in this stream use data from manual labor sectors from the early 20th century, concurring with Levin, Oler, and Whiteside (1985) that day shifts are more productive than night shifts and that more accidents occur during night shifts, particularly in the latter hours of long shifts. (Folkard and Tucker, 2003; Keller, 2009; Levin). Expanding the literature beyond factory work and manual labor, multiple studies find that the performance of medical residents is negatively affected by sleep deprivation (Veasey et al., 2002, Philibert, 2005; and Weinger and Ancoli-Israel 2002). It is yet unknown within the literature whether these results hold in occupations that predominantly consist of cognitive tasks.

The second stream in the workplace productivity literature examines the relationship between working hours and productivity, comparing different shift lengths. Pencavel (2015) examines the work hours and output of female munitions workers during World War I. He finds constant returns to hours worked when the shift is less than one hour and decreasing returns to hours worked above the one-hour threshold. Brachet et al. (2012) similarly find decreasing

returns to hours worked when analyzing short and long paramedic shifts within a difference-in-differences framework. These studies are unable to determine whether the decline in productivity across a shift is a function of fatigue or time of day.

Collewet and Sauermann (2017) attempt to unite these two streams, noting that time of day likely biases the results of the working hours literature. Using employee level data from a call center with variation in both the time a shift occurs as well as its length due to centralized scheduling based on consumer demand, they suggest a model that includes both time of day and working hours, thereby separating out the two effects. The authors conclude that as the number of hours worked increases, the average handling time for a call increases — which is a decline in productivity. The authors particularly note these results hold even among part-time workers who work comparatively short shifts.

The structure of the school day schedule literature mirrors that of workplace productivity, with two different streams, separately addressing timing and length. The school start time literature addresses when the school day occurs and finds that late morning start times improve student performance over early morning start times. Dills and Hernandez- Julian (2008) find students perform worse in earlier classes, taking evidence from administrative data from Clemson University. Self-selection into courses at the collegiate level as well as the subjectivity of course grades are more suggestive of a correlation relationship. Wahistrom (2002) avoids the self-selection issue present in collegiate schedules by examining high school data. She finds little to no effect on achievement as measured by grade point average (GPA) when school start time is moved from 7:15 am to 8:40 am. Carrell, Maghakian, and West (2011) establishes a causal relationship, absent in the previous literature, between school start time and student success outcomes. They exploit two consecutive, multidirectional school start time changes at the US Air

Force Academy. The authors utilize the randomized placement of students into courses and other unique structural aspects of the academy, like mandatory attendance and standardized grading across instructors, to tease out a positive effect of starting the school day later, roughly equivalent to raising teacher quality by one standard deviation. While the school start time literature may imply that students learn best in the afternoon, this has not been empirically tested. This literature speaks to the effect of school start time on *average* learning throughout the day but cannot speak to how learning might vary across the day.

The second strand in the school schedule literature addresses the length of a school day and gives insight into how learning might vary across the school day. Hobbs (2012) finds no increase in standardized math test scores for at-risk elementary and middle school students in north Georgia who participated in 21st Century Community Learning Centers. These after school programs provide tutoring and other curricular enrichment. Both Anderson and Walker (2015) and Morton (2020) examine longer school days within a difference-in-differences framework, assessing the performance of students in rural elementary schools who have adopted a four-day school week. They find no notable decline in standardized scores, though they notably use percent proficient on a state standardized test as their outcome, an imprecise measure which may be incapable of detecting small changes. Thompson (2020) reexamines the effect of the four-day school week on learning outcomes, using student level test score data in Oregon. He finds that math test scores decrease by between 0.037 and 0.059 standard deviations while reading scores decrease by between 0.033 and 0.042 standard deviations following the switch to the four-day school week, though he is unable to untangle the effect decreased hours of schooling each week and of fatigue generated by longer school days. On a larger scale, Patall, Cooper, and Allen (2010) look at 15 empirical studies of extended school time, either in the form of longer days or

more school days, as a measure to improve academic achievement. The studies have mixed results, and the authors are unable to determine a causal relationship between increased time in school and improved achievement. Pope (2016) aims to extend this literature by addressing how student achievement varies across a school day using student-level panel data from Los Angeles Unified School District, totaling 1.8 million student-year observations from 2003-2009 of sixth through eleventh grade students. His results show greater achievement outcomes for students' first classes of the day than for their last. He finds that having math in the first two blocks of the day, as opposed to the last two, increases the math GPA by 0.072 and increases the math CST score by 0.021 standard deviations. Having English in the morning increases the English GPA by 0.032 but does not significantly increase the English CST. As all schools in Pope's analysis have the same start-time, he is unable to determine whether this result is primarily driven by fatigue, time of day, or another mechanism entirely. As I use a data set that includes school start time variation, my work is better able to differentiate between fatigue and time of day as the most likely mechanism behind the positive effect of having a morning class.

Williams and Shapiro (2018) is the first to include time of instruction and fatigue in a unified model within the school day literature, using randomized student schedules and achievement data from the United States Air Force Academy (USAFA). Observing 6,981 students from 2004 to 2008, they find that students perform better in the afternoon and that the grade of two students in the same class may differ as much as 0.15 standard deviations owing only to their previous schedule that day. Additionally, they take into account the schedule of the instructor, finding that instructor fatigue is minimal and that instructors who teach the same course multiple times improve with each repetition, what Vernon (1921) terms "practice-efficiency" within the work hours productivity literature. While certain aspects of the USAFA

environment are ideal for determining causality, it is a demographically homogenous group of students at a highly specialized, elite university. It is not readily apparent that these results can be extrapolated to secondary education. I add to this literature by examining productivity of teachers and students in the secondary school environment.

Data

This study relies on student-level data from public high schools in North Carolina which were generously provided by the North Carolina Department of Public Instruction via the North Carolina Education Research Data Center (NCERDC). North Carolina has an annual student enrollment of over 1.6 million and contains over 600 public high schools across 115 different local educational authorities (LEAs). The data contains demographic information for students including English Language Learner (ELL) status, gender, grade level, teacher, course, semester, and course block. I summarize student characteristics in Table 2.1. Additionally, I utilize end-of-course (EOC) and end-of-grade (EOG) standardized test scores as well as individual course GPA as the measure of achievement outcome.

EOCs and EOGs are administered at the end of each term, across the state to all students in third through twelfth grade, within a five-day window, and at a time of day determined individually by each school, though consistent within a school. I normalize both the EOG and EOC scores, reporting the effects in terms of standard deviation units. Students receive a grade of A, B, C, D, or F for each course and I record these on an unweighted scale from 0 to 4 such that A=4, B=3, etc.

These data come from high schools that operate on the “block” schedule. Each student takes four classes in the fall semester and four different classes in the spring semester. The school day start time varies by LEA but ranges from 7:00 to 9:00am. This heterogeneity in start

time allows me to determine whether circadian rhythm may mute or amplify the results. I sort the school start times into four different categories: before 7:30am, 7:30 – 7:59am, 8:00 – 8:29am, and 8:30am and later. The day is made up of four blocks of approximately 90 minutes each. Third block is extended to two hours to incorporate a 25-minute lunch.

The scheduling process of students into courses for North Carolina public high schools aids in the causal identification of my research question. Students do not select courses from an existing schedule of times and instructors as is the case at most universities, nor are they assigned to a class and a schedule by an administrator as is often the case at the elementary and middle school levels. Students sign up for eight courses based upon the courses offered by a school. These courses are not yet assigned to instructors or times. Based upon the number of interested students for each course and staffing constraints, administrators determine the number of sections of each course to be offered and assign instructors to those sections. Times are not yet associated with the course and teacher pairings. Staff enter student selections, course offerings, and instructor assignments into scheduling software which then generates a master schedule matrix such that scheduling conflicts are minimized. Student scheduling conflicts are then resolved individually by school administrators and guidance counselors. Student schedule change requests are possible, though through discussions with guidance counselors tasked with managing student schedules from across the state, I note that schools have policies that limit students' ability to request schedule changes and requests for taking a particular course at a different time are exceedingly rare. To maintain this scheduling environment, I exclude charter schools, early colleges, and specialized high schools whose schedules and scheduling process differ from traditional public high schools.

In Figure 2.1, I display the normalized mean math and mean English score by block by start time. I set first block equal to zero and display the difference in mean for each of the other blocks for each of the four start time categories. There appears to be a noticeable decline in performance between third and fourth block for math classes while there appears to be a notable improvement after first block for English classes.

Methodology

While students do not directly select class times, the assignment of students to classes may not be as good as random for a variety of reasons. Schools face staffing constraints. they may for instance, employ a part time calculus teacher in the morning which necessarily means the classes they teach must fall into the first two blocks of the day. Students may notice this pattern and make course choices accordingly. I therefore employ a fixed effects framework to mitigate this issue.

Owing to the richness of the data and the random assignment of class times, I proceed with the following simple model:

$$A_{i,t} = \alpha + \beta Block_{i,0} + \delta D_{i,t} + \eta A_{i,t-1} + \theta L_{i,t} + \gamma TC_{i,t} + \rho_t + \varepsilon_{i,t}, \quad (1)$$

where $A_{i,t}$ is the achievement outcome of interest and is defined by first student i 's standardized exam score and then class GPA in year t while $A_{i,t-1}$ is defined by a student's previous achievement scores. $Block_{i,0}$ is a vector of binary variables that takes on the value of 1 if student i 's class is in a particular block and 0 otherwise. $Block_{i,0}$ serves as the treatment variable. Its coefficients compare the outcome of interest, exam score, of students with a class in second, third, or fourth block respectively versus those students enrolled in a first block class, which acts as the reference category. The vectors L_i , TC_i , and ρ_t respectively allow for grade level, teacher by course, and year fixed effects such that I am comparing within-group variation of students of

the *same* grade level who take the *same* course-teacher combination with the *same* year but in a *different* class block of the day. The vector D_i contains basic demographic controls such as gender and ELL status and parental education level. The random error term is ε_i .

Results

In this section, I begin by estimating the effect of class block on math and English EOC scores using equation (1). I then partition the results by school start time, noting how the effect of $Block_{i,0}$ might vary with what time of day that block occurs. I next proceed by examining the results by student subgroup, looking for heterogeneous effects and possible efficiency gains. Finally, I look at GPA as the outcome of interest, assessing how productivity may vary across school subjects.

Main Results

Table 2.2 displays the estimates of equation (1) for Math EOC scores in columns (1) to (4) and English EOC scores in columns (5) to (8). Columns (1) and (5) show the difference in the mean math and English EOC scores between students with math or English in block 1 and those in block 2, block 3, or block 4. When prior test scores are taken into account in columns (2) and (6), the estimates for each block for math become larger suggesting there is some positive selection of higher performing students into later block math classes. The estimates for each block for English suggest some selection of higher performing students into earlier blocks. This may be a function of how some advanced courses are scheduled. With the full set of controls in columns (4) and (8), the estimates of having math class during block 2, 3, and 4 instead of first block decreases a student's math score by 0.009(0.002), 0.002 (0.002), and 0.043(0.002) respectively. The estimates of having English class during block 2, 3, and 4 instead of first block decreases a student's English score by 0.015(0.002), -0.004 (0.002), and 0.014(0.002)

respectively. The largest effect size appears to be for having math during block 4, -0.043. Not having math during block 4, thereby increasing one's EOC score by 0.043 standard deviations which is equivalent to increasing teacher quality by one half of a standard deviation (Rockoff, 2004); or put another way, closing the gender gap (Hyde et al., 2008). I display the results of columns (4) and (8) in Figure 2.2. While the trend for English classes is less clear, and in some cases not significant, for math scores there appears to be a slight bonus for having math during block 1 and a large penalty for having math during block 4. The block 3 measures should be interpreted with caution as block 3 contains a lunch block, though when that lunch block occurs varies by classroom. Some students in the data set may experience a lunch break before block 3, while others may experience a lunch break at the end of block 3. Focusing on ease of interpretation in the next sections, I set blocks 2 and 3 as the reference category and focus on the first class bonus and last class penalty.

By Start Time

A particular block captures two things, a time of day and how many classes a student has already taken (fatigue). As school start times vary across districts, and therefore the timing of each block varies, the effect of each block might vary with that start time. While in the main model there appears to be a small bonus for having math during first block, I offer suggestive evidence this effect may not persist with less favorable start times. Widely known in the school start time literature, circadian rhythm specifies that there are certain times of day in which a person is most alert. For adolescents, melatonin levels (causing sleepiness) approach their peak at 7:00 am, stopping production at 8:00 am, with greater alertness beginning in the late morning (Carrell, Maghakian, and West; 2011). Melatonin then increases again in the late afternoon, once again causing sleepiness. In Table 2.3, I now partition the math score results by school start time,

setting blocks 2 and 3 as the reference category to demonstrate the particular effect start time has on the first class of the day bonus and last class of the day penalty. While the last class penalty is evident regardless of start time, an early start time appears to erode the first class bonus. I then proceed to test whether these differences in start time are statistically significant and present the results in Table 2.4.B. I do not find the difference in the first class of the day “bonus” between schools that start before 7:30am and those that start after. Due to the relatively small number of districts that have these earliest start times, my model is unable to precisely estimate the effect of having a first block class for these districts. I do however note that the last class of the day effect is present, regardless of the start time, suggesting that this is more likely driven by the fatigue generated by attending classes than unique to the time of day.

Subgroups

I now test whether the last class of the day effect differs by student characteristics. I first examine the results for females versus males. I perform the analysis for each of the two subsamples. I report the results in panel A of Table 2.4. The results are not significant for either gender for the English EOC. Males appear to suffer from a larger last class of the day penalty in mathematics than their female counterparts.

I also split the sample by student exceptionality in panel A of Table 2.4. Students who receive services for a documented learning disability is a broad category which includes any student whom the state deems in need of an individual education plan (IEP) owing to a reading, writing, or math learning disability. It can include students with autism, dyslexia, a hearing impairment, etc. I do not find a significant last class penalty for English for either group. I do, however, find the last class penalty persists for math, though the effect size is approximately the same for both groups.

I further split the sample into terciles by eighth grade test scores. I do not observe statistically significant differences for English. However, for mathematics, I test whether the coefficients are equal across all three terciles and find that they are not. Notably, the negative effects of having math in the last block of the school day is higher for the middle tercile of students than for both low and high achieving students. I find it implausible that this non-linear relationship is solely a function of their prior achievement. It may be picking up the interaction between their ability and the rigor of their course selections, which I am unable to observe. I investigate this further with the available data by also splitting the sample by students designated as academically and intellectually gifted (AIG) as well as general education students and report the results in panel A of Table 2.4. I once again find very little of note for English, while I find the last class penalty to be significant for math for both groups and significantly larger for those designated AIG. There may be a variety of mechanisms involved here. This difference in the size of the last class penalty for math may be picking up that students designated as AIG may select into more rigorous courses that induce more fatigue than the schedules of the general student population.

Finally, I examine the last block effect by classification as an English language learner. While I notice little difference between the last class effect on math for these students, there is a pronounced difference for English. Students who are designated as English language learners experience a stronger negative effect of having English class in the last block of the day.

Grade Point Average

I now turn to the effect of block on GPA for math, English, social studies, and science classes and present the results in Table 2.5. I make one adjustment to the previous model. As GPA is a more subjective measure, it may be the case that teachers match the distribution of

grades in a particular term to their historic distribution of grade assignments, comparing students to their current classmates rather than to an objective standard. I attempt to account for this by examining the difference between scores by block of the students of teachers who teach the same course twice in the *same term*. I therefore transition from a teacher by course fixed effect to a teacher by course by term fixed effect. An important assumption here is that to the extent that teachers assign a consistent distribution of grades, they would likely do this within a course-term, not within a course-class block combination, such that I can attribute the difference in scores between class blocks of students who take the same course from the same teacher in the same term to time of day or fatigue. I do not find evidence of a “first class” bonus for GPA across all subject areas. There appears to be a small, though perhaps not significant bonus for social studies and science whereas math and English experience a small but insignificant “first class” penalty. I do find evidence of the last class penalty across all subjects. Examining the coefficients on block 4 in row 3 of table 5, the last class penalty is largest for math and English, though still both practically and statistically significant for science and social studies.

Balance Checks

I also test whether the assignment to different blocks of different teacher and student characteristics has the equivalent balance of what we would expect under randomization, which could cast doubt on the randomization around what block a student is assigned to for a given class. I thus estimate regressions such as (1) with different student, teacher, and course characteristics as the dependent variables:

$$X_{i,t} = \alpha + \beta Block_{i,0} + \theta L_{i,t} + \gamma TC_{i,t} + \rho_t + \varepsilon_{i,t}, \quad (2)$$

In particular, I check for balance in gender, 8th grade standardized test scores, teacher years of experience, prior year test scores, and race. I report the results in Table 2.6. Many, though not

all, of the observed covariates appear to be fairly well-balanced across the class blocks. Notably, teacher experience, 8th grade math, prior test score, and gender are significant at or above the 5% level. Performing 21 hypotheses tests in Table 2.6, the probability of erroneously rejecting a true null hypothesis in 6 out of cases is less than 1%. For teacher experience, classes during block 2 have on average a teacher with one one-hundredth of a year more teaching experience. Females are approximately one half of a percentage point less likely to have a math or English class during block 4. Students in block 2 have on average one one-hundredth of a standard deviation higher prior year test scores.

Conclusion

This paper finds that productivity declines in the last class of the day regardless of when that class occurs, suggesting this result is more likely driven by fatigue than any effect circadian rhythm has on alertness. I find that having an English or math class in the last block of the day decreases a student's GPA by 0.062 (0.007) and 0.064 (0.007), respectively. I further find having a math during the first class of the day increases math EOC scores by as much as improving teacher quality by one fourth of a standard deviation while having math during the last class of the day decreases math test scores by as much as decreasing teacher quality by one half of a standard deviation (Rockoff, 2004). The first class of the day effect disappears if that class occurs before 7:30am, likely due to the effects of circadian rhythm on adolescents' early morning alertness. I find larger test score effects in mathematics than Pope (2016), though I am unable to determine if this is a result of the school start time variation present in my sample of the difference in fatigue generated by "block" schedule as opposed to the traditional six block day, or some other unknown mechanism.

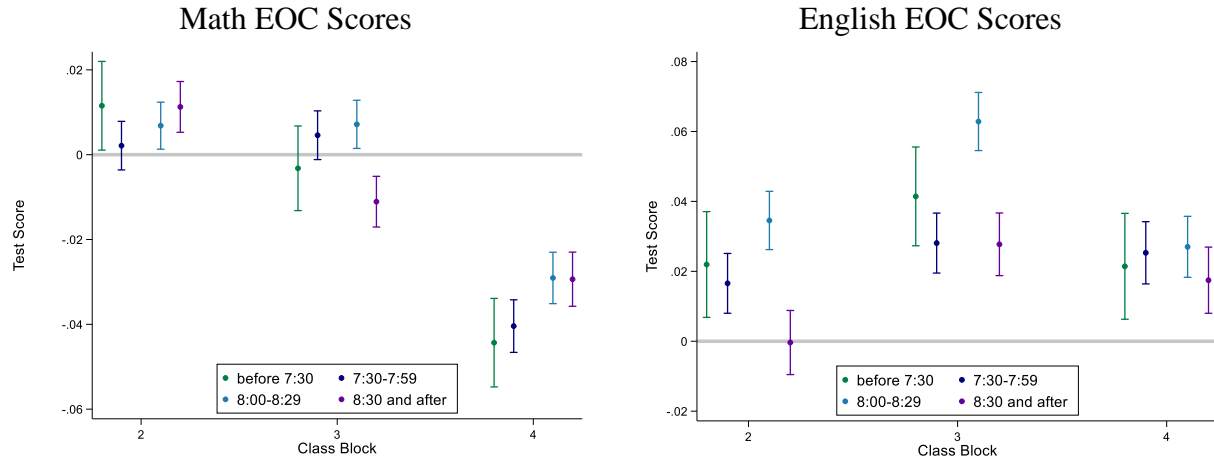
As the effect of block varies by subject area as well as student characteristics, administrators can see efficiency gains by exploiting this heterogeneity. They might do this in several different ways. First, they might place courses deemed more important earlier in the day, placing core courses such as English and math earlier in the day than electives like physical education or art. They may also place courses for which the last class penalty is larger at the beginning of the day, such as mathematics. Additionally or alternately, they might prioritize students who may experience the greatest last class penalty, such as males or those designated AIG, for early in the day math classes. School administrators could see math test score equivalent to increasing teacher quality by one half a standard deviation for one fourth of students simply by assigning all math teachers to block 4 planning blocks, thereby shifting the classes they teach away from block 4.

These results may have policy implications for rural schools in particular. As rural school districts across the United States face outmigration and declining state funding, the four-day school week offers a way to cut transportation and overhead cost while decreasing teacher and student absenteeism (Anderson and Walker, 2015). States mandate the number of hours per year a student must be in school, therefore decreasing the number of days a student attends school necessitates increasing the number of hours a student spends in school each of the remaining days. As this study shows, however, not all school hours are equally productive. Mapping the time-of-day and fatigue effects students experience can help inform policymakers both about the potential effect of extended school days as well as how to mitigate any potential ill effects. There are several alternate hypotheses that might explain the trends I find for the first and last classes of the day. The first is that the losses observed in the last block of the day, as well as the gains seen in the first block of the day, may be due to “mean-reversion.” With mean-reverting

measurement error, gains by students who scored unusually low in the previous year are not likely to be normally distributed around the initial score, but rather, those low scoring students in the previous year are likely to experience larger than average gains the following year. I cannot rule out that this is a driver of the results I find. A second hypothesis could be student or instructor fatigue. Though I note previous research examines instructor schedules and finds that they suffer from minimal fatigue and gain “practice efficiency” if teaching multiple sections of the same course, therefore student fatigue is the more likely driver of these results (Williams and Shapiro, 2018). A third hypothesis is that this effect is driven by differential attendance whereby students are disproportionately absent during the last block of the day. However, if students are disproportionately absent during the first half of the day, then attendance may be muting the result I find. These results may also in part be a function of student and instructor circadian rhythm. Though as circadian rhythm suggests instructors (adults) are most alert during mid-morning and students (adolescents) are most alert in the afternoon, circadian rhythm alone cannot be the only driver of the block 4 decline in productivity. At this phase, my work is unable to clearly define the mechanism behind this “first class” bonus and last class penalty. These results do, however, suggest that the interplay between start time and the fatigue generated by a student’s particular course choices are critical to understanding the efficiency gains possible by reorganizing school schedules. I seek to answer this question in full in the next phase of my work.

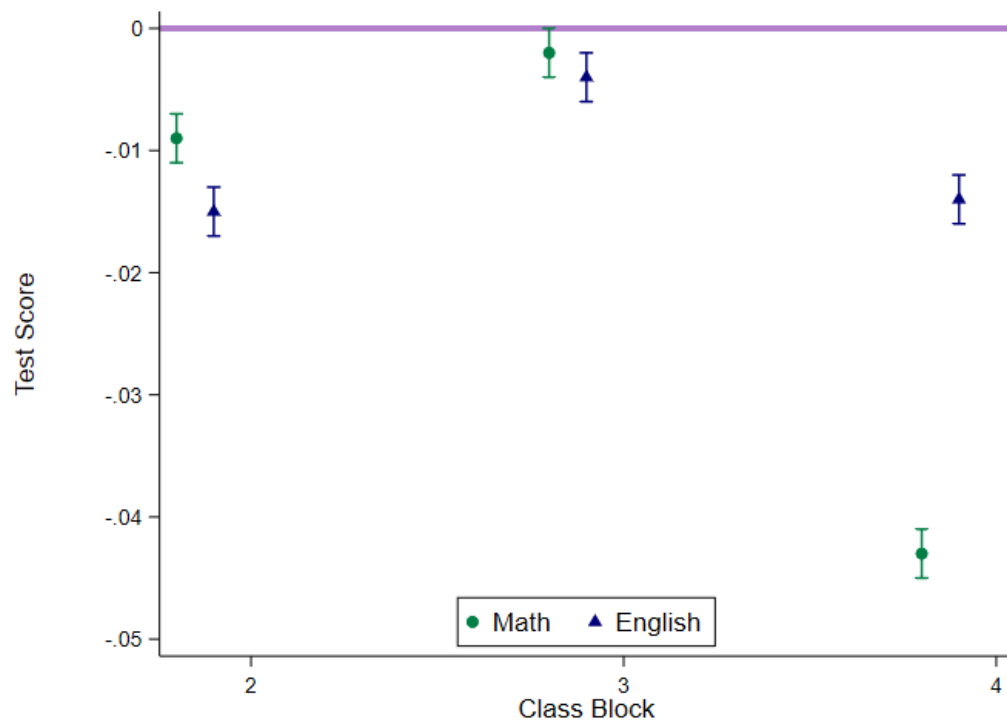
TABLES AND FIGURES

Figure 2.1. Descriptive Results



Normalized test score by 90-minute block, as compared to the mean of block 1. Note: Figure 2.1 shows the difference in the mean test score for each of four sequential 90-minute class blocks from block 1 along with a 95 percent confidence interval for each of the four start time categories.

Figure 2.2. Estimation Results



This displays the estimates of the effect of which block a class is taken on the outcome variables. All estimates use equation (1). Block 2, Block 3, and Block 4 are binary variables equal to 1 if the individual's Math/English class takes place during that Block and 0 otherwise. Block 1 is the reference category. Standard errors clustered at the school level.

Table 2.1. Summary Statistics

Variable	Mean	SD
Female	0.52	0.50
English Language Learner	0.11	0.41
Academically Gifted	0.08	0.28
Asian	0.02	0.14
Black	0.26	0.44
Multi-racial	0.02	0.14
Hispanic	0.05	0.22
Native American	0.01	0.11
White	0.62	0.48
Parent Education: Less than HS	0.03	0.17
HS Grad	0.12	0.32
Some College	0.07	0.26
Comm College	0.12	0.32
4yr College	0.12	0.33
College Plus	0.05	0.21
No Educ Response	0.49	0.50
Number of Individuals	578,805	

Table 2.2. Later Blocks versus First Block Classes

Variables	Math EOC Score				English EOC Score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Block 2	0.00245	-0.012***	-0.011***	-0.009***	0.021**	- 0.012***	-0.011***	-0.015***
	(0.006)	(0.004)	(0.002)	(0.002)	(0.010)	(0.003)	(0.002)	(0.002)
Block 3	0.001	-0.006	-0.005**	-0.002	0.044***	0.002	-0.001	-0.004*
	(0.007)	(0.004)	(0.002)	(0.002)	(0.011)	(0.003)	(0.002)	(0.002)
Block 4	-0.034***	-0.056***	-0.052***	-0.043***	0.026**	- 0.014***	-0.014***	-0.014***
	(0.007)	(0.004)	(0.003)	(0.002)	(0.012)	(0.004)	(0.002)	(0.002)
Prior math score		0.669***	0.627***	0.621***		0.278***	0.274***	0.260***
		(0.005)	(0.001)	(0.001)		(0.002)	(0.001)	(0.001)
Prior English score		0.142***	0.130***	0.126***		0.578***	0.568***	0.545***
		(0.002)	(0.001)	(0.001)		(0.002)	(0.001)	(0.001)
Less than HS			0.136***	0.103***			0.0301*	0.0367**
			(0.023)	(0.014)			(0.017)	(0.015)
HS Grad			0.132***	0.103***			0.0755***	0.0758***
			(0.023)	(0.014)			(0.017)	(0.015)
Some College			0.139***	0.108***			0.105***	0.0979***
			(0.023)	(0.014)			(0.018)	(0.016)
Comm College			0.151***	0.118***			0.151***	0.140***
			(0.022)	(0.014)			(0.017)	(0.015)
College Grad			0.192***	0.147***			0.173***	0.153***
			(0.022)	(0.014)			(0.017)	(0.015)
College Plus			0.238***	0.180***			0.203***	0.173***
			(0.023)	(0.010)			(0.017)	(0.015)
ELL			-0.073***	-0.070***			0.0523***	0.0498***
			(0.012)	(0.009)			(0.009)	(0.010)
Grade FE			X	X			X	X
Year FE			X	X			X	X
TeacherxCourse FE				X				X
Observations	1,195,384	1,195,384	1,195,384	1,195,384	578,805	578,805	578,805	578,805
R-squared	0.000	0.531	0.536	0.483	0.000	0.679	0.683	0.604

All columns use equation (1) for Math EOC and English EOC scores, respectively. Block 2, Block 3, and Block 4 are binary variables equal to 1 if the individual's Math/English class takes place during that Block and 0 otherwise. Block 1 is the reference category. The excluded parental education binary variable is no response. Standard errors clustered at the school level are in brackets. Statistical significance is shown by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2.3. By School Start Time

Variables	Math EOC Scores				English EOC Scores			
	before 7:30	7:30-7:59	8:00-8:29	8:30&after	before 7:30	7:30-7:59	8:00-8:29	8:30&after
First class	0.001 (0.007)	0.012*** (0.004)	0.015*** (0.004)	0.013*** (0.004)	0.000 (0.007)	0.015*** (0.004)	0.005 (0.003)	0.011*** (0.004)
Last class	-0.042*** (0.005)	-0.035*** (0.004)	-0.037*** (0.004)	-0.038*** (0.004)	0.003 (0.007)	-0.002 (0.004)	-0.009** (0.004)	-0.005 (0.005)
Prior math score	0.536*** (0.005)	0.541*** (0.003)	0.562*** (0.003)	0.541*** (0.003)	0.260*** (0.004)	0.260*** (0.002)	0.265*** (0.002)	0.254*** (0.002)
Prior English score	0.123*** (0.004)	0.121*** (0.002)	0.111*** (0.002)	0.114*** (0.002)	0.531*** (0.004)	0.547*** (0.002)	0.545*** (0.002)	0.548*** (0.002)
Less than HS	0.038 (0.070)	0.099*** (0.020)	0.151*** (0.031)	0.071*** (0.025)	-0.011 (0.062)	0.034 (0.022)	0.033 (0.032)	-0.011 (0.062)
HS Grad	0.045 (0.069)	0.099*** (0.020)	0.146*** (0.030)	0.077*** (0.025)	0.008 (0.061)	0.077*** (0.020)	0.075** (0.032)	0.008 (0.061)
Some College	0.070 (0.068)	0.113*** (0.020)	0.150*** (0.030)	0.087*** (0.024)	0.076 (0.063)	0.129*** (0.021)	0.125*** (0.032)	0.076 (0.063)
Comm College	0.063 (0.067)	0.120*** (0.020)	0.158*** (0.030)	0.088*** (0.025)	0.067 (0.061)	0.140*** (0.021)	0.139*** (0.032)	0.067 (0.061)
College Grad	0.116* (0.068)	0.149*** (0.020)	0.180*** (0.030)	0.121*** (0.024)	0.098* (0.059)	0.160*** (0.021)	0.143*** (0.032)	0.098* (0.059)
College Plus	0.136** (0.068)	0.182*** (0.020)	0.209*** (0.030)	0.159*** (0.025)	0.113* (0.060)	0.164*** (0.020)	0.175*** (0.033)	0.113* (0.060)
ELL	-0.055 (0.087)	-0.005 (0.029)	-0.084*** (0.029)	-0.027 (0.037)	0.087** (0.036)	0.060*** (0.017)	0.033** (0.015)	0.047* (0.024)
Grade FE	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X
TeacherxCourse FE	X	X	X	X	X	X	X	X
Observations	176,505	227,204	208,426	193,102	153,479	177,514	186,149	161,663
R-squared	0.498	0.501	0.497	0.479	0.591	0.617	0.593	0.600

All columns use equation (1) for Math EOC and English EOC scores, respectively. First class is a binary variable equal to 1 if a student has math/English class during Block 1 while Last class is a binary variable equal to 1 if a student has math/English class during Block 4. Blocks 2 and 3 serve as the reference category. The excluded parental education binary variable is no response. Standard errors, clustered at the school level, are in brackets. Statistical significance is shown by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2.4.A. Student Subgroups

Variable	Math EOC	English EOC
A. Gender		
Last for Females	-0.037*** (0.003)	-0.003 (0.003)
Last for Males	-0.056*** (0.003)	-0.004 (0.003)
Difference	0.019***	0.001
<i>P</i> -value	0.000	0.120
B. Exceptionality		
Last for General Education Students	-0.041*** (0.004)	-0.001 (0.003)
Last for Students who Receive Services for a Learning Disability	-0.040*** (0.007)	-0.003 (0.004)
Difference	0.001**	0.002
<i>P</i> -value	0.050	0.210
C. Academically and Intellectually Gifted		
Last General Education Students	-0.034*** (0.003)	-0.004 (0.002)
Last for Students Designated AIG	-0.046*** (0.008)	-0.012* (0.006)
Difference	0.012***	0.008*
<i>P</i> -value	0.000	0.070
D. English Learner Status		
Last English Learner Students	-0.044 *** (0.005)	-0.020*** (0.005)
Last Non-English Learners	-0.038*** (0.003)	-0.000 (0.003)
Difference	0.006	0.020***
<i>P</i> -value	0.749	0.000
E. Eight Grade Scores		
Last Lowest Tercile	-0.034*** (0.004)	0.003 (0.004)
Last Middle Tercile	-0.040*** (0.004)	-0.007 (0.004)
Last Highest Tercile	-0.033*** (0.004)	-.007 (0.004)
<i>P</i> -value	0.004**	0.299

Table 2.4.B. Start Time

Variable	Math EOC	English EOC
A. Last Class		
Last 7:00-7:29 Start Times	-0.042*** (0.009)	0.001 (0.008)
Last 7:30-9:00 Start Times	-0.034*** (0.003)	-0.005 (0.003)
Difference	0.008	0.006
<i>P</i> -value	0.693	0.108
B. First Class		
First 7:00-7:29 Start Times	0.002 (0.009)	0.003 (0.006)
First 7:30-9:00 Start Times	0.011*** (0.003)	0.010*** (0.003)
Difference	0.009	0.007
<i>P</i> -value	0.265	0.183

Table 2.5. GPA across Subjects

Variables	Math	English	Social studies	Science
Block 2	0.003 (0.004)	0.001 (0.004)	-0.007 (0.005)	0.002 (0.005)
Block 3	0.003 (0.005)	-0.006 (0.005)	-0.020*** (0.005)	-0.013** (0.005)
Block 4	-0.064*** (0.007)	-0.062*** (0.007)	-0.053*** (0.007)	-0.058*** (0.007)
Prior Math	0.631*** (0.006)	0.413*** (0.005)	0.372*** (0.004)	0.491*** (0.005)
Prior English	0.050*** (0.003)	0.197*** (0.003)	0.271*** (0.004)	0.202*** (0.004)
Honors level course	0.239*** (0.016)	0.156*** (0.011)	0.335*** (0.012)	0.240*** (0.012)
Female	0.298*** (0.005)	0.432*** (0.005)	0.201*** (0.005)	0.242*** (0.005)
Year FE	X	X	X	X
Grade FE	X	X	X	X
Teacher by course FE by term	X	X	X	X
Observations	556,413	525,632	454,416	438,626
R-squared	0.260	0.242	0.291	0.294

All columns use equation (1) for Math GPA and English GPA scores, respectively. Block 2, Block 3, and Block 4 are binary variables equal to 1 if the individual's Math/English class takes place during that Block and 0 otherwise. Block 1 is the reference category. Standard errors clustered at the school level are in brackets. Statistical significance is shown by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2.6. Balance in Covariates

	Block 2	Block 3	Block 4
8 th grade English score	0.006* (0.003)	0.003 (0.003)	0.006* (0.003)
8 th grade math score	0.007** (0.003)	0.006* (0.003)	0.007** (0.004)
Prior year score	0.011** (0.004)	0.003 (0.005)	0.000 (0.005)
Female	0.000 (0.001)	-0.004** (0.002)	-0.006*** (0.002)
Black	-0.002 (0.001)	0.000 (0.001)	-0.001 (0.001)
Hispanic	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)
Teacher experience	0.014** (0.005)	-0.008 (0.005)	0.002 (0.005)
Observations	1,774,189		

All columns use equation (2). Block 2, Block 3, and Block 4 are binary variables equal to 1 if the individual's Math/English class takes place during that Block and 0 otherwise. Block 1 is the reference category. Each coefficient within a column is from a different regression. dependent variable is indicated in each row header. Standard errors clustered at the school level are in brackets. Statistical significance is shown by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

BIBLIOGRAPHY

- Anderson, D. Mark, and Mary Beth Walker, (2015), “Does Shortening the School Week Impact Student Performance? Evidence from the Four-Day School Week,” *Education Finance and Policy*, 10 314–349.
- Becker, Gary S, (2009) Human capital: A theoretical and empirical analysis, with special reference to education, University of Chicago press.
- Brachet, T., David, G., Drechsler, A.M., (2012), “The effect of shift structure on performance,” *American Economic Journal: Applied Economics*, 4 219–246.
- Carrell, Scott E., and Mark Hoekstra, (2014), " Are school counselors an effective education input?" *Economics Letters*, 125 (2014) 66–69.
- Carrell, Scott E., Teny Maghakian, and James West, (2011), “A’s from Zzzz’s? The Causal Effect of School Start Time on the Academic Achievement of Adolescents,” *American Economic Journal: Economic Policy*, 3 62–81.
- Collewet, Marion, Jan Sauermann, (2017), “Working Hours and Productivity,” *Labour Economics*, 47, 96–106.
- Crowley, Stephanie, Christine Acebo, and Mary Carskadon, “Sleep, Circadian Rhythms, and Delayed Phase in Adolescents,” *Sleep Medicine* 8 (2007), 602–612.
- Dills, Angel K., and Ray Hernandez-Julian, (2008), “Course Scheduling and Academic Performance,” *Economics of Education Review*, 27, 646–654.
- Dolton, P., Howorth, C., Abouaziza, M., 2016, “The Optimal Length of the Working Day: Evidence from Hawthorne Experiments,” ESPE conference paper.
- Folkard, Simon, and Philip Tucker, (2003), “Shift Work, Safety and Productivity,” *Occupational Medicine*, 53, 95–101.
- Goldin, Claudia and Lawrence F Katz, (2010), *The race between education and technology*, Harvard University Press.
- Hobbs, Connie Lynn, (2012), "Effects of an Afterschool Program on Elementary and Middle School Math Achievement in Georgia Schools," *Doctoral Dissertations and Projects*. 509. <http://digitalcommons.liberty.edu/doctoral/509>.
- Hyde, Janet, Sarah Lindberg, Marcia Linn, Amy Ellis, and Caroline Williams, (2008), “Gender Similarities Characterize Math Performance,” *Science* 321, 494 – 495.
- Keller, Simone M, (2009), "Effects of Extended Work Shifts and Shift Work on Patient Safety, Productivity, and Employee Health," *American Association of Occupational Health Nurses*, 57, 497 - 502.

- Levin, Lester, Jacqueline Oler, and Jeffery R. Whiteside, (1985), "Injury Incidence Rates in a Paint Company on Rotating Production Shifts," *Accident Analysis and Prevention* 17, 67–73.
- Mincer, Jacob, "Human capital and economic growth," *Economics of education review*, 1984, 3 (3), 195–205.
- Morton, Emily, (2020), "Effects of Four-day School Weeks on School Finance and Achievement: Evidence from Oklahoma. *Educational Researcher*, Vol. 50, No. 1, 30–40
- Patall, Erika A., Harris Cooper and Ashley Batts Allen, (2010), " Extending the School Day or School Year: A Systematic Review of Research (1985-2009)," *Review of Educational Research*, Vol. 80, No. 3, 401-436.
- Pencavel, John, (2015), "The Productivity of Working Hours," *The Economic Journal*, 125, 2053–2076.
- Pencavel, John, (2016), "Recovery from Work and Productivity of Working Hours," *Economica*, 83, 545–563.
- Persson, J., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2007), "Cognitive fatigue of executive processes: Interaction between interference resolution tasks," *Neuropsychologia*, 45(7), 1571–1579.
- Pope, Nolan G., (2016), "How the Time of Day Affects Productivity: Evidence from School Schedules," *The Review of Economics and Statistics* 98, 1–11.
- Rockoff, Jonah, (2004), "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review, May Papers and Proceedings*, 94 (2004), 247 – 252.
- Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: Circadian rhythms in human cognition. *Cognitive Neuropsychology*, 24(7), 755–789.
- Schultz, T Paul, (1988), "Education investments and returns," *Handbook of development economics*, 1, 543–630.
- Thompson, Paul N. (2020). "Is Four less than Five? Effects of Four-day School Weeks on Student Achievement in Oregon," *Journal of Public Economics*, 192 (2020), 1 – 20.
- Vernon, H.M., (1921) *Industrial fatigue and efficiency*, Routledge & Sons, London.
- Wahistrom, Kyla, (2002), "Changing Times: Findings from the First Longitudinal Study of Later High School Start Times," *NASSP Bulletin*, 86(633), 3-21.
- Williams, Kevin M., Teny Maghakian Shapiro, (2018), "Academic Achievement across the Day: Evidence from Randomized Class Schedules," *Economics of Education Review*, 67, 158–170.

Wolfson, Amy, and Mary Carskadon, "Sleep Schedules and Daily Functioning in Adolescents,"
Child Development 69 (1998), 875–887.

CHAPTER 3: SPENDING & ACHIEVEMENT EFFECTS OF INCREASED FUNDING TO RURAL SCHOOL DISTRICTS: EVIDENCE FROM WISCONSIN

Introduction

Rural school districts in the United States face unique challenges relative to their urban and suburban counterparts, such as frequent staffing turnover, high transportation costs, and limited economies of scale (Sipple and Brent, 2015; Showalter et al., 2019). These characteristics may reduce how much rural districts can spend on specialized staff (e.g., social workers and guidance counselors) and curriculum (e.g., AP courses and career and technical education), potentially contributing to the well-documented disparities in educational outcomes between rural and non-rural students. For instance, rural students are four percentage points less likely to attend college and seven percentage points less likely to earn a bachelor's degree than their non-rural peers (Wells et al., 2019).

As a potential remedy to these inequalities, 34 states provide additional funding to rural districts through grants and multipliers that account for their low enrollment, low density of students, and/or isolated location (Education Commission of the States, 2021). These funding programs vary widely in terms of their eligibility requirements and generosity (Gutierrez and Terrones, 2023). But despite their prevalence and their relevance to policymakers, there are few attempts in the literature to estimate how school districts use the additional funds provided by these programs and how they influence student outcomes. Understanding the spending and achievement impacts of these state funding programs is important because, while prior literature documents that, on average, increases in school funding improve student outcomes (Jackson, 2020), the heterogeneity in observed effects (Jackson and Mackevicius, 2023) and unique challenges of rural education make the efficacy of increased funding to rural districts less certain.

For example, given rural districts' distinct cost structures, they may allocate additional funds differently than their urban or suburban counterparts, generating different effects on student achievement and educational attainment.

In this paper, we evaluate the impact of Wisconsin's Sparsity Aid program, which is one of the largest state-level school finance programs targeting small, rural districts. Currently, the program provides \$28 million in additional, unrestricted funding to 185 districts in the state. Our empirical approach leverages the introduction of the program in 2008 and subsequent expansion in 2010 in event study and difference-in-differences designs that compare the outcomes of school districts that were eligible and ineligible for sparsity aid, before and after the policy changed. Using comprehensive school finance data from the U.S. Department of Education's Common Core of Data (CCD), we first show that receiving a sparsity aid payment increases annual spending on elementary and secondary education by approximately \$226 per student, or 2% of average spending.

To further understand how district spending responds to an increase in funding, we analyze detailed, district-level spending data from the Wisconsin Department of Public Instruction (DPI). We document that districts use sparsity aid dollars in a variety of ways and, in general, tend to allocate funds to non-instructional areas: we estimate positive, statistically significant increases in spending on administration, food service, and general operations, as well as non-salary spending on instruction outside of core academic subjects (e.g., extracurricular activities), due to the sparsity aid program. Notably, we find minimal effects on teacher staffing—including student/teacher ratios, average salaries, and experience levels—but find increases in the staffing of administrative positions. Furthermore, we find substantial heterogeneity in how districts allocate funds, with increases in spending in most categories inversely related to

districts' baseline budget shares. This finding suggests that the unrestricted nature of the sparsity aid program allows administrators to put the additional funding towards areas, particularly non-instructional areas, that were relatively underfunded prior to the program's introduction. In a survey of school district administrators, we confirm that districts used these funds flexibly and rarely earmarked them for specific purposes.

We then assess the impact of this increased spending on educational outcomes using student-level data from the DPI. We find little evidence that the increased spending resulting from the program substantially improved student performance on state standardized tests. Our preferred point estimate for average scores across grades and subjects is statistically indistinguishable from zero, which may be explained by the minimal increase in teacher staffing and instructional spending that we note above. We also see little effect of the program on behavioral outcomes, such as attendance and disciplinary incidence, and on postsecondary enrollment and completion rates for the full sample of students. For the subset of students eligible for free or reduced-price lunch (FRL), we find suggestive evidence that the sparsity aid program improved enrollment and completion at two-year colleges.

Despite our largely null effects on student outcomes, we note that our findings are not inconsistent with the existing school finance literature. In a meta-analysis of 31 studies that causally identify the effects of increased school spending on student outcomes, Jackson and Mackevicius (2023) document that, on average, a policy increasing spending by \$1,000 per student improves test scores by 0.032 standard deviations and increases college-going by 2.8pp. Given that the Wisconsin sparsity aid program increases school spending by about \$250 per student, we would expect test score gains of about 0.008 standard deviations and college-going gains of about 0.7pp if the returns to school spending in rural Wisconsin were similar to the

returns of previously studied policies. Our 95% confidence intervals do not exclude positive effects as large as these. Thus, we cannot rule out the possibility that larger increases in school spending in rural areas can have similar effects to increased spending in non-rural areas. However, our results highlight the importance of understanding how districts use increased funding and the potential for differences in districts' needs and allocation decisions across different contexts.

Our findings contribute to several lines of literature on rural education and public investments in K-12 schools. First, we contribute to a growing set of studies on a variety of education policies targeted at rural schools and students. In recent years, an increasing number of small and rural school districts have adopted four-day school weeks as a cost-saving measure (Thompson et al., 2021). While these adoptions have reduced expenditures (Thompson, 2021a), they have also caused a reduction in student achievement (Thompson, 2021b). Other rural communities have either chosen or been mandated to rein in costs by consolidating school districts (Duncombe and Yinger, 2007). Such consolidations do not necessarily affect districts' economies of scale (Gordon and Knight, 2008), but they may lead to increased student performance (McGee et al., 2022) at the expense of population growth and property values (Smith and Zimmer, 2022). Some states have also begun to allocate funding to rural communities to specifically address teacher shortages, which Tran and Smith (2021) find modestly reduces teacher turnover. Other states have attempted to retain teachers and bolster student success by providing professional development in specific content areas to teachers in rural areas, realizing significant math gains (Barrett et al., 2015). While these studies consider narrowly prescribed interventions, we examine the efficacy of additional, unrestricted funding that allows rural districts to respond to their unique needs.

We also add to a large literature on how school spending affects student outcomes. Much of this literature addresses the endogeneity of school spending by exploiting variation in court-ordered school finance reforms that weakened the correlation between district wealth and per-student spending. These studies generally find positive effects of increased school spending on test scores (Papke, 2005; Roy, 2011; Lafortune et al., 2018), high school completion (Candelaria and Shores, 2019), educational attainment (Hyman, 2017), adult earnings (Jackson et al., 2016), and income mobility (Biasi, 2021b). These effects are generated by sustained increases in school spending that are much larger than the increases induced by the Wisconsin sparsity aid program or by other grant programs that target rural school districts. For example, Lafortune et al. (2018) estimate that, on average, a school finance reform increases state funding for low-income districts by \$1,225 per student per year and for high-income districts by \$527 per student per year. The sparsity aid program, in contrast, increases funding by \$226 per student per year during the time frame of our analysis. In this way, our findings are more closely aligned with studies of smaller-scale investments in schools, such as funding for new technology purchases (Bass, 2021) or textbooks (Holden, 2016), which have been shown to generate improvements in student achievement. Our paper is distinguished from this prior work by the fact that the sparsity aid program allows for unrestricted spending—along with its focus on rural schools.

To our knowledge, only one existing study considers the effect of a sparsity aid policy like Wisconsin's.⁶ Kreisman and Steinberg (2019) exploit two features of Texas' school finance system that provide additional funding to geographically large districts with low enrollment.

⁶ We note that some other papers, such as Hyman (2017) and Rauscher (2020) consider heterogeneous effects of funding in rural areas, as opposed to urban or suburban setting. However, they do not study policies specifically targeted at rural districts.

Using regression discontinuity and regression kink designs, they find that an additional \$1,000 in funding per year over students' schooling years improves reading scores by 0.1 standard deviations, improves math scores by 0.07 standard deviations, decreases high school dropout rates by 1.6 percentage points, and increases college enrollment among students who take college preparatory exams (e.g., SAT, AP exams) by 10 percentage points. Kreisman and Steinberg are not able to fully measure enrollment changes in the two-year sector because many students who enroll in a two-year college do not take these exams.

We build on this prior work by providing new evidence on the impacts of increased education funding to rural communities within a different demographic and policy context. The Wisconsin policy we study serves districts that are, on average, much smaller and less dense than the Texas districts studied by Kreisman and Steinberg. The policy is also structured as a stand-alone grant, rather than being embedded within the state's general aid formula, which may affect how the program is perceived and used by district administrators. In addition, our empirical approach that leverages the introduction and expansion of the program allows us to estimate effects across all eligible districts, rather than those situated at the enrollment and density cutoffs, which are the largest and least sparse of districts receiving additional funding. In doing so, our paper considers how the effects of additional funding to rural school districts may vary across settings and provides new insights into how programs similar to Wisconsin's may affect the outcomes of similarly small and sparse rural districts.

Policy Background & Data

Policy Introduction

Wisconsin is home to 421 unique school districts, each of which is funded by a combination of state aid (46.1%, on average), local property taxes (42.2%), federal funding

(7.2%), and other local revenue sources (4.5%) (Kava and Pugh, 2019). The majority (79%) of state aid is allocated via a “general aid” formula that distributes funds based on districts’ per student value of taxable property to equalize funding across districts with low and high property tax revenues. The remainder (21%) of state aid is allocated via categorical aid programs, which are designed to fund specific costs faced by districts, such as special education, transportation, and limited-English proficiency (LEP) programs. Unlike general aid, categorical aid programs are distributed without regard to the district’s local property tax revenues and are not subject to state revenue limits that cap the total amount of general state aid and local property tax revenues a district can receive.⁷ As such, categorical aid programs can increase a district’s resources even if they are not eligible to receive additional general aid.

In 2007, under Wisconsin Act 20, the state legislature added a new categorical aid program called the Sparsity Aid Program. The goal of the program was to provide additional unrestricted funds to rural school districts experiencing small economies of scale.⁸ Initially, districts were eligible to receive sparsity aid funds if (1) they had a pupil membership of no more than 725 in the prior year, (2) they had a density of fewer than 10 members per square mile in the prior year; and (3) at least 20% of the district’s students in the prior year were eligible for free or reduced-price lunch. The program first became active in the 2008-2009 school year, during which eligible districts whose FRL percentage fell between 20 and 50 percent were slated to receive \$150 per student while those whose FRL percentage exceeded 50 percent received \$300 per student. However, the total program appropriation was not large enough to make these payments, so payments were prorated to \$67 per student and \$134 per student, respectively. Beginning in 2009, the legislature removed the bifurcation and all eligible school districts were

⁷ For a full discussion of revenue limits in Wisconsin and the impact on students of raising them, see Baron (2022).

⁸ See the Wisconsin Sparsity Aid Program website for a full legislative history.

eligible to receive \$300 per student. Once again, the program was underfunded and, due to proration, eligible districts only received \$69 per student.

In 2010, the Wisconsin legislature significantly expanded funding for sparsity aid, from \$3.5 million per year to nearly \$15 million per year. Today, the program allocates almost \$28 million per year to 185 districts and is one of the largest categorical aid programs in the state.⁹ From the 2010-2011 school year onward, eligible districts received between \$237 and \$400 per student per year, including any necessary prorations. The only other change in program eligibility during our time frame occurred in 2015 when the FRL requirement (which was not binding for any otherwise eligible districts) was removed and the enrollment eligibility threshold increased from 725 to 745. Figure 1 plots the average sparsity aid amount received by districts, both in total and on a per-student basis, that are consistently eligible for sparsity aid from 2008 to 2017.¹⁰ Prior to 2008, districts received no sparsity aid. In 2008 and 2009, districts received an average of about \$32,480 annually, or \$72 per student per year. Since 2010, districts have received an average of \$115,350 annually or \$269 per student per year. For context, this total amount is approximately equal to 2.5 times the average full-time equivalent (FTE) teacher's salary in sparsity-eligible districts. Given that these districts employ an average of 35 teachers, this additional funding —while small in per- student terms relative to previously studied school finance interventions —represents a meaningful increase in districts' available resources.

Data Sources

Our primary data source for student outcomes is the Wisconsin DPI student-level records

⁹ Source: Wisconsin DPI. The 188 count of districts in the 2022-2023 includes two who receive “stop-gap” sparsity aid (introduced in the 2017-2018 school year to provide 50% of the prior year's sparsity aid grant to schools who lose eligibility) and 33 who receive Tier 2 sparsity aid funding (introduced in the 2021-2022 school year to provide a reduced sparsity aid grant to schools with membership between 745 and 1,000).

¹⁰ Of the 106 districts initially eligible for the program in 2008, 104 remain eligible across the entire time series

from 2005-2006 through 2017-2018. These records contain demographic information, enrollment history, attendance data, disciplinary infractions, and standardized test scores for every student who attended a Wisconsin public school in the time period. The state further links these records to postsecondary enrollment and completion records from the National Student Clearinghouse (NSC).¹¹ We supplement the student-level data with several sources of district-level data. We obtain a rich set of district-level demographics, enrollment, and financial information from the National Center for Education Statistics (NCES) Common Core of Data (CCD), along with annual sparsity aid payments and additional school finance outcomes —such as the revenue districts receive from local, state, and federal sources and the amount they spend on instruction, support, and administration —from the DPI. We also obtain district-level information on teacher and administrator staffing, including full-time equivalent (FTE) staffing levels, average salaries, and average years of experience, from the DPI. In addition, we follow Bayer et al. (2021) to aggregate annual census tract-level house price index data from the Federal Housing Finance Agency (FHFA) to the school district level to track district-level house price indices over time.¹²

To validate and expand upon results from these sources of administrative data, we additionally conducted a survey of rural school district leaders throughout Wisconsin. The survey primarily contained questions regarding the usage of sparsity aid funds, but it also measured general awareness of the program and the expected effects of receiving sparsity aid funding. We distributed the survey electronically to all superintendents/district administrators, principals, and financial officers who were employed by a district receiving sparsity aid funding

¹¹ Additional information about Wisconsin’s use of NSC data is available on the DPI website: <https://dpi.wi.gov/wisedash/districts/about-data/ps-enrollment>.

¹² All indices are measured relative to the first year the FHFA tracts data for a given tract. See Bogin et al. (2018) for more detail on the construction of this dataset.

in 2022. The Wisconsin Rural School Alliance (WiRSA) also advertised the survey in an email newsletter. Out of the 409 employees we attempted to contact via email, 39 (9.5%) completed the survey, representing 37 distinct school districts (and one unnamed school district). For our analysis, we drop 3 observations that reported not working in a sparsity aid-eligible district. Appendix B contains the full text of the survey and recruitment email, and we reference the results throughout the text. Appendix Table B.1 further reports baseline summary statistics for the sample of districts represented by the survey respondents and the sample of sparsity aid-eligible districts without a survey respondent. The two samples are similar across a variety of dimensions, including size, finances, and student achievement. As such, we interpret our survey responses as representative of how a typical sparsity aid district perceived and used the sparsity aid program.

Sample Restrictions

Because the goal of our analysis is to compare the outcomes of similar districts that did and did not receive sparsity aid funding, we limit our sample to school districts that offer all grades K-12 and are either always or never eligible for the sparsity aid program.¹³ We drop 13 districts that have poor house price index coverage and 5 districts with implausibly large spikes in either revenue or spending that we believe are due to data reporting errors. We then drop 113 districts that were in the top 30% of the enrollment or density distributions prior to the policy's introduction in 2008. This restriction eliminates Wisconsin's largest cities and suburbs, which differ in multiple dimensions from rural areas, and may provide poor estimates for the

¹³ The majority of Wisconsin school districts offer all grades K-12. In 2007, there were 426 local school districts in the state, 369 (86.6%) of which offered grades K-12. 46 districts only offered grades K-8, while 1 only offered grades 6-12 and 10 only offered only grades 9-12.

counterfactual outcomes of rural districts had rural districts not received sparsity aid.¹⁴ Finally, because we will consider specifications with flexible region-specific time trends, we drop 14 non-sparsity districts located in regions where no districts in the region both meet the above criteria and receive sparsity aid payments.¹⁵

Our final sample consists of 89 districts that received sparsity aid payments beginning in 2008 and 99 districts that never received sparsity aid payments. Figure 3.2 identifies the locations of these districts. Both sparsity-eligible and ineligible districts in our sample are geographically distributed throughout Wisconsin and, often, eligible and ineligible districts are located next to one another. The only area of the state that our sample does not cover is the southeast region, in and around the Milwaukee metropolitan area.

Summary Statistics

Table 3.1 provides summary statistics on school districts in Wisconsin and in our analysis sample, averaged across the academic years 2003-2007. Panel A provides information on the size and location of sparsity-eligible and ineligible districts. Unsurprisingly given the eligibility guidelines of the program, sparsity-eligible districts are smaller and less dense than the districts in our comparison group. However, the comparison group itself is relatively small and sparse compared to Wisconsin as a whole: the average Wisconsin school district enrolled 2,035 students and 45.9 students per square mile before 2008, whereas our comparison group enrolled an average of 1,231 students and 9.31 students per square mile. Sparsity districts also tend to have fewer school buildings and fewer students per building than their non-sparsity peer districts in

¹⁴ In our results section and Appendix Figure A3.10, we consider alternative sample restrictions and show that our effects are similar across different enrollment and density criteria.

¹⁵ Throughout the analysis, we define school district regions using Wisconsin's Cooperative Educational Service Agency (CESA) definitions. CESAs are collections of adjacent school districts that facilitate communication and cooperation across districts in the same area of the state. More information is available on the Wisconsin DPI website.

the comparison group, with the comparison districts still being smaller than the overall Wisconsin average. 97.8% of school districts in our sample are characterized as “rural” by the NCES defined as an area outside of an urban cluster, as are the majority (70.7%) of districts in our comparison group.¹⁶

Panel B of Table 3.1 then provides summary statistics on the demographic characteristics of the two groups of school districts. The racial demographics of our treatment and comparison groups are similar, with both enrolling approximately 94% white students. In contrast, the average Wisconsin district was 88.7% white between 2003 and 2007. Students attending sparsity districts are somewhat more disadvantaged than our comparison districts, with an average FRL rate of 34.7% (vs. 23.7%) and an average local child poverty rate of 14.5% (vs. 9.5%). The house price index in sparsity districts is also about 36 percentage points lower than that in the comparison districts, though the comparison district group average is also lower than the Wisconsin average.

Next, Panel C compares district finances in sparsity-eligible districts to the comparison districts in our analysis sample. Sparsity districts both receive and spend more per student than the comparison group —and the state average —prior to the introduction of the sparsity aid program, which is surprising given that sparsity districts tend to be less wealthy than non-sparsity districts. However, in Appendix Figure A.3.1 we show that there is a striking non-linear relationship between district size and per-student finances: smaller districts receive and spend much more per student than larger districts do, perhaps due to fixed costs and economies of scale that allow districts to reduce per-student costs as enrollment increases. Thus, because sparsity

¹⁶ 28.3% of our comparison districts are located in “towns” (defined as an area within an urban cluster, but outside of the primary urbanized area) and 1% are located in suburban areas. No districts in our comparison group are located in urban areas. In contrast, 7% of Wisconsin school districts statewide are located in urban areas, with an additional 14.7% located in suburban areas.

districts are, by definition, smaller than their peer districts that are not eligible for sparsity aid, they tend to have higher revenues and expenditures on a per-student basis. Sparsity districts also tend to spend a larger share of their budgets on administrative and other operational costs (e.g., transportation, food service), as compared to instruction and support services.¹⁷

Panel D then compares teacher and administrator staffing across sparsity-eligible and ineligible districts. Due to their small size, sparsity districts employ fewer teachers and administrators (superintendents, principals, and directors/coordinators), but sparsity districts have similar —or a bit higher —staffing levels on a per-student basis. In addition, sparsity and non-sparsity districts employ teachers with similar levels of experience. Despite this similarity in experience, and the fact that sparsity districts spend more per-student overall, teachers in sparsity-eligible districts are, on average, paid roughly \$2,900 (6.7%) less than teachers in the comparison group. This disparity suggests that the higher spending levels in sparsity districts are not due to higher investments in teacher pay, but potentially a result of the higher per-student operating costs sparsity districts face due to their small size and lack of economies of scale.

Finally, Panel E highlights differences in baseline achievement between sparsity districts and the comparison group. While we might expect sparsity districts to outperform their sparsity-ineligible peers because of their higher levels of spending and smaller school sizes (Kuziemko, 2006; Gershenson and Langbein, 2015; Egalite and Kisida, 2016), prior to the sparsity aid reform, they tended to have lower levels of achievement than their non-sparsity peer districts and the state average. A smaller share of students grades 3-8 and grade 10 were rated as “proficient” on state math and reading exams, and a smaller share of students enrolled in college within one year of graduating from high school or completed college by the end of our data’s time frame.

¹⁷ We derive the budget shares in Table 1 from detailed, district-level financial data submitted to the Wisconsin DPI. We provide more information on this data source and the construction of budget shares in the staffing section.

These disparities lend credence to the idea that sparse, rural school districts face additional challenges in educating their students and motivate our analysis of how districts utilize increased state funding and whether a policy like the sparsity aid program can improve student outcomes.

Empirical Strategy

Difference-in-Differences Framework

Our empirical strategy leverages the introduction and subsequent expansion of the sparsity aid program, which generated exogenous increases in district funding for eligible districts and should not have affected ineligible districts. To demonstrate that the program indeed increased district revenues and expenditures, we begin by estimating the effect of the program on district-level outcomes by estimating equations of the following form:

$$Y_{dt} = \beta \text{SparsityAid}_{dt} + \mathbf{Z}_{dt} \boldsymbol{\Pi} + \theta_d + \delta_t + \varepsilon_{dt} \quad (1)$$

where Y_{dt} is an outcome (e.g., revenues or expenditures per student) for district d in year t and SparsityAid_{dt} indicates whether district d received sparsity funding in year t . This variable “turns on” for all eligible districts in 2008 and remains zero for all ineligible districts throughout the entire time frame of the sample. \mathbf{Z}_{dt} are time-varying district covariates (e.g., enrollment, student demographic composition) that may also affect a district’s outcomes over time. We discuss our choice of control variables in the context of our identification assumptions in the next section. θ_d are district-level fixed effects that capture any time-invariant characteristics of school districts (e.g., location within the state) and δ_t are year fixed effects that capture any state-wide changes in district finances or student outcomes over time. ε_{dt} is the error term. Throughout the analysis, we cluster all standard errors at the school district level.

The coefficient of interest in equation (1) is β , the difference-in-differences (DID) estimate of how a school district’s outcomes change when it becomes eligible for the sparsity aid

program. In order for β to represent the causal effect of the sparsity aid program on outcomes, it must be the case that sparsity-eligible districts' outcomes would have evolved the same as non-sparsity-eligible districts' outcomes had the sparsity aid program never been implemented.¹⁸

While this counterfactual assumption is inherently untestable, we assess its plausibility by extending our DID equation to the following event study specification:

$$Y_{dt} = \sum_{k=2003, k \neq 2007}^{2017} \beta_k \text{SparsityEligible}_d * 1[t = k] + \mathbf{Z}_{dt} \boldsymbol{\Pi} + \theta_d + \delta_t + \varepsilon_{dt} \quad (2)$$

where $\text{SparsityEligible}_d$ indicates that a district will be eligible for sparsity aid funding when the policy is implemented and k indexes years. The β_k coefficients, therefore, trace out differences in the trends between sparsity and comparison districts' outcomes before and after the sparsity aid policy was implemented in 2008. If the two groups were trending similarly prior to the policy, we expect that the β_k coefficients will be equal to 0 up until 2006.

We also extend our district-level regression from equation (1) to consider student-level outcomes for students in grades 3-12 by estimating equations of the following form:

$$Y_{igsdt} = \beta \text{SparsityAid}_{dt} + \mathbf{Z}_{dt} \boldsymbol{\Pi}_g + \mathbf{X}_{it} \boldsymbol{\Gamma}_g + \lambda_{sg} + \delta_{tg} + u_{igsdt} \quad (3)$$

where Y_{igsdt} is an outcome (e.g., standardized test score or college enrollment) for student i , who is enrolled in grade g in school s in district d in year t . SparsityAid_{dt} is equal to 1 if student i 's district d receives sparsity aid funding in year t . \mathbf{Z}_{dt} are the same time-varying district covariates as equation (1) and \mathbf{X}_{it} are student characteristics that may or may not vary over time, such as their race, gender, FRL status, and special education status. In specifications that include multiple grade levels, we allow the effects of both sets of covariates to vary by grade level. λ_{sg} are school- by-grade fixed effects that capture any time-invariant characteristics of individual

¹⁸ Since the treated districts in our sample all receive treatment at the same time and never lose their treated status, we do not face the econometric problems associated with staggered treatment timing identified in recent DID methodological research. See Roth et al. (2022) for a recent literature review.

schools at each grade level and δ_{tg} are year-by-grade fixed effects that capture any secular trends by grade level. When we estimate specifications with only one grade level—for example, postsecondary outcomes for graduating seniors—these fixed effects collapse to the school and year levels, as in our district-level regressions. $u_{igsd,t}$ is the error term. We continue to cluster standard errors at the district level and also extend equation (3) to an event study equation to test for pre-trends.

Identification Assumptions

Our DID empirical framework relies on the assumption that school districts ineligible for the sparsity aid program serve as valid counterfactuals for school districts eligible for the sparsity aid program. Functionally, this assumption can be broken down into two parts. The first part is that the outcomes of school districts eligible and ineligible for sparsity aid were trending similarly prior to the introduction and expansion of the program. The β_k coefficients in the event study specifications from equation (2) allow us to test this assumption directly.

The second part of our identification assumption is the untestable assumption that there are no changes in unobserved determinants of our outcome measures that occur concurrently with the introduction of the sparsity aid program and which differentially affect sparsity-eligible and ineligible districts. This assumption could be threatened if there are (1) policy changes surrounding the introduction of the sparsity aid program that differentially affect sparsity eligible or ineligible districts and/or (2) if there are underlying demographic and economic trends that differ between sparsity districts and our comparison group. We address both sets of identification challenges in the sections that follow.

Concurrent Policy Changes

While we are unaware of any policy changes that occurred alongside the introduction of

the sparsity aid program and specifically targeted sparsity-eligible or ineligible districts, there were several other education policy changes in Wisconsin during the time frame of our sample that may threaten our empirical approach. One of the largest education-related policy changes in Wisconsin during the past 20 years was the passage of Act 10 in 2011, which discontinued collective bargaining requirements over teachers' salaries. As school districts' existing collective bargaining agreements (CBAs) expired in the years following 2011, they were able to pay teachers outside of standard salary schedules. Biasi (2021a) shows that the end of these CBAs and the subsequent introduction of flexible pay raised salaries of teachers with high value-added (VA) measures, increased cross-district teacher mobility to districts with flexible pay, and improved student achievement. Biasi and Sarsons (2022) further show that the adoption of flexible pay schemes following Act 10 induced a gender wage gap in teacher salaries.

Because Act 10 occurred at the state level and did not target rural school districts, it is not obvious that its introduction would threaten our identification strategy. However, the policy change could have differentially impacted sparsity-eligible districts if they (1) had collective bargaining agreements that expired earlier or later than those in the comparison districts and/or (2) if they employed higher or lower VA teachers, who faced different incentives to move to flexible play districts after Act 10. While we lack the data to answer these questions precisely, we provide evidence below that a variety of teacher-related staffing outcomes—including the number of teachers, salary distributions, and experience—trended similarly in sparsity eligible and ineligible districts from 2003 through 2017. Thus, we do not believe that the introduction of Act 10 poses a threat to our identification of the effects of the sparsity aid program.

Besides Act 10, there were two smaller policy changes in Wisconsin during our sample period that may have affected sparsity-eligible and ineligible districts differently. First,

beginning in the 2014-2015 academic year, Wisconsin changed its standardized testing regime three times in three years due to a combination of technical troubles in transitioning to online exams and a series of legislative decisions related to the national Common Core curriculum (Mason, 2016). It is reasonable to expect that sparsity-eligible districts—which are smaller, have fewer specialized staff, and may face additional barriers to internet access—were less equipped to deal with these regime changes than our comparison districts. In addition, it is unclear how to compare exam results over time given the changes in content and modality. As such, we limit our analysis of test scores to those through the 2013-2014 academic year.

The second policy change that may have differentially affected sparsity districts was the addition of a “high-cost pupil transportation aid” categorical aid program beginning in 2013-2014.¹⁹ As of 2019, the program stipulates that districts receive additional transportation funding if their transportation cost per student is greater than 145% of the state average in the prior year and their density is less than or equal to 50 students per square mile (Wisconsin Legislative Fiscal Bureau, 2019). The grant is not provided on a per-student basis, but the average per-student amount per school was \$185 in the 2019-2020 school year. While sparsity districts may be more likely to meet these criteria, districts in our comparison group are also relatively sparse and, thus, may also qualify for the program. Indeed, using data from the DPI on eligibility and payments for the program, we find that 86% of sparsity districts received payments from the program in at least one year from 2013-2017 and 33% of comparison districts did. Given this variation, we present specifications that control for districts’ receipt of additional transportation aid. We find that doing so minimally changes our results, indicating that this policy change is not a main driver of our findings.

¹⁹ Additional information on the high cost pupil transportation aid program is available on the DPI website: <https://dpi.wi.gov/sfs/aid/categorical/high-cost-pupil-transportation-aid>.

Demographic & Economic Trends

While we believe our results are robust to the various policy changes we discuss above, we also consider whether our identification assumptions may be threatened by demographic and economic trends over our analysis period. We note that sparsity-eligible districts experience more pronounced decreases in membership—in relative terms—during the time frame of our analysis than their sparsity-ineligible counterparts. Between 2003 and 2017, sparsity-eligible districts saw an average membership decrease of 76.3 students—or 15.3% of their baseline membership. In contrast, membership in sparsity-ineligible districts declined by an average of 61.7 students which, due to their larger size, represents only a 5% decrease in their baseline membership. Appendix Figure A.3.2 presents these membership trends, both in raw numbers and in relative decreases from districts’ 2003 baselines. Panel A of Appendix Figure A.3.3 then presents event study estimates of districts’ log membership before and after the sparsity aid program’s introductions, both with year fixed effects and with year-by-region fixed effects to capture the fact that different regions of the state may be experiencing different migration and fertility trends over time. Both sets of estimates show a consistent decline in membership in sparsity districts that begins before the program began and continues after.

Appendix Figures A.3.2 and A.3.3 provide little evidence that the decline in membership in sparsity districts differentially changes when the sparsity aid program is introduced. This consistent downward membership trend, combined with the relatively small amount of funding provided to districts—particularly in the first years of the program—makes it unlikely that households re-sorted between school districts in response to the policy. As further evidence against household resorting in response to the policy, Panel B of Appendix Figure A.3.3 presents event study specifications of districts’ retention rates: the share of students in grades K-11

enrolled in the district in year $t - 1$ that continue to be enrolled in the district in year t . We see little evidence that districts' retention of students changes differentially across sparsity and non-sparsity districts when the sparsity aid policy is introduced. If anything, we see a slight increase in the retention of students in sparsity districts across our analysis period that does not differentially change when the policy begins. The fact that we see few changes in districts' retention of their students also indicates that the downward membership trends we document in Panel A are not driven by increasing rates of students leaving sparsity districts. Rather, the declines in membership are the result of smaller and smaller cohorts entering the districts, which are likely reflective of broader birth rate and population declines in Wisconsin's most rural areas in the 2000s and 2010s (Forward Analytics, 2020).

While it is unlikely that the underlying membership trends we document are related to the sparsity aid program, they could raise two concerns for our empirical strategy. First, larger relative membership declines in sparsity districts may reflect or induce changes in districts' demographic characteristics that are related to academic achievement outcomes. Second, changes in membership will mechanically affect districts' per-student financial outcomes, such as revenues and spending per student. As such, we control for districts' log membership in all of our empirical specifications. In addition, we directly test whether districts' demographic characteristics change differentially in sparsity and non-sparsity districts over our sample period. The remaining panels of Appendix Figure A.3.3 present event studies for select characteristics. In Panel C, we see little change in the share of students who qualify for free or reduced-price lunch, indicating that the declines in membership in sparsity districts occur evenly across lower- and higher-income students. Similarly, in Panel D, we see little change in the share of students identified as eligible for special education services. In Panel E, we see a slight increase in the

share of students in sparsity districts, as compared to our comparison group, who are white. This trend appears to be the result of some- what faster racial diversification in our comparison group: in 2003, both sparsity and non-sparsity districts were approximately 95% white, while in 2017, sparsity districts were 89% white and non- sparsity districts were 87% white. To capture these modest compositional changes, our preferred empirical specifications control for districts' racial composition (% white, % Black, % Hispanic, and % Asian), along with their FRL percentage, special education percentage, and the local child poverty rate.

A related concern to declines in membership is the potential for different effects of the Great Recession on sparsity-eligible and ineligible districts. Given the large role of local property taxes in Wisconsin's school finance system, differential changes in local home values during the housing and financial crisis could result in differential changes in school district resources over the same time period. Appendix Figure A.3.4 plots changes in districts' house price indices (HPIs) over time. Panel A presents averages of the HPIs, while Panel B standardizes each district's index relative to 2003. While home prices in sparsity-eligible and ineligible districts are generally trending similarly prior to the start of the sparsity aid program, sparsity-eligible districts did not experience as large of a decline in the 2007-2012 period as sparsity-ineligible districts, which had higher prices prior to the start of the Great Recession.

Panel A of Appendix Figure A.3.5 presents event study estimates of districts' log-HPI, which are somewhat attenuated by the inclusion of year-by-region fixed effects. Panel B shows similar effects for the log of total property values in the district, as reported by the DPI. Panels C and D then provide event study estimates for per-student property values and property taxes, which mechanically capture both the changes in property values and the changes in membership described above. Given the concurrent decline in membership and slight increase in home

values, both measures exhibit upward pre-trends which continue after the introduction of the sparsity aid program. While the increase in property taxes per student does not affect eligible districts' receipt of sparsity aid funds, it could affect the total state revenue they receive since Wisconsin's state finance system equalizes resources between districts with low and high property tax revenues. To show that this increasing trend does not contaminate our results, in Appendix Figure A.6 we show that these differential trends largely disappear if we include year-by-region FEs and control for both a district's log membership and log-HPI, which we include in our preferred empirical specifications. Further, we later show that our estimates of the effects of the sparsity aid program on districts' finances are robust for controlling directly for districts' property value per student or property tax revenue per student.

In summary, to address potential threats to our identification assumption, our preferred specifications include district-level control variables that capture relevant changes in districts' demographic and economic conditions over time that may be related to their financial and student achievement outcomes. Specifically, we control for a district's log membership, log house price index, the number of school buildings, racial composition, % FRL, % special education, and the local child poverty rate, as well as region-by-year FEs. In the results that follow below, we show that our estimated effects of the sparsity aid program on districts' finances are generally similar with and without these controls, further validating our choice of the comparison group and difference-in-differences framework.

Effects of Sparsity Aid Program on District Finances

We begin our analysis by showing how the sparsity aid program affected eligible districts' revenues and overall spending. In the next section, we investigate districts' allocation of sparsity funds and changes in school inputs, particularly staffing decisions. Finally, we

estimate how these spending changes affected student outcomes.

Figure 3.3 presents the event study estimates from equation (2) for districts' financial outcomes. First, in Panel A, we consider the relationship between a district's sparsity aid eligibility and the state revenue they receive from sources other than the general aid formula. From 2003 to 2007, there is no differential trend in non-formula state revenue between sparsity-eligible and ineligible districts. Then, beginning in 2008, we see that sparsity eligibility districts see an increase in non- formula state revenue per student that is approximately the same size as the sparsity payments. This effect persists when the sparsity aid program is expanded in 2010 and becomes somewhat larger than the sparsity payments beginning in 2015. This shift is due to an expansion of high-cost pupil transportation aid in 2015, for which sparsity aid districts were more likely to be eligible. As discussed in Section above, we present specifications that control for districts' receipt of funding from the high-cost pupil transportation program when evaluating student-level outcomes.

In Panels B and C, we present event study estimates of a district's total revenue per student and total current spending on elementary and secondary education per student. For revenues, we see no evidence of pre-trends prior to the sparsity aid policy and clear increases in total revenues per student after the policy that are consistent with the size of sparsity aid payments. For expenditures, we again see evidence of parallel trends before 2008, but we do not see increases in the first two years of the program. This lack of a spending response could be driven by uncertainties over whether the program would be permanent, or it could be the case that other changes during the height of the Great Recession muted any effects that the sparsity aid program had on spending. Our survey respondents confirmed that uncertainty surrounding sparsity aid was common in the early years of the program; 20 of our 36 survey respondents said

they were unaware of the sparsity aid program when it was introduced, and another 8 believed it was unlikely the program would continue into the future. Nevertheless, beginning with the 2010 program expansion, we see increases in spending per student that align with the size of sparsity payments.

Table 3.2 summarizes these effects using the difference-in-differences specification from equation (1). Column (1) presents estimates only controlling for log membership, while column (2) adds demographic controls, column (3) interacts the year FEs with twelve different school district region indicators, and column (4) controls for whether a district receives additional transportation funding. The results are similar across specifications, indicating that demographic changes, transportation funding changes, nor regional trends are driving our results. In the most saturated specification in column (4), we find that receiving sparsity aid increases non-formula state revenue by \$217 per student (a 26.6% increase), total revenues by \$252 per student (a 1.9% increase), and current spending on elementary and secondary education by \$226 (a 2% increase). Given that the average sparsity-eligible district enrolled 434 members following the sparsity aid program's implementation, these increases translate to additional funding of approximately \$109,000 per year and additional spending of \$98,000 per year, more than twice the average teacher salary of sparsity-eligible districts during this time period.

We now conduct several robustness checks of these results. First, in Appendix Table A.3.1, we conduct placebo tests to further verify that these increases in revenue and spending are not driven by changes to other revenue sources or expenditures. In Panel A, we repeat our difference-in-differences specifications for per-student revenues from all sources other than non-formula state aid, which includes local property tax revenues and appropriations from the state general aid formula. Across specifications, we find no evidence that revenues from other sources

increased differentially in sparsity districts, relative to non-sparsity districts, indicating that the sparsity aid program alone was responsible for increasing districts' revenues. In Panel B, we further consider district spending in all areas other than elementary & secondary education, including capital outlays, community and adult education programs, payments to other government entities, and debt interest payments. Spending in these areas is typically financed via other revenue sources and, as such, we do not find that spending in these areas changed in sparsity districts following the implementation of the sparsity aid program. Together, the results in Appendix Table A.3.1 bolster our claim that the increases in revenues and expenditures we document in our main results are the result of the sparsity aid program.

Appendix Figure A.3.7 verifies that our revenue increase is driven by the sparsity aid program by plotting the difference-in-differences coefficients for all 35 revenue sources included in the CCD dataset. While we see some substitution between state general aid revenues and local revenues—which is consistent with declining membership and increased per-student property tax revenue we document above—the increase in total revenues is primarily driven by state revenue for other programs, which contains sparsity aid payments. In Appendix Figure A.3.8 we show that our total revenue and current spending event study estimates hardly change if we include additional controls for a district's total property value per student or local property tax revenue per student. Moreover, Appendix Figure A.3.9 further shows that our difference-in-differences estimates for total revenues and current spending are robust to controlling for districts' per-student revenue from local, state formula, or federal sources. Thus, the increased revenue and spending in sparsity-eligible districts appears to come from the sparsity aid program and not other changes in revenue sources during the time frame of the data.

Finally, we provide evidence that our estimates are not driven by our sample selection

criteria. Recall that our preferred sample removes districts that were in the top 30% of the enrollment or density distribution prior to the start of the sparsity aid program in 2008. In Appendix Figure A.3.10, we repeat the estimation of β from column (4) in Table 3.2, varying the percentage of districts dropped from 0% (meaning we do not drop any districts based on their pre-2008 enrollment or density) to 50% (meaning we drop all districts in the top half of the density or enrollment distributions). Across our four outcomes, the estimates vary only slightly across different sample definitions and in no case are statistically different from our preferred specification.

Allocation of Spending and Staffing

We now investigate how the sparsity aid program affected districts' spending across budget categories and how these changes in budget allocation are reflected in staffing levels. These results provide important insights into how districts used the increased state funding they received under the sparsity aid program, which may generate different predictions of how the program affected student achievement outcomes.

Budget Categories

To examine how districts allocated their sparsity aid dollars, we leverage annual financial reports submitted by Wisconsin districts to the DPI that summarize all transactions occurring in a district in a fiscal year.²⁰ To structure our analysis, we bundle spending in eight distinct areas tracked by DPI: (1) general instruction in core curricular areas, (2) instruction in all other curriculum areas (e.g., physical education and co-curricular activities), (3) pupil support (e.g., health, guidance), (4) instructional staff support (e.g., curriculum development, training), (5) administration, (6) transportation, (7) food service operations, and (8) general operations (e.g.,

²⁰ These annual reports are available publicly on the DPI website: <https://dpi.wi.gov/sfs/reporting/safr/annual/data-download>.

maintenance, fiscal services).²¹

Panel A of Table 3.3 repeats the specification from column (4) in Table 3.2 for each of the above spending categories. Overall, we find that districts allocated a larger share of their additional spending to non-instructional areas than instructional areas. Specifically, receiving sparsity aid increased districts' general instructional spending by \$56.96, general operations spending by \$56.71, administrative spending by \$45.01, and food service spending by \$23.46.²² The latter three estimates are statistically significant at the 5% level, and these categories received the largest relative increases, with the estimates each representing 4-6% of sparsity-eligible districts' pre-program means. Panels B and C separate the changes in spending in each category into spending on employee salaries and benefits and all non-personnel spending. We find that districts increased salary and benefit spending related to administration and general operations by statistically significant amounts, whereas they increased non-personnel spending related to other instruction (e.g., extracurriculars), student transportation, and food service.

As a whole, the results in Table 3.3 show that districts did not primarily allocate sparsity aid funding to instruction and most likely did not use sparsity aid funds for a single use. Instead, they allocated the increased funding to a wide variety of areas. Our survey data further confirm this finding. 24 of 36 respondents said the sparsity dollars were rarely or never set aside for a specific purpose, and only 4 respondents said they were always earmarked for a specific purpose.

²¹ Consistent with measures of current elementary and secondary spending in the NCES Common Core of Data (CCD), we exclude capital outlays, debt service payments, inter-fund transfers, and the purchase of investment assets from our spending measures. We also follow Kelly and Farrie (2023) and exclude payments to other governmental entities and schools (e.g., charter schools), except for payments to other Wisconsin public school districts for special education services. The correlation between per-student revenues in the CCD and Wisconsin annual report data is 0.992 and, for expenditures, it is 0.982.

²² It is possible that this increase in food service spending could be driven in part by changes to the federal nutrition program following the passage of the Healthy, Hunger-Free Kids Act of 2010. However, in Appendix Figure A.11, we show that our estimated effects of the sparsity aid program on spending allocation are robust to controlling for districts' per-student revenues from local, state formula, or federal sources.

Given this finding, and the fact that districts are unrestricted in how they spend sparsity aid dollars, the average effects presented in Table 3.3 may mask important differences in how different districts use the funds. To explore heterogeneity across districts, we augment our main difference-in-difference specification with an interaction term that allows the effect of receiving sparsity aid to vary based on a district's average budget share in a given spending category in the pre-program period (2003- 2007). We define these budget shares as the total spending in a category divided by a district's total revenue and standardize them to have a mean of 0. We scale our effects such that the coefficients represent the change in the sparsity receipt effect per 1pp increase in pre-period budget share.

Table 3.4 presents these results and uses the linear interaction terms to estimate how spending effects vary across the budget share distribution of each spending category. For the categories of general instruction, other instruction, pupil support, instructional staff support, and food service, we estimate negative and statistically significant interaction terms that indicate that districts' spending allocations differ based on their baseline budget shares. Specifically, districts increase their spending more when they have a low baseline budget share in a given category. For example, districts at the 25th percentile of the general instruction budget share distribution increase spending on general instruction by a statistically significant \$109.40, whereas districts at the 75th percentile do not increase their spending in this category. In addition, while there are no effects on spending on instructional or pupil support services on average, districts at the bottom of the budget share distributions for these categories increase their spending by statistically significant and economically meaningful amounts. An opposite pattern emerges for student transportation spending, where the interaction term is positive and statistically significant, indicating that districts with high base- line transportation budget shares further increase their

transportation spending when they receive sparsity aid funds. As a whole, these heterogeneous effects provide evidence that the unrestricted nature of Wisconsin’s sparsity aid program allows districts to allocate funds toward areas that were relatively underfunded prior to the program’s introduction.

District Staffing

In Table 3.3, we see that receiving sparsity aid increases spending on salary and employee benefits, particularly in the administration budget category. Specifically, we find that districts increased administrative personnel spending by \$41.26 per student, or approximately \$18,000 for the average-sized district. To quantify how this spending increase affected the number and type of staff employed by sparsity-eligible districts, we estimate our difference-in-differences and event study specifications for the number of administrative FTEs per 100 students in three categories: superintendents, principals, and other administrative staff members. In interpreting these results, it is important to note that sparsity-eligible districts frequently employ less than one FTE in administrative positions and/or employ a single staff member across multiple positions. For example, prior to 2007, 35.7% of sparsity-eligible districts employed less than 1 FTE principal and 91.2% employed less than 1 FTE in other administrative positions, such as curriculum and special education directors.

Table 3.5 presents difference-in-difference results for our measures of administrative staff.²³ Panel A presents our main results, while Panel B adds an interaction term with districts’ standardized baseline administrative budget share, analogous to the specifications in Table 3.4. In column (1), we find that receiving sparsity aid increased total district administrators — inclusive of superintendents and assistant superintendents, principals and assistant principals, and

²³ In Appendix Figure A.12, we present event study estimates of all outcomes in Table 5, none of which indicate that the results are driven by differential pre-trends between sparsity and non-sparsity districts.

other directors and coordinators —by 0.021 per 100 students, or 9.1% of an FTE for the average-sized district. This effect was equal to approximately 4.5% of baseline staffing and, in Panel B, we show that it was larger in districts with low baseline administrative spending: districts at the 25th percentile of the baseline administrative budget share distribution increased staffing by 0.036 FTEs per 100 students (15.6% of an FTE for the average-sized district), while districts at the 75th percentile of the distribution did not meaningfully increase their administrative staffing.

Columns (2) through (6) of Table 5 estimate the effects of sparsity aid receipt on staffing increases across different administrative positions. In column (2) we find that sparsity aid receipt has little effect on the number of superintendents and assistant superintendents per 100 students. This finding is not surprising as, at baseline, over 98% of sparsity-eligible districts report employing a superintendent and only 1.1% of districts report having more than one. Thus, it is unlikely that districts used sparsity aid funds to hire additional superintendent positions.

In columns (3) through (6) we consider sparsity aid effects on principal and other administrator staffing. For each, we consider both the likelihood that a district reports employing non-zero FTEs and the total number of FTEs per 100 students. For principals, we find a statistically insignificant but positive effect on the likelihood that districts employ a principal (which 91.2% report at baseline), as well as on the number of principals per 100 students. Consistent with our spending patterns in Table 3.3, these effects are larger for districts with low baseline administrative spending. Districts at the 25th percentile of the baseline administrative budget share distribution are 4.8pp more likely to employ non-zero principal FTEs and employ 0.021 more FTEs per 100 students after receiving sparsity aid. The latter estimated effect is statistically significant at the 5% level and represents an increase of 8.9% of the baseline mean.

We find even larger effects for other administrative positions. In column (5), we find that

receiving sparsity aid funds increases the likelihood that a district employs non-zero FTEs in other administrative positions by 14.9pp. This effect is highly statistically significant and represents an effect size of over 50% of the baseline mean. This effect is slightly larger for districts with lower baseline administrative spending but remains statistically significant at the 75th percentile of the baseline distribution. In column (6), we find an increase of 0.013 FTEs per 100 students, which is less precise, but similarly large, representing 38% of the baseline mean. While we lack the precision to estimate these effects across more granular positions, the types of administrative staff that sparsity districts are most likely to employ in the post-2008 period include special education directors, business managers, and instruction/curriculum coordinators. Thus, it is reasonable to expect that sparsity aid funds allowed districts to hire staff in these positions.

In Appendix Table A.3.2 and Appendix Figure A.3.13, we estimate similar regressions for the number of teacher FTEs per 100 students and teacher salaries, benefits, and experience. Consistent with our imprecise and relatively small effect on teacher-related salary and benefit spending found in Table 3.3, we find little effect of receiving sparsity aid funds on teacher staffing, salaries, benefits, or experience. These results indicate that the sparsity aid program did not substantially change instructional inputs in eligible districts. Further, they help us rule out that Wisconsin Act 10 differentially affected sparsity and non-sparsity districts and, thus, is unlikely to be driving our results.

Taken together, our results indicate that sparsity aid funding allowed districts to increase spending on a variety of non-instructional areas, particularly those with low baseline budget shares. Districts used a portion of these additional funds to hire non-instructional staff, such as administrative positions. It is possible that these spending patterns may not meaningfully change

academic outcomes, such as test scores and college enrollment, which we consider in the next section.

Effects of Increased Spending on Student Outcomes

To estimate the impact of increased spending from the sparsity aid program on student outcomes, we rely on student-level administrative records from the Wisconsin DPI. We limit our sample to students in grades 3-12 who have non-missing demographic information and who continuously enroll in Wisconsin public schools. These restrictions produce a final sample of 308,630 unique students and 1,484,856 unique student-year observations, of which approximately one-quarter are enrolled in sparsity-eligible districts and three-quarters are not.

K-12 Outcomes

Table 3.6 presents estimates of β in equation (3) for student test score outcomes. Because of changes in Wisconsin's testing regime over time, our analysis is limited to math and reading test scores among students in grades 3-8 and grade 10 from 2005 to 2013, along with science, social studies, and writing scores for students in grades 4, 8, and 10 in the same years. We standardize all test scores at the year, grade, and subject level across the universe of test-takers to have a mean of zero and a standard deviation of one. Thus, our β estimates can be interpreted as the percent of a standard deviation change due to the sparsity aid program. We estimate effects for each grade level and test subject, as well as average effects across grade levels and across test subjects.

Panels A through E present our estimated effects for each test subject. For reading, science, social studies, and writing exams, we estimate small and statistically insignificant effects across all grade levels. In addition, our average effects across grade levels are close to zero and not statistically significant. However, for math exams, our average effect is negative

and statistically different from zero at the 10% level, mostly driven by a large negative effect among 6th graders. Panel F then presents average effects across all test subjects taken by a grade level, and across all grades and test subjects. For each grade level, our effects are small and statistically insignificant at conventional levels. The point estimates are particularly small and close to zero for grades 7, 8, and 10. Overall, we do not detect a statistically significant effect on test scores and, with 95% confidence, can rule out that the sparsity aid program increased test scores by more than 0.75% of a standard deviation. Moreover, our confidence intervals generally include the effect size we would expect from the Jackson and Mackevicius (2023) meta-analysis given the size of the spending increase in the sparsity aid program, suggesting that our estimates are not inconsistent with the school finance literature as a whole.²⁴

In Appendix Figure A.3.14 we present analogous event study estimates for each test subject to test whether our effects are driven by existing differential trends in test scores between students in sparsity-eligible and ineligible districts. While the pre-trends are generally flat, the availability of only three pre-treatment periods limits our ability to determine whether the negative effects we estimate —particularly for math test scores —may be due to a longer-run decline in performance in sparsity districts. To further assess the role of pre-trends in our estimates, we obtain a longer panel of school-level test score data for grades 4, 8, and 10 from the Wisconsin DPI. Appendix Figure A.3.15 plots the average school-level test scores for these grades for schools in sparsity- eligible and ineligible districts from 2002 to 2013 and Appendix Figure A.3.16 estimates event study specifications over this time frame. While the estimates are

²⁴ Jackson and Mackevicius document that, on average, an increase in school spending of \$1,000 per student increases test scores by 0.31 standard deviations. Given that the typical sparsity aid payment is about one-quarter of this size (\$250 per student), we would expect test scores impacts of about 0.008 (0.8%) standard deviations if the returns to this increased school spending were similar to that of previously studied policies and could be scaled linearly.

quite noisy, we do not see strong evidence of declining math scores prior to treatment years. In addition, we see little evidence of changes in reading test scores for sparsity districts before or after the sparsity aid program began, which aligns with our small, statistically insignificant results in Table 3.6. Finally, despite the substantially smaller size of the sparsity aid grant during the first two years of the program, we do not observe significantly different effects on math and reading test scores when looking at 2010 and onward as opposed to 2008 and 2009.

As a whole, we take our test score results to suggest that the spending patterns induced by the sparsity aid program did not meaningfully improve academic achievement in sparsity districts. If anything, test scores declined somewhat in sparsity districts, relative to non-sparsity districts, during the 2008–2013-time frame. These effects are not surprising given that districts did not increase instructional investments in response to the program’s funding. In addition, despite our robust set of control variables, we cannot rule out that our effects are driven by differential impacts of the Great Recession on sparsity and non-sparsity districts. Stated differently, our results indicate that increased, unrestricted spending to sparse, rural school districts during the Great Recession did not improve their standardized test scores.

A null effect on test scores does not rule out the possibility that the increased funding improves students’ non-cognitive and/or behavioral outcomes, particularly since districts appear to use their increased funds to hire administrative positions that may oversee these sorts of student outcomes—such as special education directors. In Appendix Table A.3.3 we consider how sparsity aid funding affects a variety of these outcomes, including students’ annual attendance rate, the likelihood they are involved in a disciplinary incident, the likelihood they repeat a grade level, and, for 10th–12th graders, the likelihood they dual-enroll in a college

course while enrolled in a Wisconsin public high school.²⁵ Across the different outcomes and grade levels, we estimate precise null effects. Appendix Figure A.3.17 presents event study estimates of these outcomes. The estimates are noisy but do not suggest that the main results are driven by differential pre-trends between sparsity-eligible and ineligible districts. As such, we interpret our findings as indicating that the sparsity aid program also had little effect on students' behavior in schools.

Postsecondary Enrollment & Completion

We now turn to estimating how exposure to the sparsity aid program affects students' longer-run educational attainment by estimating effects on both postsecondary enrollment and completion. While we do not observe positive effects on student achievement or behavioral outcomes, it is still possible that the spending patterns we document in Section 5 could improve postsecondary outcomes. For example, the administrative positions that districts hire may oversee college and career preparation activities.

Table 3.7 presents estimates of β from equation (3) on the sample of all seniors in our sample from 2005 to 2017. We estimate effects for the entire sample of students and separately for students who are and are not eligible for free and/or reduced-price lunch (FRL) to test whether additional resources boosts college attendance and completion for low-income students who, at baseline, are less likely to attend and complete college.

Panel A of Table 3.7 presents our estimated effects for the full sample of students. Columns (1) to (3) show effects for enrollment in college within 12 months of high school

²⁵ We consider outcomes for grades 6-12 in this table, as there is little variation in attendance and disciplinary incidence in elementary grades.

graduation.²⁶ Overall, we estimate that the sparsity aid program increased college enrollment by 0.3pp, with a larger effect in two-year colleges (0.6pp) than in four-year colleges. These effects are not statistically different from zero, but similar to our test score impacts, our confidence intervals contain effect sizes that we would expect from the prior literature.²⁷ Columns (4) to (6) then estimate the effects for college completion at any point in our data's time frame. We similarly see about a 0.7pp increase in college completion, which is again not statistically different than zero.²⁸

Panels B and C then estimate separate effects for FRL eligible and ineligible students. In Panel B, we find large point estimates of the sparsity aid program on college enrollment and completion for FRL eligible (lower-income) students, especially for the two-year sector, but none of the estimated effects are statistically different from zero at conventional levels. In Panel C, we find even less evidence that the sparsity aid program improved college enrollment and completion for non-FRL eligible students—all the point estimates are practically and statistically indistinguishable from zero. In Appendix Figures A.3.18 and A.3.19, we present the corresponding event study estimates for post-secondary enrollment and completion. The estimated coefficients on the few periods we have available prior to the sparsity aid program's implementation do not reveal any clear violations of our parallel trends assumption, and we see similar effect sizes across the post- treatment period. Thus, overall, our results do not provide

²⁶ Due to changes in how Wisconsin reported high school graduation data during the time period of our analysis, we only observe high school graduation dates for students who enroll in college and are not able to consider graduation as an outcome directly.

²⁷ Jackson and Mackevicius (2023) document that, on average, an increase in school spending of \$1,000 per student increases test scores by 2.7pp. Given that the typical sparsity aid payment is about one-quarter of this size (\$250 per student), we would expect test scores impacts of about 0.7pp if the returns to this increased school spending were similar to that of previously studied policies and scaled linearly.

²⁸ Because later cohorts in our data have had fewer years to enroll in and complete college, we also estimate effects only on the 2005-2013 cohorts. Appendix Table A.4 presents these results, which are very similar to our main results.

strong evidence that the sparsity aid program increased college enrollment and completion for affected students overall, but we find suggestive evidence that there may have been modest increases for low-income students.

Conclusion

Rural schools and school districts in the United States face a myriad of distinct challenges compared to their urban and suburban peers, including high per-student costs in areas like transportation and frequent staffing turnover. States often recognize these challenges and many provide additional funding to rural districts as a result. However, there is limited work on how districts use these funding sources, nor how increased spending generated by them affects student outcomes like test scores and postsecondary enrollment.

We provide new evidence on the returns to school funding in rural districts by exploiting policy variation from Wisconsin's sparsity aid program—one of the largest state-level funding streams geared towards rural school districts. We find that the introduction and subsequent expansion of the program increased school spending by about 2% annually. Using detailed school finance data from the Wisconsin DPI, we show that the sparsity aid program increased spending more in non-instructional areas than instructional areas. Moreover, we find that districts increased spending the most in areas where, at baseline, they had low budget shares compared to other districts, suggesting that districts were able to use the sparsity aid funds flexibly to supplement areas that were relatively underfunded prior to the introduction of sparsity aid.

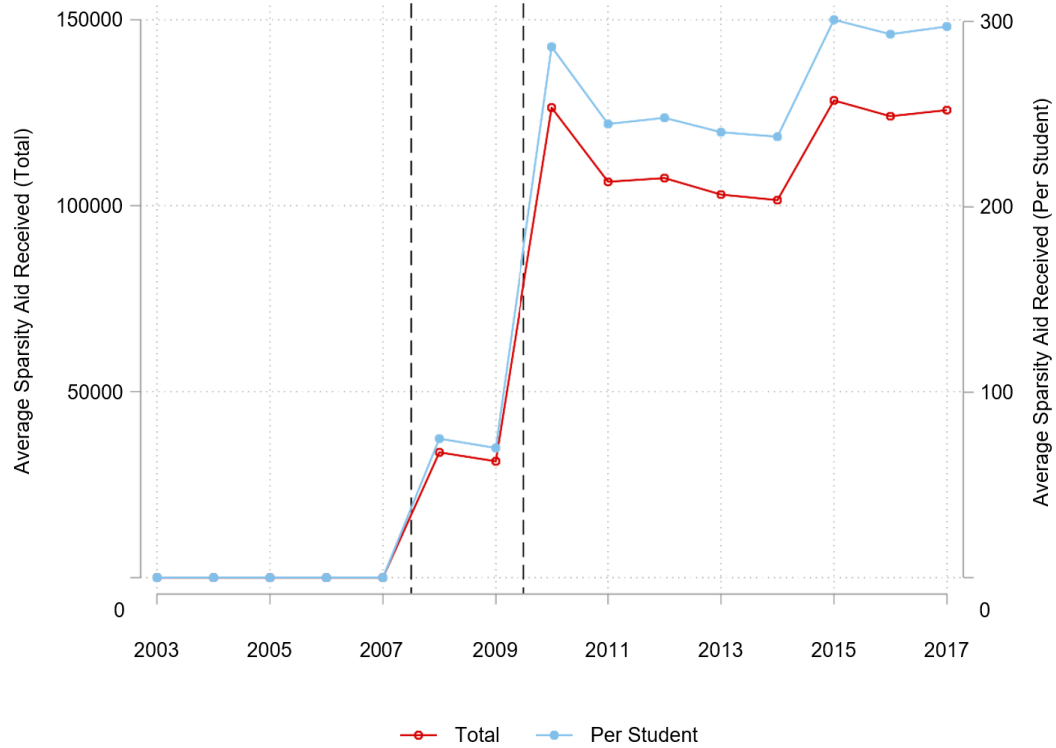
We generally find that the additional funding did not change standardized test scores, behavioral outcomes, college enrollment, nor college completion, but we find suggestive evidence that the program may have boosted college attainment for low-income students,

particularly at two-year colleges. Our largely null effects on academic outcomes are consistent with how districts allocated sparsity aid funds toward non-instructional areas. These results thus underscore the importance of understanding how districts allocate increased funding to contextualize achievement effects, particularly when policies provide districts flexibility in how funds are used.

Finally, we highlight that, despite our largely null results, our results are broadly consistent with the existing school finance literature. Across the outcomes we study, we generally cannot rule out that our effect sizes are statistically different than the effect sizes we would expect if the returns to school spending in rural districts were the same as in previously studied contexts (Jackson and Mackevicius, 2023). Thus, while the relatively small revenue increases districts received due to the sparsity aid program did not meaningfully affect student outcomes, it is possible that larger increases to rural school districts' budgets would generate positive effects on student achievement and postsecondary outcomes. Future work that considers policy interventions with larger increases to rural school districts' budgets would be a valuable contribution to the literature and could inform the structure of sparsity aid policies across states.

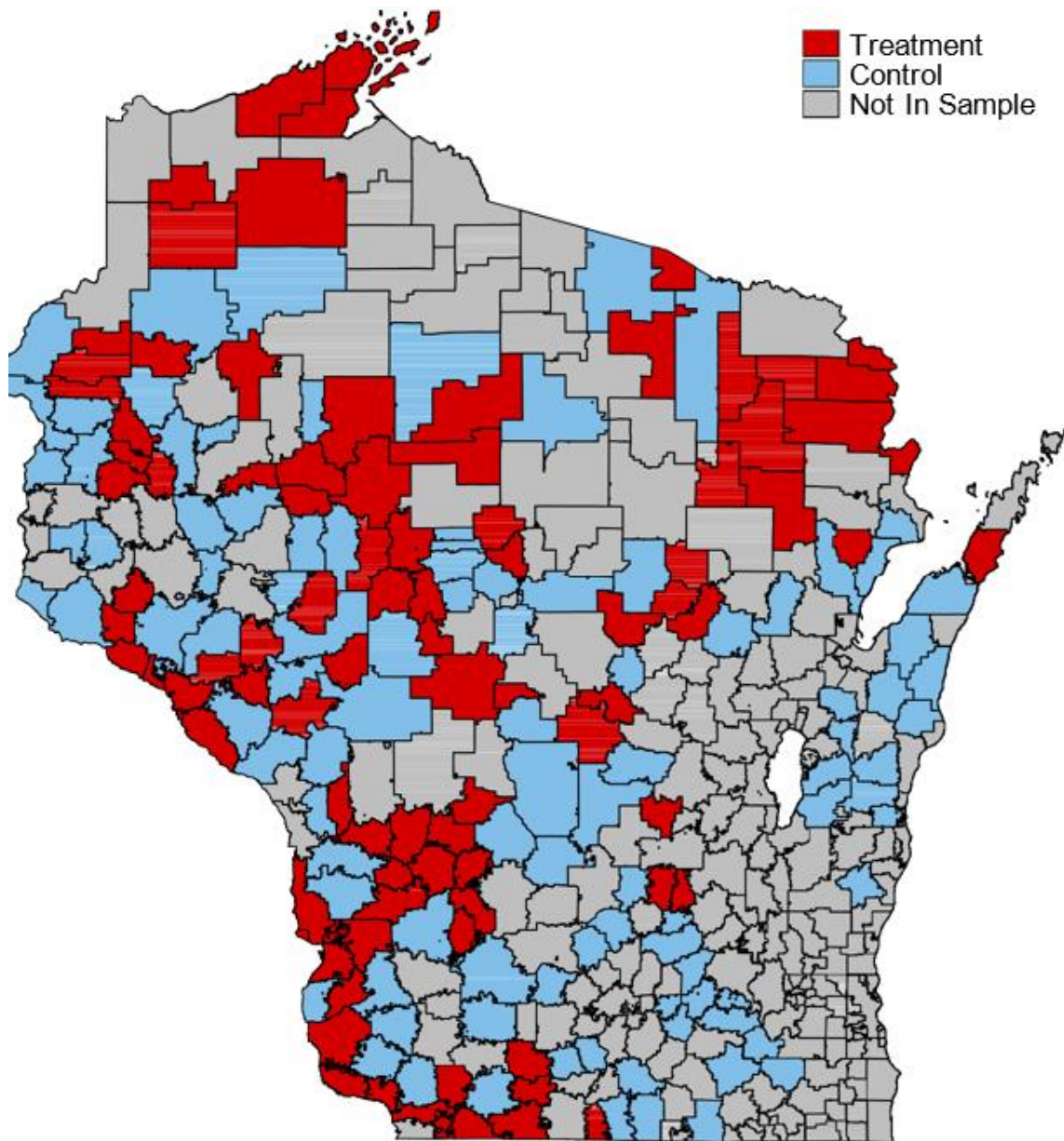
TABLES AND FIGURES

Figure 3.1. Introduction & Expansion of Sparsity Aid Program



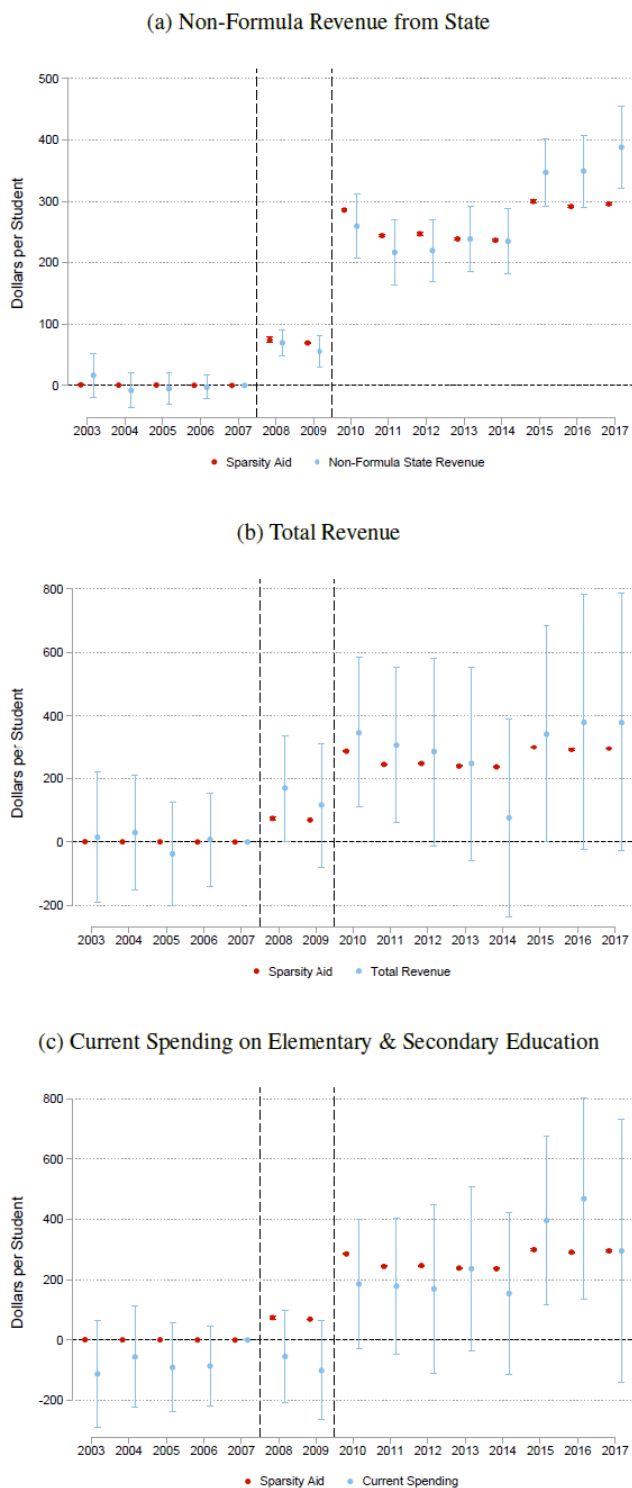
Notes: This figure shows the average sparsity aid funding received each year, both in total and per student, among districts in our sample that are always eligible for sparsity aid funding from 2008 to 2017.

Figure 3.2. School Districts in Analytic Sample



Notes: This figure shows the sparsity-eligible (treatment) and ineligible (comparison) districts in our sample.

Figure 3.3. Event Study Estimates of Sparsity Aid Program on District Finances



Notes: Each figure presents estimates of the β_k coefficients from equation (2). All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level.

Table 3.1. Baseline District Characteristics

	Wisconsin	Analysis Sample		
	Average	Sparsity	Comparison	p-value
	(1)	(2)	(3)	(4)
Panel A. Size & Location				
Membership	2035	480.1	1231.4	0.000
Membership per Square Mile	45.94	3.809	9.312	0.000
Number of Schools	4.941	2.544	3.889	0.000
Avg. Membership per School	355.9	201.1	332.9	0.000
NCES Rural Classification	0.580	0.978	0.707	0.000
Panel B. Demographics				
% White	0.887	0.940	0.940	0.979
% FRPL	0.239	0.347	0.237	0.000
% Special Education	0.145	0.157	0.144	0.000
Local Child Poverty Rate	0.098	0.145	0.095	0.000
District House Price Index	207.586	163.024	198.858	0.000
Panel C. Finances				
Revenue per student	11,789	12,502	11,122	0.000
Spending per student	9,966	10,354	9,408	0.000
% Instruction	0.698	0.669	0.689	0.000
% Support	0.073	0.064	0.073	0.000
% Administration	0.068	0.072	0.065	0.000
% Other	0.176	0.182	0.179	0.007
Panel D. Staffing				
Number of Teachers (FTE)	131.4	37.71	86.54	0.000
Teachers per 100 Members	7.324	8.027	7.045	0.000
Average Teacher Salary	44,087	40,396	43,305	0.000
Average Teacher Experience	15.476	16.029	15.946	0.589
Number of Administrators (FTE)	7.933	2.336	5.523	0.000
Administrators per 100 Members	0.467	0.496	0.450	0.000
Panel E. Educational Outcomes				
Math Proficiency Rate	0.439	0.393	0.418	0.001
Reading Proficiency Rate	0.361	0.324	0.346	0.000
College Enrollment Rate	0.542	0.518	0.538	0.028
College Completion Rate	0.403	0.386	0.405	0.021
Observations	2,310	445	495	940
Districts	462	89	99	188

Notes: Each column summarizes district-level characteristics over the 2003-2007 academic years. The test score and postsecondary outcomes are averaged over the 2005-2007 academic years. The college enrollment rate is defined as the share of high school seniors who enroll in a postsecondary institution within one year of graduating from high school and the college completion rate is defined as the share of high school seniors in the 2005-2007 cohorts who completed a postsecondary credential within the time frame of our data.

Table 3.2. Effect of Sparsity Aid on District Revenues & Spending

	(1)	(2)	(3)	(4)
Panel A. Sparsity aid dollars				
Received sparsity aid	223.6***	222.2***	221.6***	219.1***
	(1.9)	(2.2)	(2.3)	(2.5)
Observations	2,820	2,820	2,820	2,820
Panel B. Non-formula revenue from state				
Received sparsity aid	253.7***	259.7***	229.9***	217.1***
	(23.4)	(22.1)	(20.6)	(20.4)
	[31.1%]	[31.9%]	[28.2%]	[26.6%]
Observations	2,820	2,820	2,820	2,820
Panel C. Total revenue				
Received sparsity aid	261.6**	263.8**	257.6**	252.4**
	(114.4)	(104.6)	(118.9)	(118.6)
	[1.96%]	[1.97%]	[1.93%]	[1.89%]
Observations	2,820	2,820	2,820	2,820
Panel D. Current spending				
Received sparsity aid	301.5***	291.6***	247.7**	226.3**
	(103.1)	(94.6)	(99.4)	(101.6)
	[2.70%]	[2.61%]	[2.22%]	[2.03%]
Observations	2,820	2,820	2,820	2,820
Membership control	X	X	X	X
Demographic controls		X	X	X
Year-by-CESA FEs			X	X
Transportation funding control				X

Notes: Each coefficient is estimated from a separate regression and represents β in equation (1), the effect of receiving sparsity aid funding. Column (1) controls for a district's log membership, column (2) adds controls for a district's log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, and the local child poverty rate, column (3) adds year-by-region (CESA) fixed effects, and column (4) further controls for whether a district receives funding from the state's high-cost pupil transportation aid program. All standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.3. Effect of Sparsity Aid on Spending Allocations

	General Instruc.	Other Instruc.	Pupil Supp.	Instruc. Staff Supp.	Admin.	Student Transp.	Food Service	General Ops.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Effects on Total Spending</i>								
Received sparsity aid	56.96 (50.71)	24.08 (35.01)	-0.354 (14.25)	6.404 (18.24)	45.01** (18.94)	5.772 (9.290)	23.46*** (8.453)	56.71** (26.686)
Observations	2,820	2,820	2,820	2,820	2,820	2,820	2,820	2,820
Baseline Mean	4185.3	1990.4	385.88	416.28	895.57	584.90	417.21	1282.27
<i>Panel B. Effects on Salary & Employee Benefit Spending</i>								
Received sparsity aid	75.12 (48.25)	-5.047 (34.79)	-11.99 (12.12)	2.976 (14.17)	41.26** (16.26)	-16.42* (9.871)	9.731 (7.382)	35.01* (18.09)
Observations	2,820	2,820	2,820	2,820	2,820	2,820	2,820	2,820
Baseline Mean	3981.4	1810.13	272.87	282.40	778.20	174.83	221.74	662.73
<i>Panel C. Effects on Other Spending</i>								
Received sparsity aid	-18.16 (12.27)	29.13*** (10.75)	11.64 (8.914)	3.428 (8.052)	3.747 (7.012)	22.19* (13.27)	13.73* (8.126)	21.70 (21.85)
Observations	2,820	2,820	2,820	2,820	2,820	2,820	2,820	2,820
Baseline Mean	203.86	180.27	113.00	133.89	117.37	410.07	195.47	619.54

Notes: Each coefficient is estimated from a separate regression and represents β in equation (1), the effect of receiving sparsity aid funding. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.4. Heterogeneous Effects of Sparsity Aid on Spending

	General Instruc.	Other Instruc.	Pupil Supp.	Instruc. Supp. Staff	Admin	Student Transp.	Food Service	General Ops.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Received sparsity aid	43.605 (50.087)	17.34 (34.819)	-4.052 (14.879)	-7.235 (20.088)	54.56*** (18.759)	3.833 (9.081)	22.41*** (8.341)	60.12** (26.399)
Received sparsity aid x budget share	-28.640** (11.192)	-23.37** (10.149)	-36.03** (16.193)	-34.65** (15.332)	-27.15** (13.321)	13.66** (6.901)	-25.99** (12.888)	19.53 (13.598)
Baseline Mean Spending \$	4185.3	1990.4	385.88	416.28	895.57	584.9	417.21	1282.27
Baseline Mean Share	0.342	0.163	0.032	0.037	0.068	0.044	0.033	0.104
Observations	2,820	2,820	2,820	2,820	2,820	2,820	2,820	2,820
Effect at 10th Percentile	198.3***	121.7**	37.41**	54.21**	91.15***	-14.4	36.96***	21.38
Effect at 25th Percentile	109.4**	64.52	26.48*	37.63*	66.06***	-4.62	31.65***	35.64
Effect at 50th Percentile	54.25	21.90	5.305	6.751	42.72**	5.813	23.01***	55.85**
Effect at 75th Percentile	-6.66	-5.96	-21.9	-21.4	21.60	17.20	14.96*	73.44**
Effect at 90th Percentile	-55.3	-54.7	-37.1	-42.4	4.132	34.12*	3.147	100.7**

Notes: The coefficients in each column are estimated from a separate regression and represent β in equation (1), the effect of receiving sparsity aid funding, with interaction effects based on districts' pre-2008 budget shares. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.5. Effects of Sparsity Aid on Administrator Staffing

	Total Admin	Superintendent	Principals		All Other	
	FTEs per 100 Students (1)	FTEs per 100 Students (2)	Non-Zero FTEs (3)	FTEs per 100 Students (4)	Non-Zero FTEs (5)	FTEs per 100 Students (6)
Panel A. Main Specification						
Received sparsity aid	0.021*	-0.005	0.030	0.014	0.149***	0.013
	(0.012)	(0.006)	(0.020)	(0.010)	(0.046)	(0.008)
Baseline Mean	0.465	0.193	0.912	0.237	0.283	0.034
Observations	2,820	2,820	2,820	2,820	2,820	2,820
Panel B. Interaction with Baseline Admin Budget Share						
Received sparsity aid	0.028**	-0.003	0.038*	0.017*	0.152***	0.014*
	(0.012)	(0.006)	(0.023)	(0.009)	(0.046)	(0.008)
Received sparsity aid x budget share	-0.019**	-0.006	-0.023	-0.010*	-0.007	-0.003
	(0.008)	(0.004)	(0.017)	(0.006)	(0.023)	(0.004)
Baseline Mean	0.465	0.193	0.912	0.237	0.283	0.034
Observations	2,820	2,820	2,820	2,820	2,820	2,820
Effect at 25th Percentile	0.036***	-0.001	0.048*	0.021**	0.155***	0.015*
Effect at 75th Percentile	0.005	-0.010	0.010	0.005	0.143***	0.011

Notes: The coefficients in each column are estimated from a separate regression and represent β in equation (1), the effect of receiving sparsity aid funding, with interaction effects based on districts' pre-2008 budget shares. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.6. Effect of Sparsity Aid on Standardized Test Scores

	3rd Grade (1)	4th Grade (2)	5th Grade (3)	6th Grade (4)	7th Grade (5)	8th Grade (6)	10th Grade (7)	All Grades (8)
Panel A. Reading								
Received sparsity aid	-0.010 (0.024)	-0.018 (0.021)	-0.020 (0.021)	-0.007 (0.020)	0.021 (0.019)	-0.022 (0.020)	-0.010 (0.020)	-0.009 (0.012)
Observations	90,631	91,595	92,516	95,059	98,256	100,523	110,142	678,722
Mean	0.047	0.046	0.027	0.057	0.049	0.045	0.039	0.044
Panel B. Math								
Received sparsity aid	-0.015 (0.031)	-0.036 (0.028)	-0.034 (0.028)	-0.062** (0.028)	-0.026 (0.026)	-0.017 (0.024)	-0.026 (0.020)	-0.030* (0.015)
Observations	90,896	91,720	92,612	95,125	98,328	100,574	110,164	679,419
Mean	0.016	0.013	-0.014	-0.006	0.011	0.040	0.046	0.016
Panel C. Science								
Received sparsity aid		0.007 (0.022)				-0.019 (0.022)	-0.008 (0.018)	-0.007 (0.013)
Observations		91,757				100,556	110,096	302,409
Mean		0.079				0.086	0.085	0.083
Panel D. Social Studies								
Received sparsity aid		-0.014 (0.022)				-0.020 (0.021)	0.003 (0.019)	-0.010 (0.013)
Observations		91,731				100,407	110,061	302,199
Mean		0.082				0.100	0.079	0.087
Panel E. Writing								
Received sparsity aid		-0.010 (0.021)				0.002 (0.022)	-0.012 (0.018)	-0.007 (0.014)
Observations		91,607				100,457	109,935	301,999
Mean		0.032				0.025	0.022	0.026
Panel F. Average								
Received sparsity aid	-0.012 (0.025)	-0.015 (0.020)	-0.028 (0.022)	-0.035 (0.022)	-0.002 (0.020)	-0.015 (0.019)	-0.011 (0.016)	-0.016 (0.012)
Observations	90,619	91,578	92,495	95,036	98,231	100,470	110,025	678,454
Mean	0.032	0.044	0.007	0.026	0.031	0.050	0.045	0.034

Notes: Each coefficient is estimated from a separate regression and represents β in equation (1), the effect of receiving sparsity aid funding on standardized test scores. Average test scores are calculated for students with non-missing math and reading scores. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.7. Effect of Sparsity Aid on Postsecondary Enrollment & Completion

	Any (1)	Two-Year (2)	Four-Year (3)	Any (4)	Two-Year (5)	Four-Year (6)
<i>Panel A. All Students</i>						
Received sparsity aid	0.003 (0.009)	0.006 (0.008)	-0.002 (0.007)	0.007 (0.009)	0.001 (0.006)	0.007 (0.007)
Observations	165,442	165,442	165,442	165,442	165,442	165,442
Mean	0.558	0.237	0.340	0.345	0.161	0.209
<i>Panel B. FRL Eligible Students</i>						
Received sparsity aid	0.021 (0.018)	0.020 (0.017)	0.002 (0.012)	0.014 (0.013)	0.011 (0.011)	0.004 (0.011)
Observations	37,017	37,017	37,017	37,017	37,017	37,017
Mean	0.381	0.208	0.183	0.209	0.128	0.092
<i>Panel C. FRL Ineligible Students</i>						
Received sparsity aid	-0.003 (0.010)	0.003 (0.008)	-0.006 (0.008)	0.004 (0.010)	-0.000 (0.007)	0.005 (0.008)
Observations	128,419	128,419	128,419	128,419	128,419	128,419
Mean	0.610	0.246	0.386	0.385	0.170	0.243

Notes: Each coefficient is estimated from a separate regression and represents β in equation (3), the effect of receiving sparsity aid funding. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

BIBLIOGRAPHY

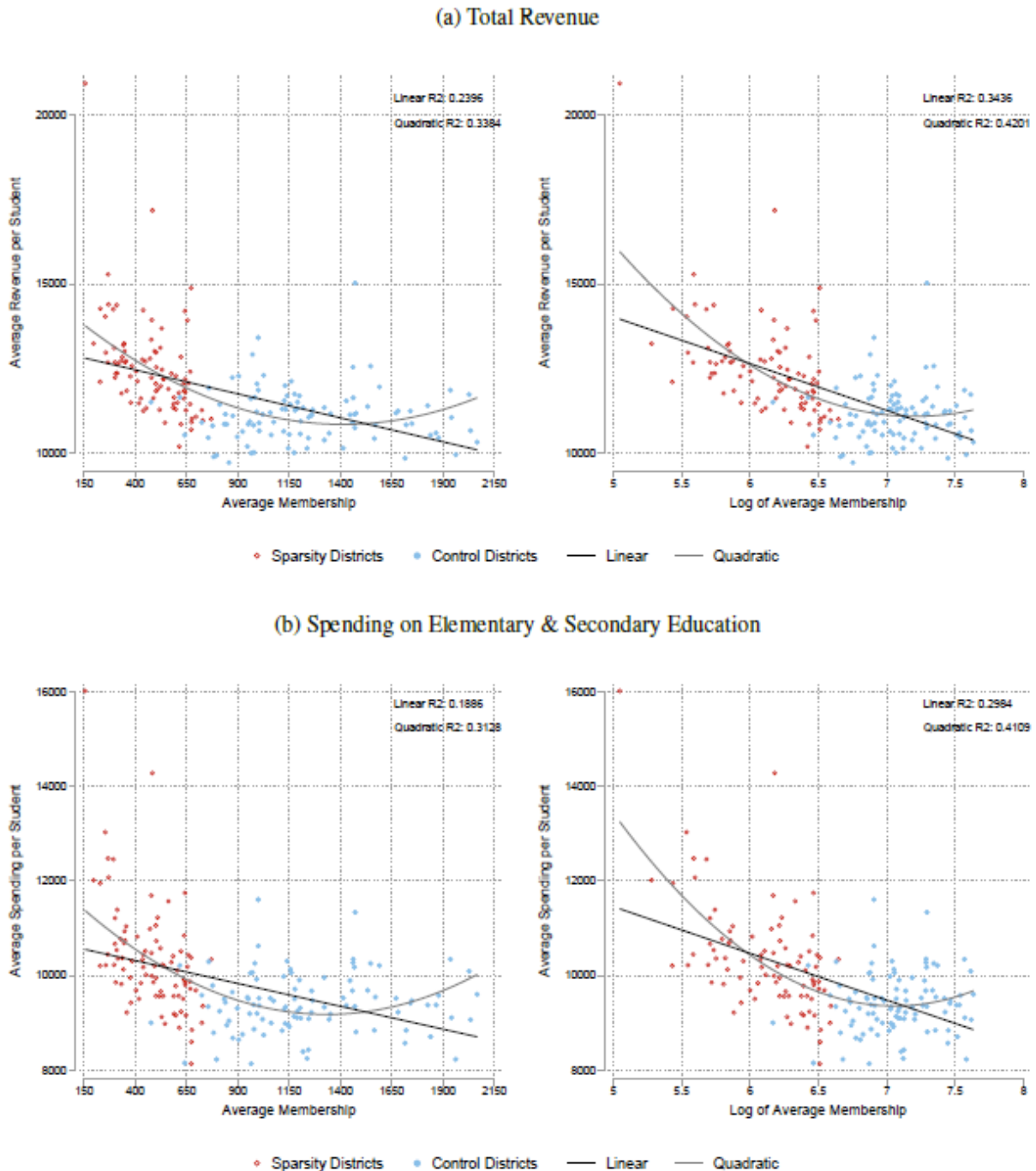
- Barrett, N., J. Cowen, E. Toma, and S. Troske (2015). Working with What They Have: Professional Development as a Reform Strategy in Rural Schools. *Journal of Research in Rural Education* 30(10), 1–18.
- Bass, B. (2021). The effect of technology funding on school-level student proficiency. *Economics of Education Review* 84.
- Bayer, P. J., P. Q. Blair, and K. Whaley (2021). Are we spending enough on teachers in the U.S.? Working Paper 28255, National Bureau of Economic Research.
- Biasi, B. (2021a). The labor market for teachers under different pay schemes. *American Economic Journal: Economic Policy* 13.
- Biasi, B. (2021b). School Finance Equalization Increases Intergenerational Mobility. *Journal of Labor Economics*, forthcoming.
- Biasi, B. and H. Sarsons (2022). Flexible wages, bargaining, and the gender gap. *Quarterly Journal of Economics* 137.
- Candelaria, C. A. and K. A. Shores (2019). Court-Ordered Finance Reforms in the Adequacy Era: Heterogeneous Causal Effects and Sensitivity. *Education Finance and Policy* 14(1).
- Duncombe, W. and J. Yinger (2007). Does school district consolidation cut costs? *Education Finance & Policy* 2(4), 341–375.
- Education Commission of the States (2021). K-12 and Special Education Funding. <https://reports.ecs.org/comparisons/k-12-and-special-education-funding-08>.
- Egalite, A. J. and B. Kisida (2016). School size and student achievement: a longitudinal analysis. *School Effectiveness and School Improvement* 27(3), 406–417.
- Forward Analytics (2020). The Rural Challenge: Depopulation and Its Economic Consequences. <https://www.forward-analytics.net/research/the-rural-challenge-depopulation-and-its-economic-consequences/>.
- Gershenson, S. and L. Langbein (2015). The effect of primary school size on academic achievement. *Educational Evaluation and Policy Analysis* 37(1S), 135S–155S.
- Gordon, N. and B. Knight (2008). The Effects of School District Consolidation on Educational Cost and Quality. *Public Finance Review* 36.
- Gutierrez, E. and F. Terrones (2023). Small and sparse: Defining rural school districts for k-12 funding. Research Paper, Urban Institute.

- Holden, K. (2016). Buy the Book? Evidence on the Effect of Textbook Funding on School-Level Achievement. *American Economic Journal: Applied Economics* 8(4).
- Hyman, J. (2017). Does money matter in the long run? Effects of school spending on educational attainment. *American Economic Journal: Economic Policy* 9.
- Jackson, C. K. (2020). Does school spending matter? The new literature on an old question. In L. Tach, R. Dunifon, and D. Miller (Eds.), *Confronting inequality: How policies and practices shape children's opportunities*, pp. 165–186. American Psychological Association.
- Jackson, C. K., R. C. Johnson, and C. Persico (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics* 131(1).
- Jackson, C. K. and C. Mackevicius (2023). What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the U.S. *American Economic Journal: Applied Economics*, forthcoming.
- Kava, R. and C. Pugh (2019). State Aid to School Districts. Informational Paper 24, Wisconsin Legislative Fiscal Bureau.
- Kreisman, D. and M. P. Steinberg (2019). The effect of increased funding on student achievement: Evidence from Texas's small district adjustment. *Journal of Public Economics* 176.
- Kuziemko, I. (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review* 25(1), 63–75.
- Lafortune, J., J. Rothstein, and D. W. Schanzenbach (2018). School Finance Reform and the Distribution of Student Achievement. *American Economic Journal: Applied Economics* 10(2), 1–26.
- Mason, K. C. (2016). Wisconsin's Second New State Test In 2 Years Rolls Out Quietly. *Wisconsin Public Radio*. <https://www.wpr.org/wisconsins-second-new-state-test-2-years-rolls-out-quietly>.
- McGee, J. B., J. N. Mills, and J. S. Goldstein (2022). The Effect of School District Consolidation on Student Achievement: Evidence From Arkansas. *Educational Evaluation and Policy Analysis*.
- Papke, L. A. (2005). The effects of spending on test pass rates: evidence from Michigan. *Journal of Public Economics* 89.
- Roy, J. (2011). Impact of School Finance Reform on Resource Equalization and Academic Performance: Evidence from Michigan. *Education Finance and Policy* 6.
- Showalter, D., S. L. Hartman, J. Johnson, and B. Klein (2019). Why Rural Matters 2018-2019. Report, Rural School and Community Trust.

- Sipple, J. W. and B. O. Brent (2015). Challenges and Strategies Associated with Rural School Settings. In H. F. Ladd and M. E. Goertz (Eds.), *Handbook of Research in Education Finance and Policy*, pp. 607–622.
- Smith, S. A. and R. Zimmer (2022). The Impacts of School District Consolidation on Rural Communities: Evidence from Arkansas Reform. EdWorkingPaper 22-530, Annenberg Institute for School Reform at Brown University.
- Thompson, P.N. (2021a). Does a day lost equal dollars saved? The effects of four-day school weeks on school district expenditures. *National Tax Journal* 74.
- Thompson, P. N. (2021b). Is four less than five? Effects of four-day school weeks on student achievement in Oregon. *Journal of Public Economics* 193.
- Thompson, P. N., J. M. Schuna Jr., K. Gunter, and E. J. Tomayko (2021). Are all four-day school weeks created equal? A national assessment of four-day school week policy adoption and implementation. *Education Finance and Policy* 16.
- Tran, H. and D. A. Smith (2021). How Hard-to-Staff Rural School Districts Use State Funds to Address Teacher Shortages. *Journal of Education Finance* 47.
- Wells, R. S., C. A. Manly, S. Kommers, and E. Kimball (2019). Narrowed Gaps and Persistent Challenges: Examining Rural-Nonrural Disparities in Postsecondary Outcomes over Time. *American Journal of Education* 126.
- Wisconsin Legislative Fiscal Bureau (2019). High Cost Transportation Aid (DPI –Categorical Aids). https://docs.legis.wisconsin.gov/misc/lfb/jfcmotions/2019/2019_05_23/001_public_instruction/019_paper_580_high_cost_transportation_aid.

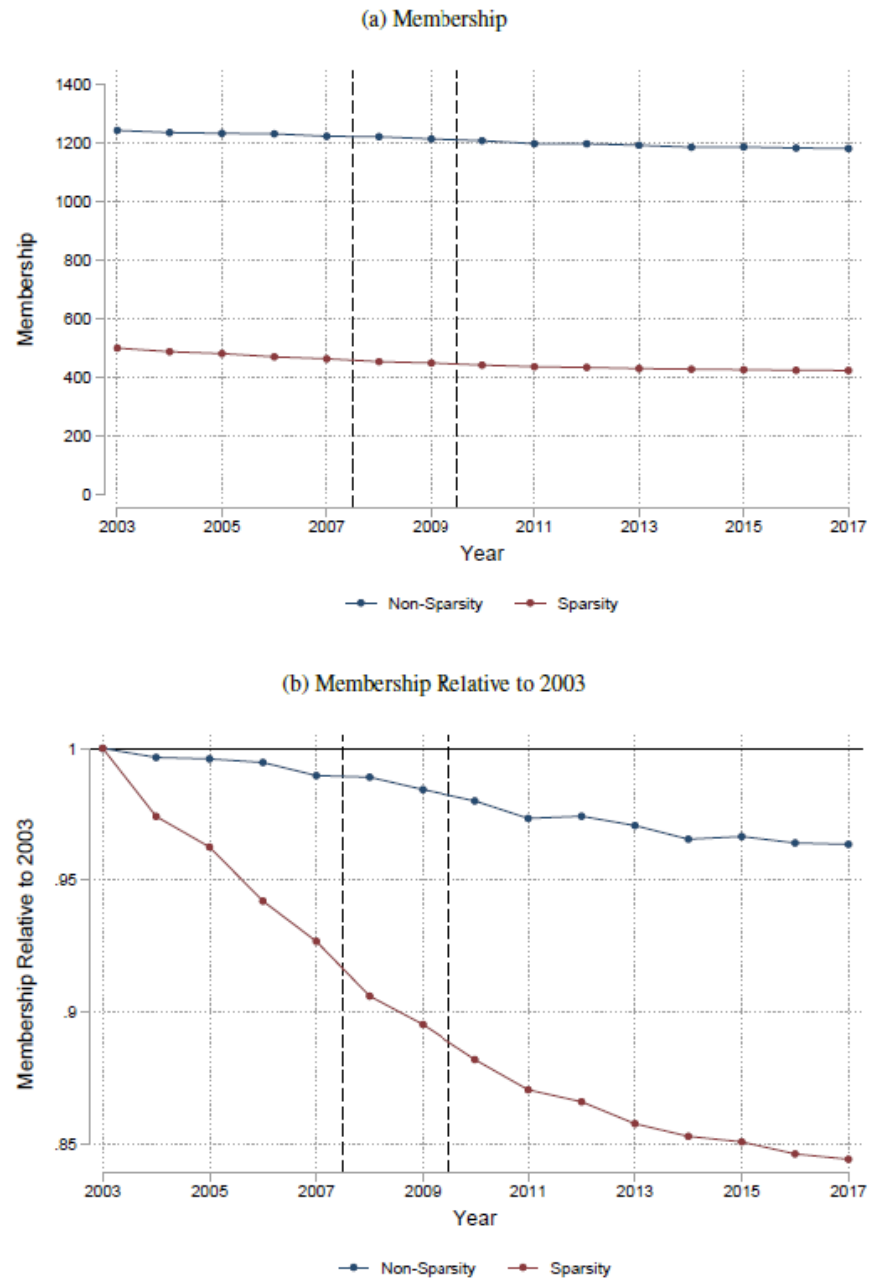
APPENDIX

Figure A3.1. Relationship Between District Size & Finances, 2003-2007



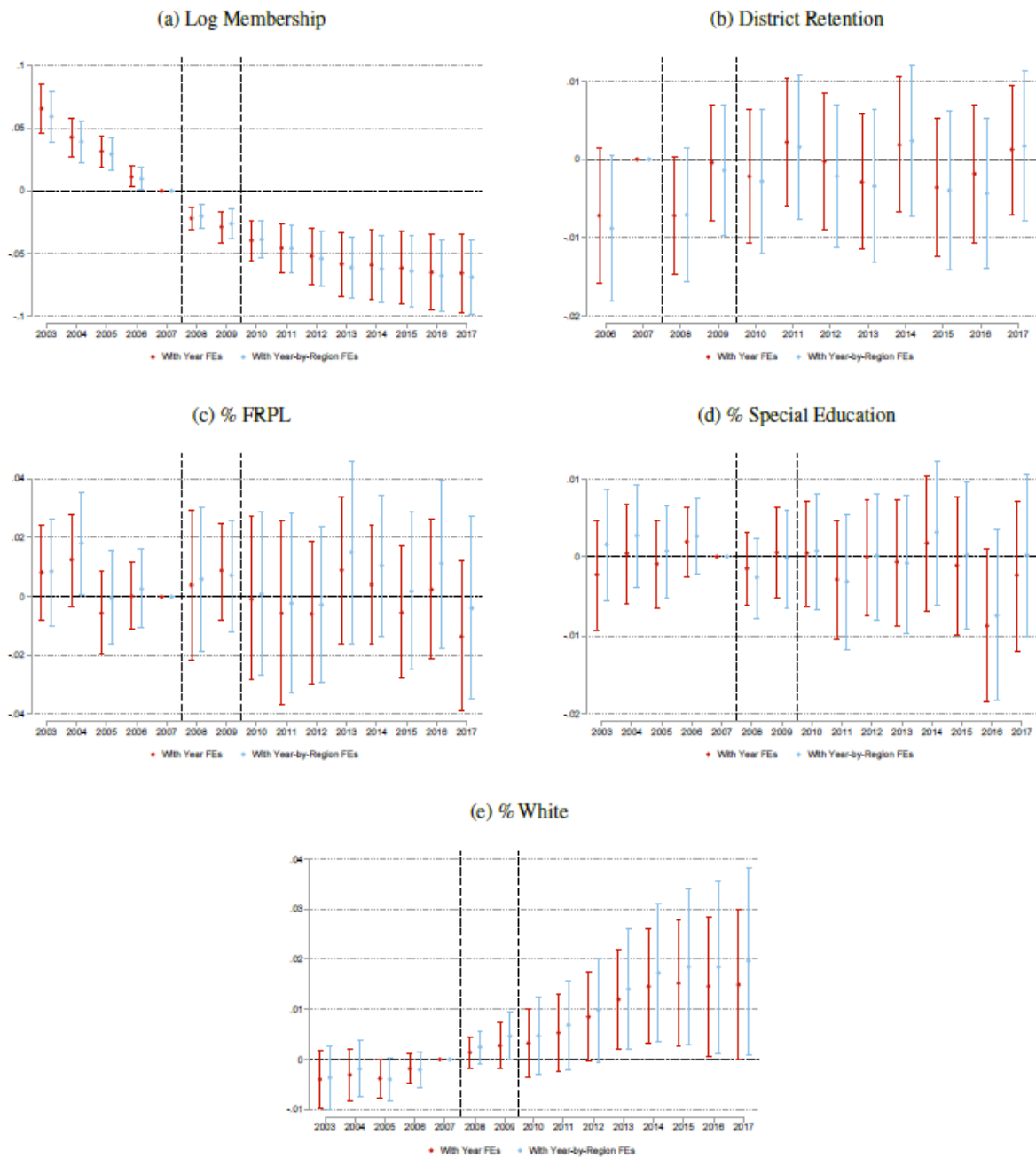
Notes: Each figure presents the relationship between districts' average resources per student and average membership in academic years 2003-2007. Districts that become eligible for the sparsity aid program in 2008 are shaded red, while districts that are never eligible for the sparsity aid program are shaded blue.

Figure A3.2. Membership in Sparsity-Eligible & Ineligible Districts, 2003-2017



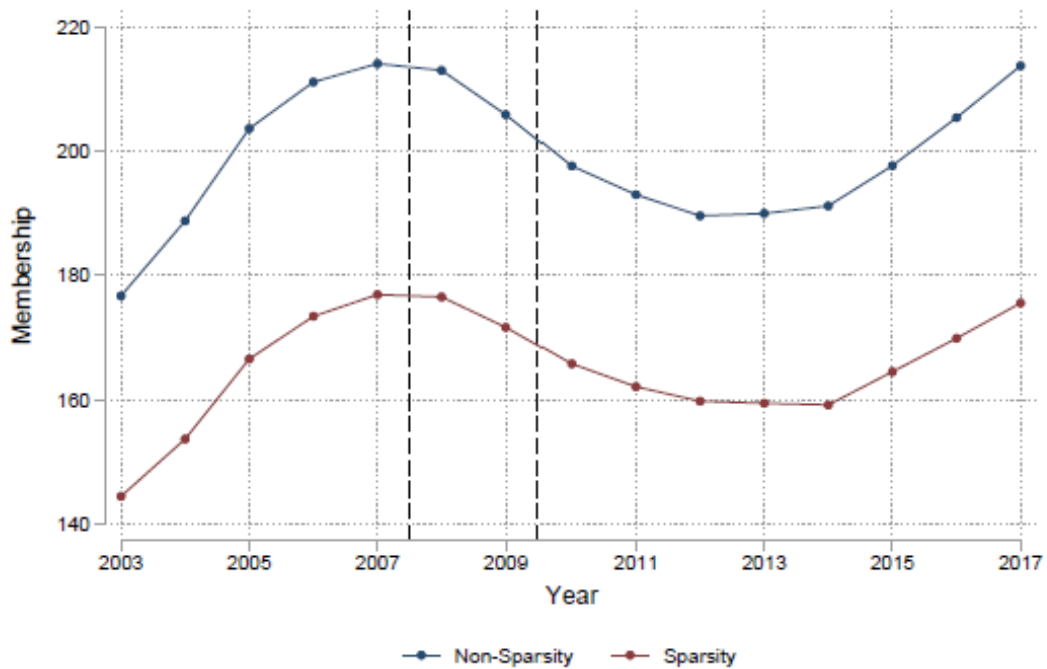
Notes: Panel A plots the average membership in sparsity-eligible and ineligible districts across academic years 2003-2017. Panel B repeats this plot measuring districts' membership relative to 2003.

Figure A3.3. Event Study Estimates of Sparsity Aid Program on District Demographics

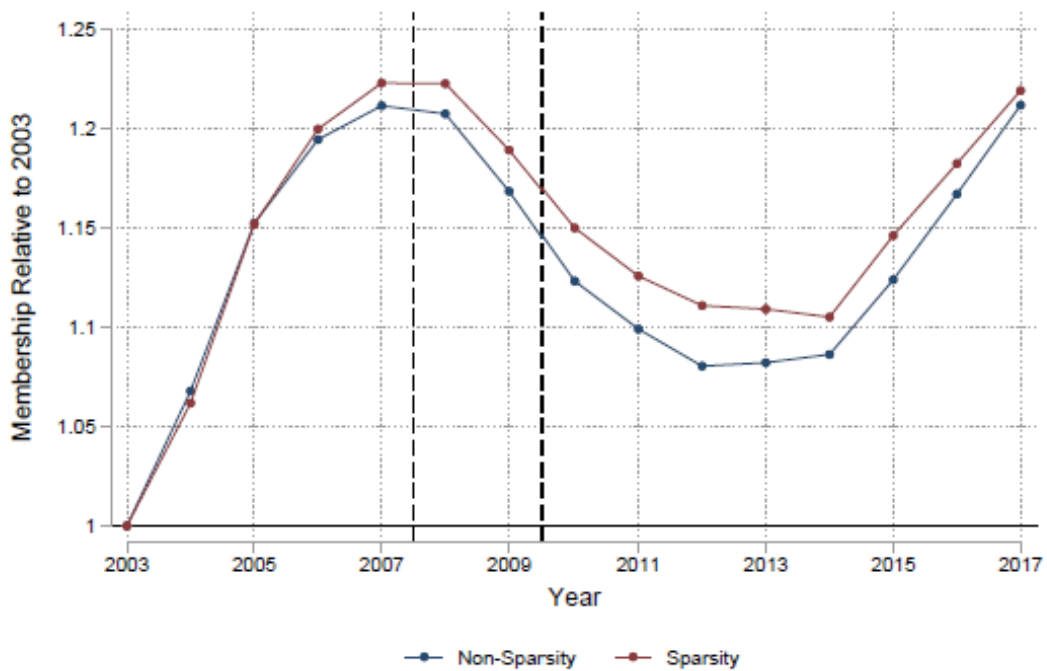


Notes: Each figure presents estimates of the β_k coefficients from equation (2), including either district and year FEs or district and year-by-region FEs. All standard errors are clustered at the school district level.

Figure A3.4. House Price Index in sparsity-eligible & Ineligible Districts, 2003-2017
(a) District-Level House Price Index

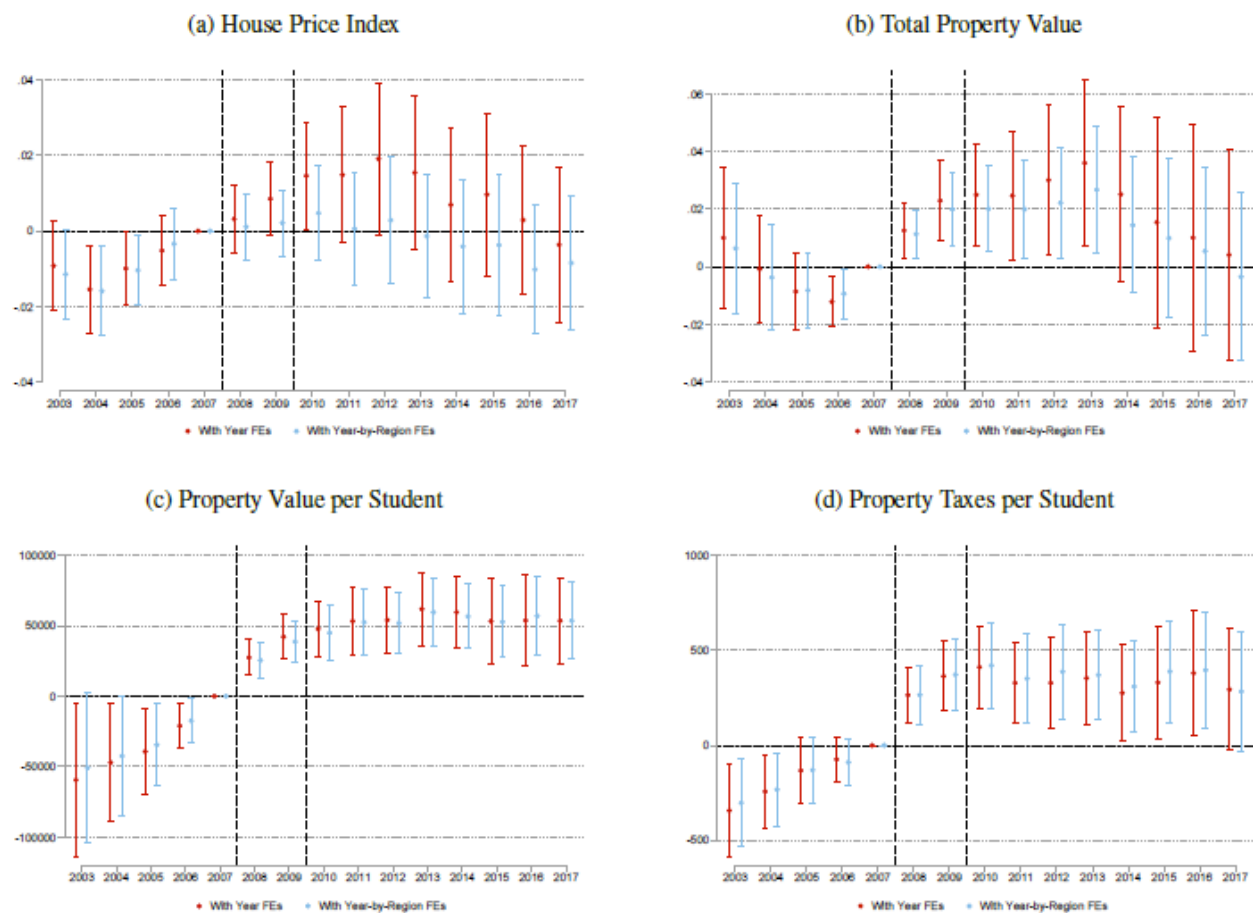


(b) District-Level House Price Index Relative to 2003



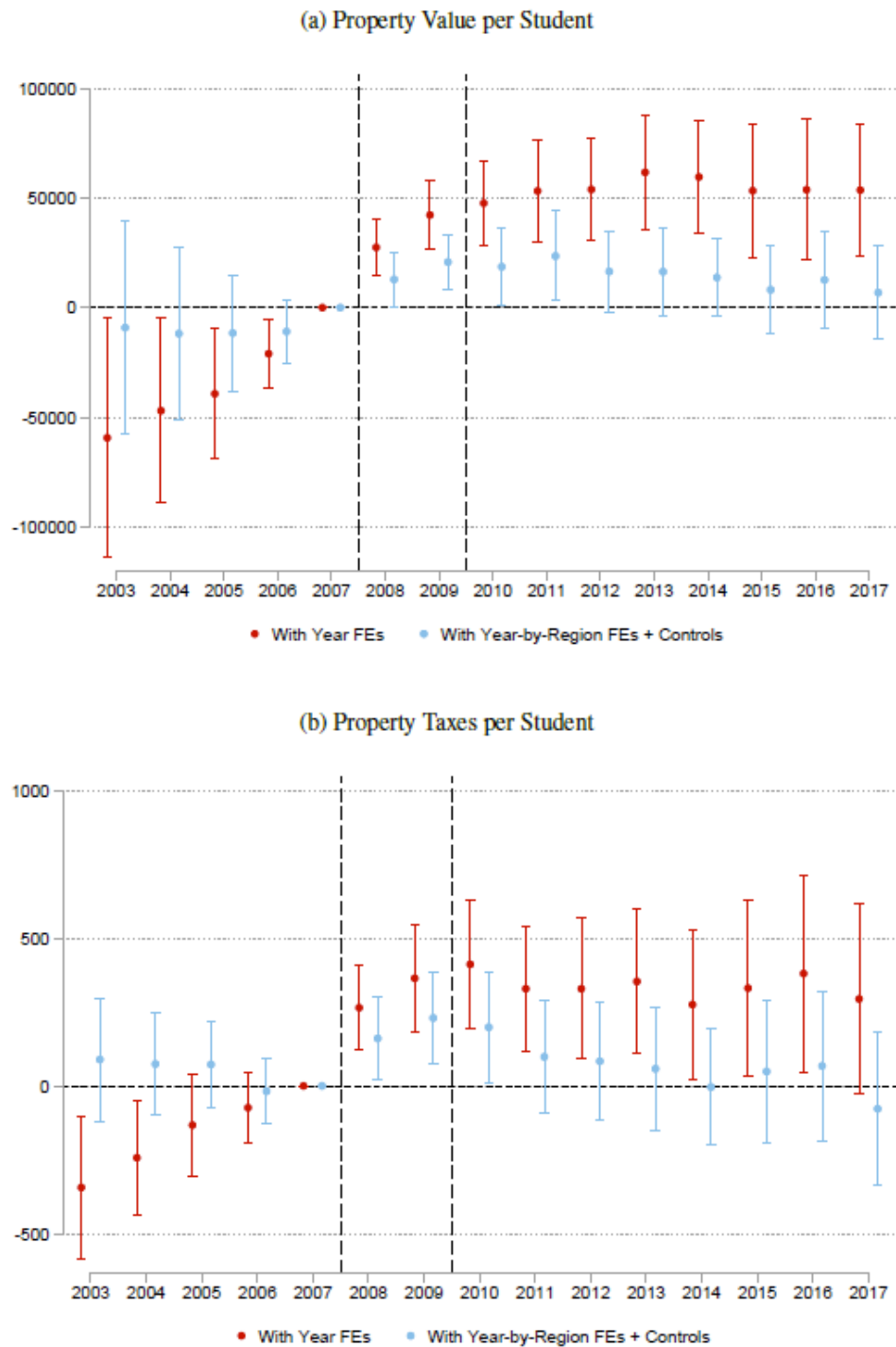
Notes: Panel A plots the average district-level house price index in sparsity-eligible and ineligible districts across academic years 2003-2017. Panel B repeats this plot measuring districts' house price indices relative to 2003.

Figure A3.5. Event Study Estimates of Sparsity Aid Program on Local Resources



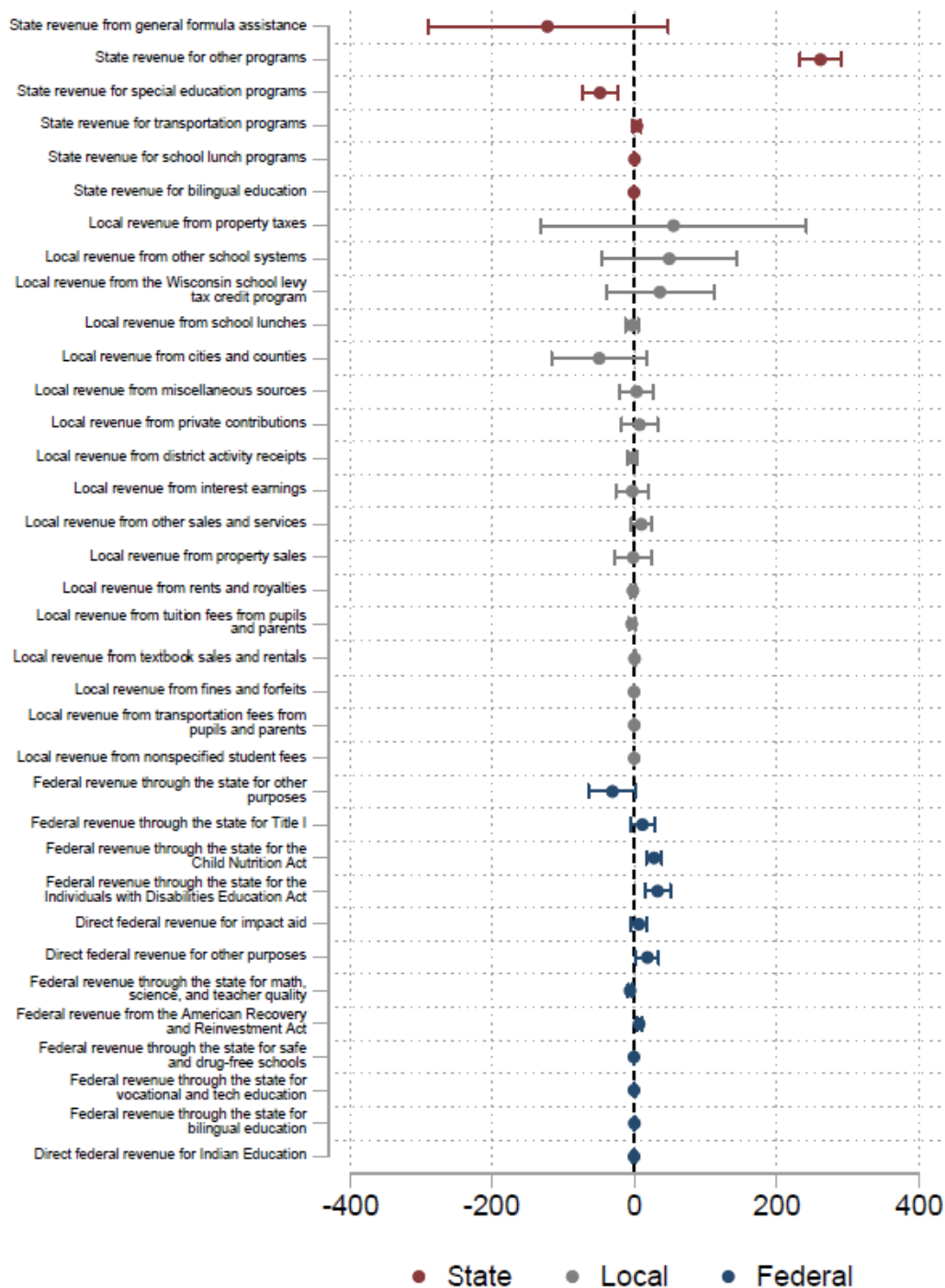
Notes: Each figure presents estimates of the β_k coefficients from equation (2), including either district and year FEs or district and year-by-region FEs. All standard errors are clustered at the school district level.

Figure A3.6. Event Study Estimates of Sparsity Aid Program on Local Resources, With Additional Controls



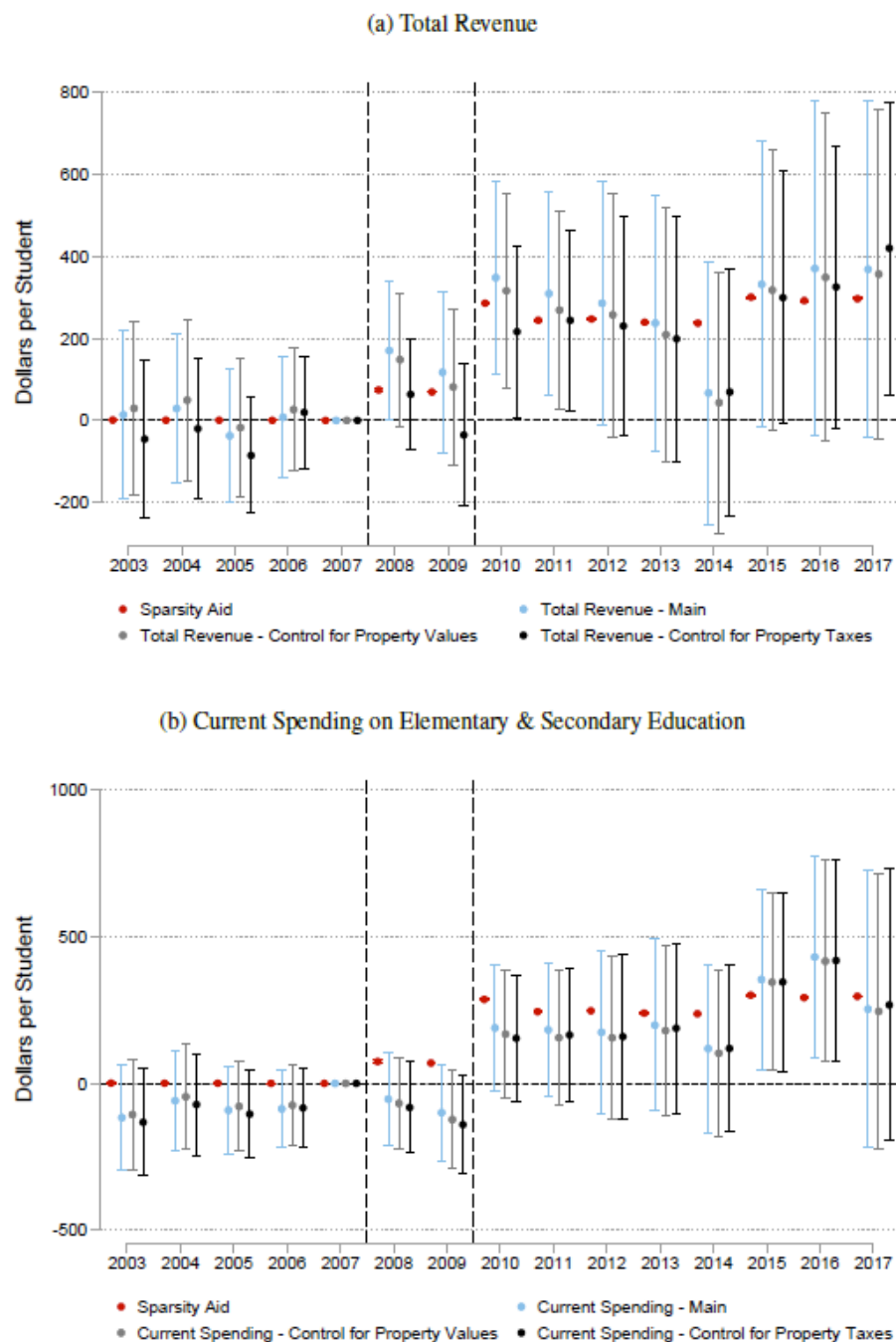
Notes: Each figure presents estimates of the β_k coefficients from equation (2), including either district and year FEs or district and year-by-region FEs, along with districts' log membership and log house price index. All standard errors are clustered at the school district level.

Figure A3.7. Effects of Sparsity Aid on District Revenues, by Source



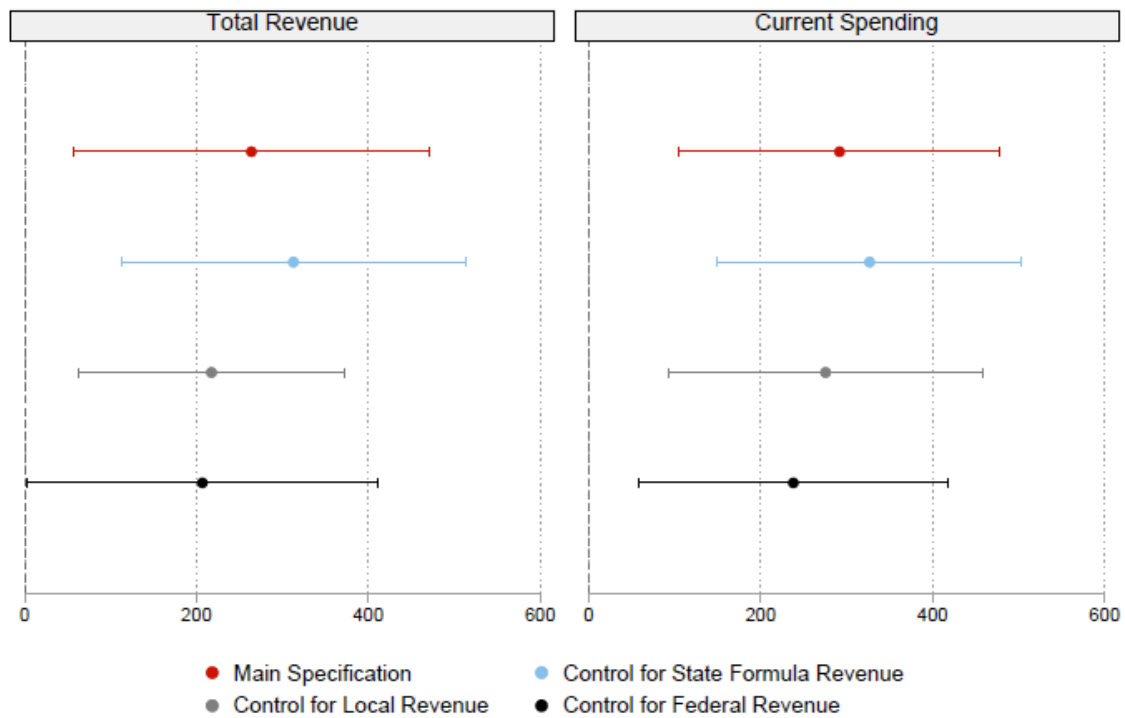
Notes: Each coefficient is estimated from a separate regression and represents β in equation (1), the effect of receiving sparsity aid funding. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level.

Figure A3.8. Event Study Estimates of Sparsity Aid Program on District Finances, With Additional Controls



Notes: Each figure presents estimates of the β_k coefficients from equation (2). All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. Additional specifications control for a district's total property value per student or total property tax revenue per student, as indicated. All standard errors are clustered at the school district level.

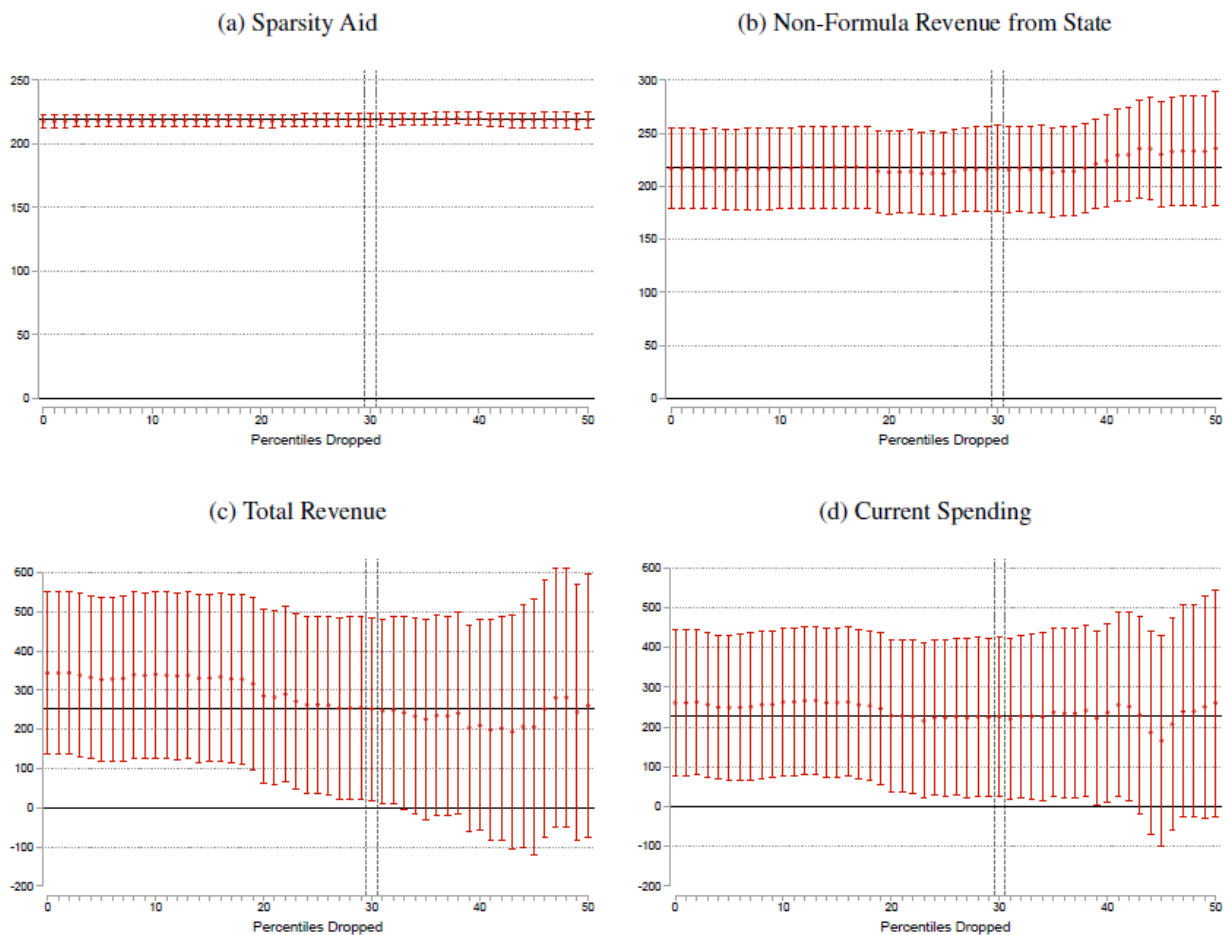
Figure A3.9. Effect of Sparsity Aid Program on District Finances, With Additional Revenue Controls



Notes: Each coefficient is estimated from a separate regression and represents β in equation (1), the effect of receiving sparsity aid funding. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. Additional specifications control for a district's state formula revenue per student, total local revenue per student, or total federal revenue per student, as indicated. All standard errors are clustered at the school district level.

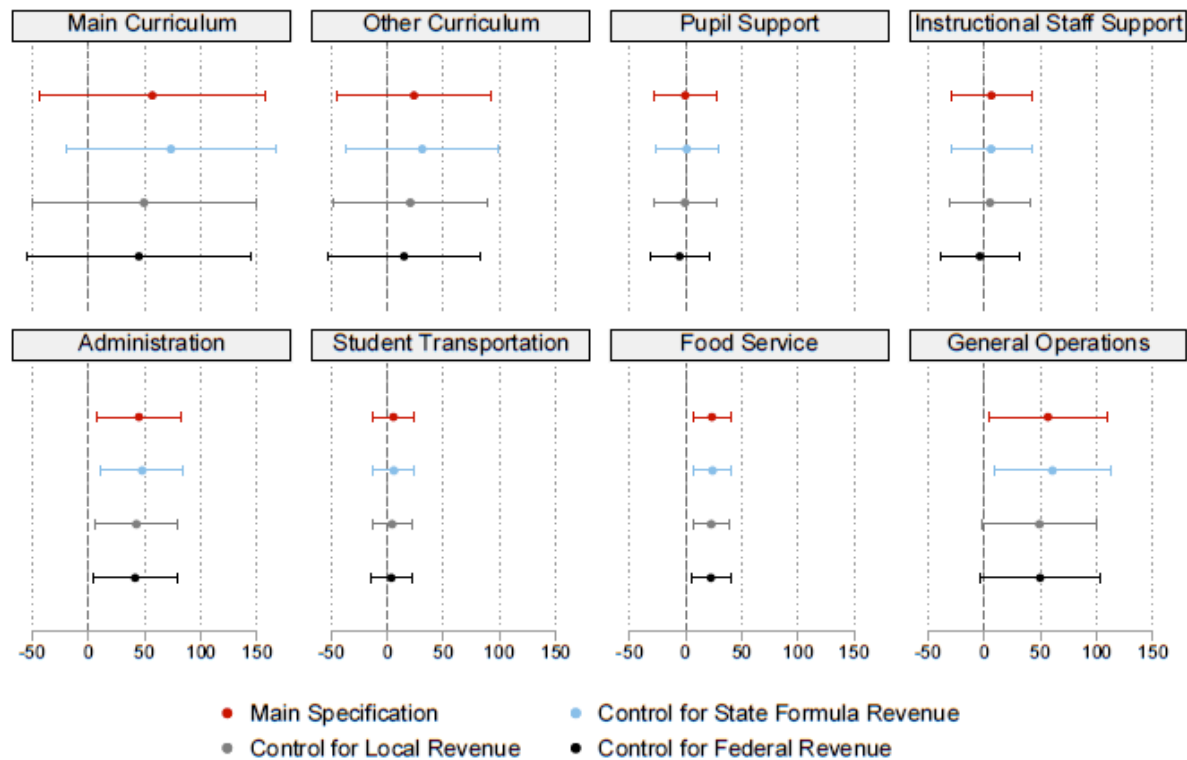
Figure A3.10. Sensitivity of School Finance Estimates to Density & Membership

Bandwidth Selection



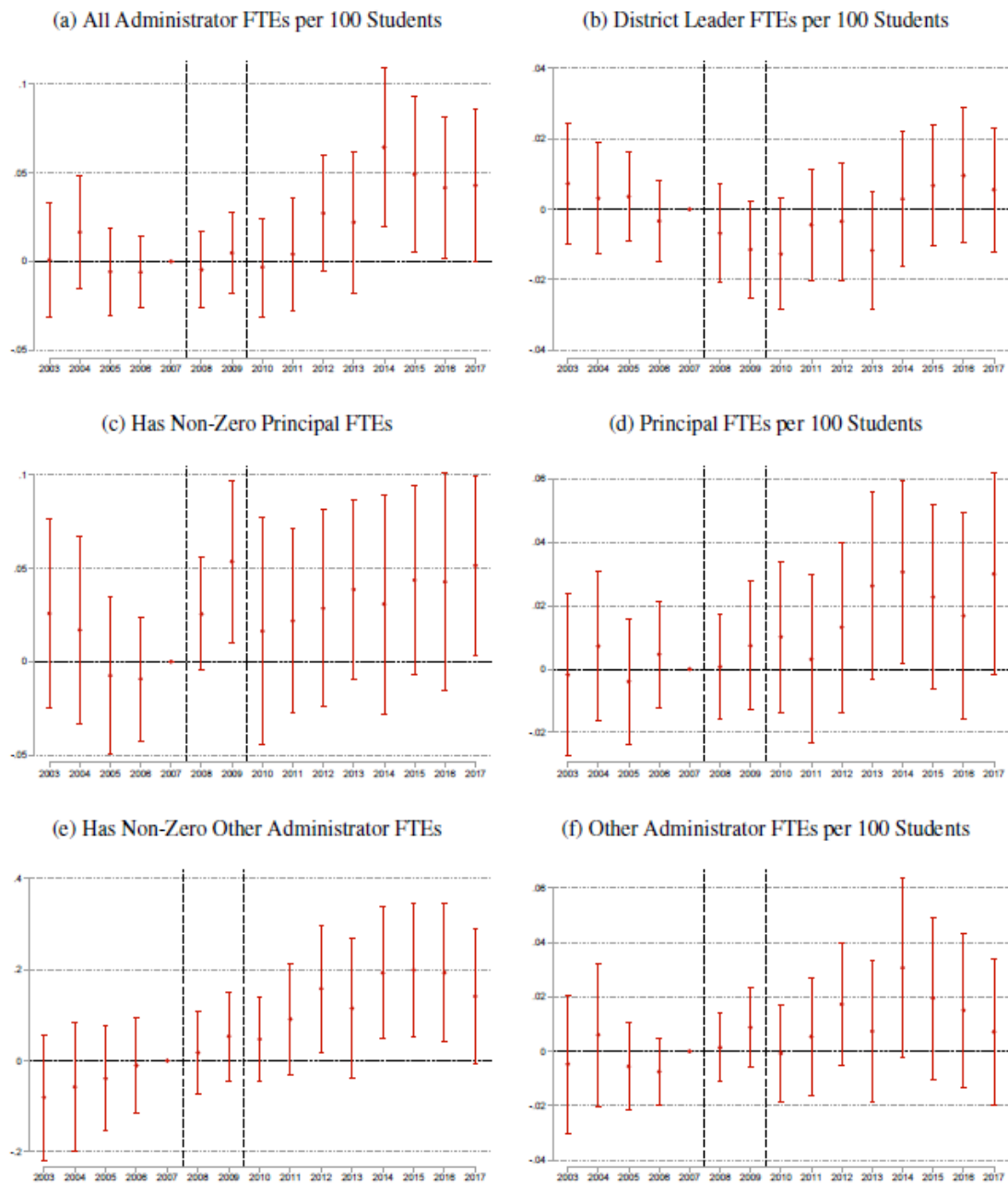
Notes: Each figure presents estimates of β in equation (1), the effect of receiving sparsity aid funding. Each coefficient is estimated from a separate regression, where we restrict the sample by dropping districts at the top of the pre-2008 density and membership distributions. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level. The grey dashed lines highlight our baseline bandwidth of the 30th percentile.

Figure A3.11. Effect of Sparsity Aid Program on Spending Allocations, With Additional Revenue Controls



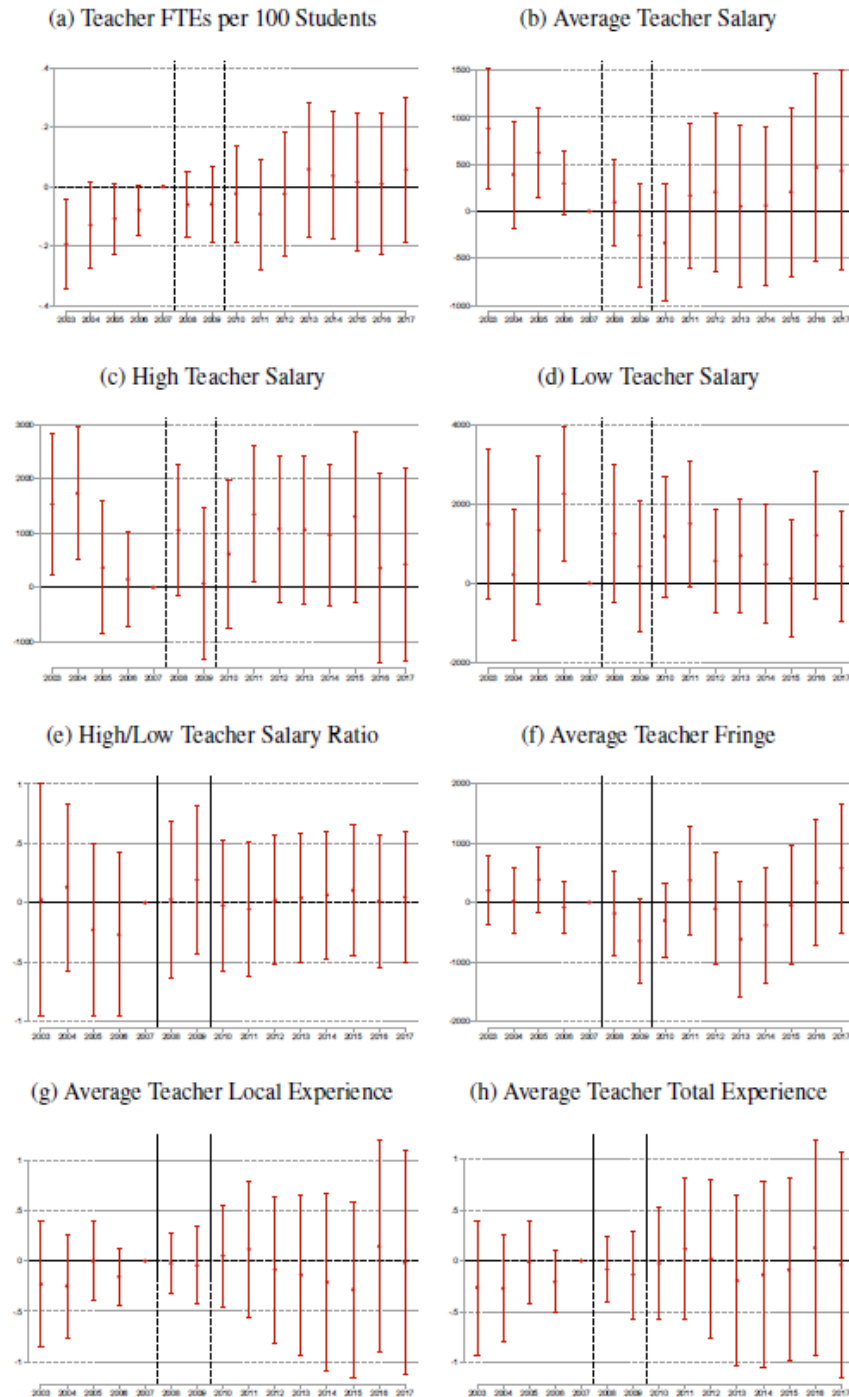
Notes: Each coefficient is estimated from a separate regression and represents β in equation (1), the effect of receiving sparsity aid funding. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. Additional specifications control for a district's state formula revenue per student, total local revenue per student, or total federal revenue per student, as indicated. All standard errors are clustered at the school district level.

Figure A3.12. Event Study Estimates of Sparsity Aid Program on Administrator Staffing



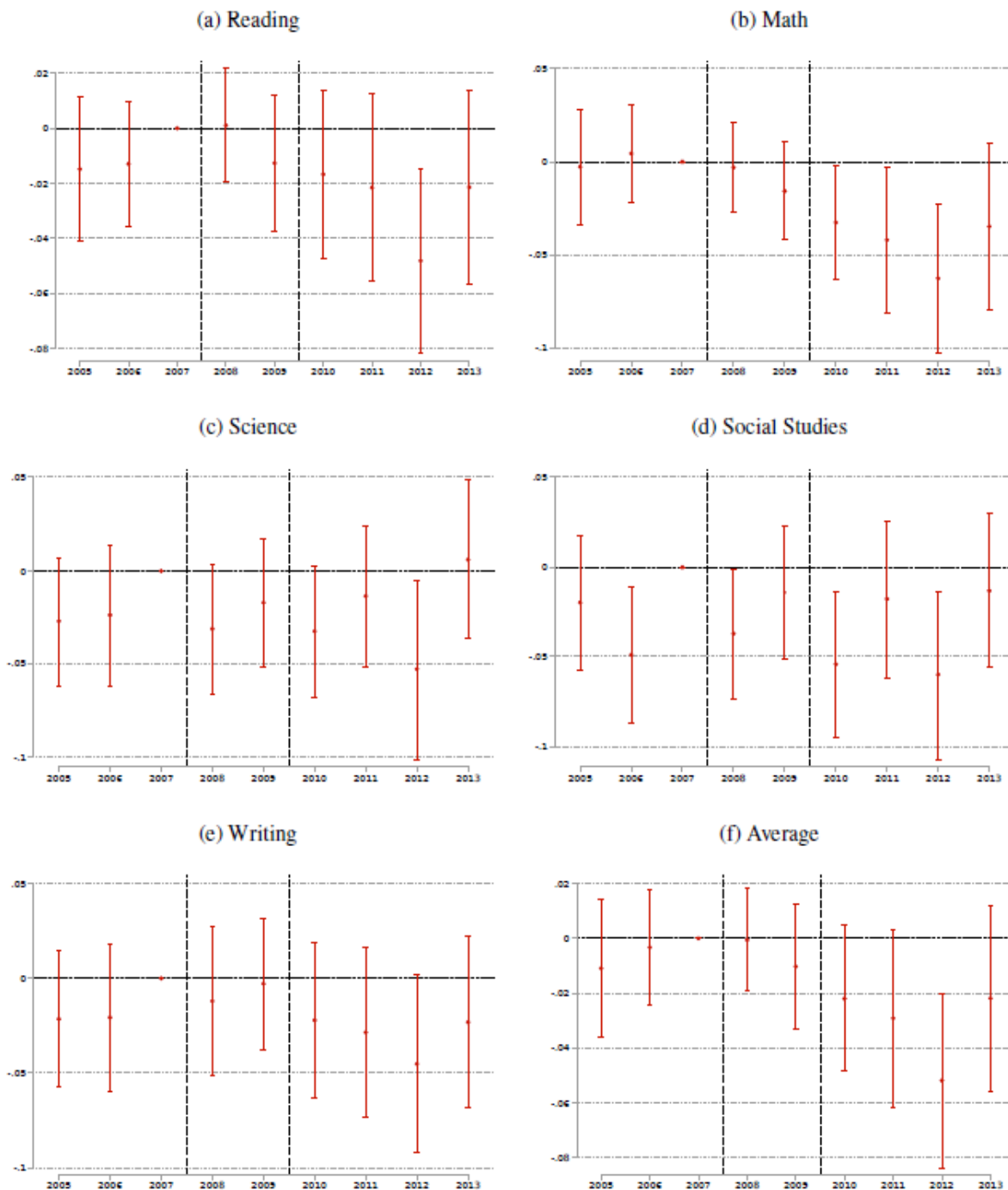
Notes: Each figure presents estimates of the β_k coefficients from equation (2). All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level.

Figure A3.13. Event Study Estimates of Sparsity Aid Program on Teacher Staffing



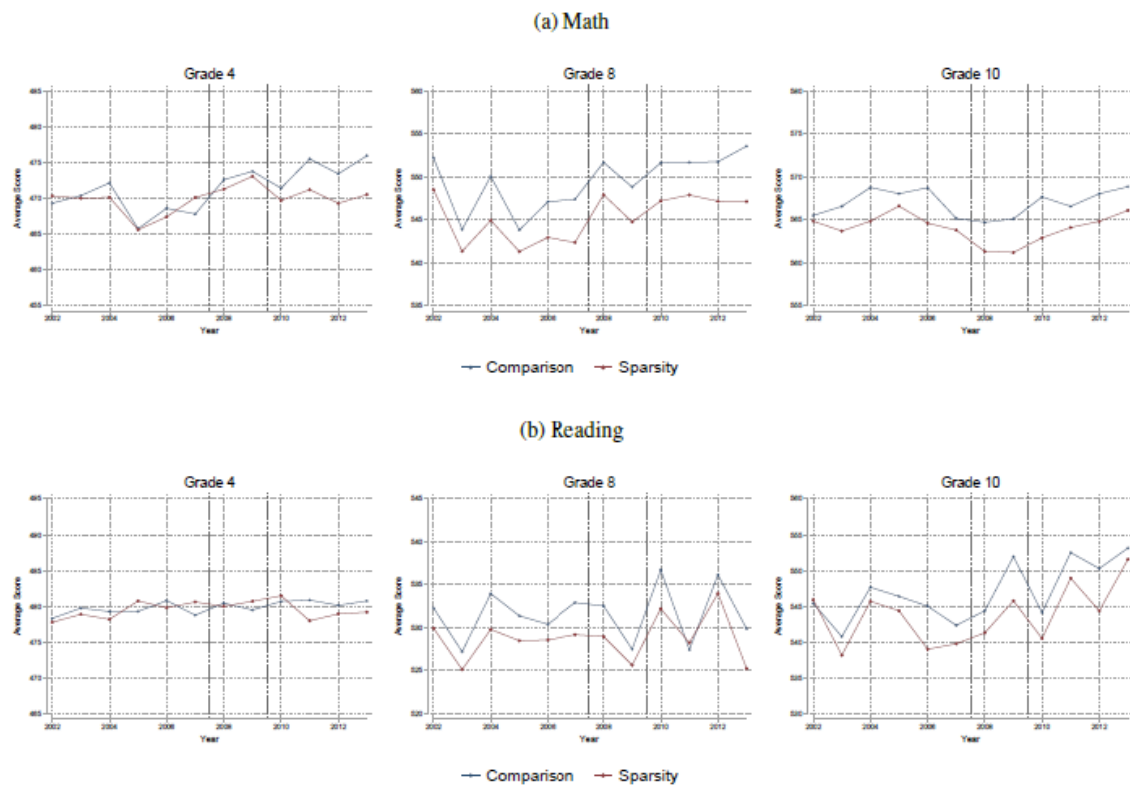
Notes: Each figure presents estimates of the β_k coefficients from equation (2). All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level.

Figure A3.14. Event Study Estimates of Sparsity Aid Program on Test Scores



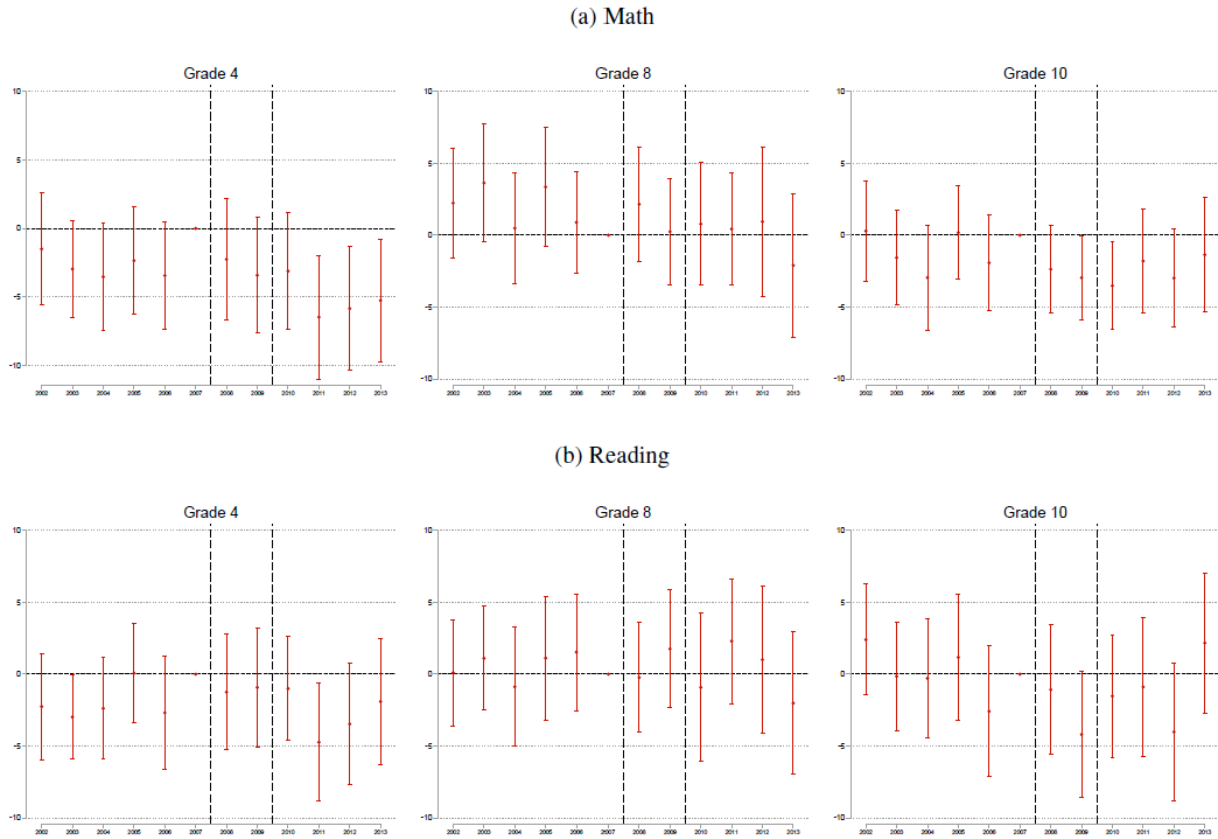
Notes: Each figure presents event study estimates for student-level outcomes. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level.

Figure A3.15. School-Level Test Scores, 2002-2013



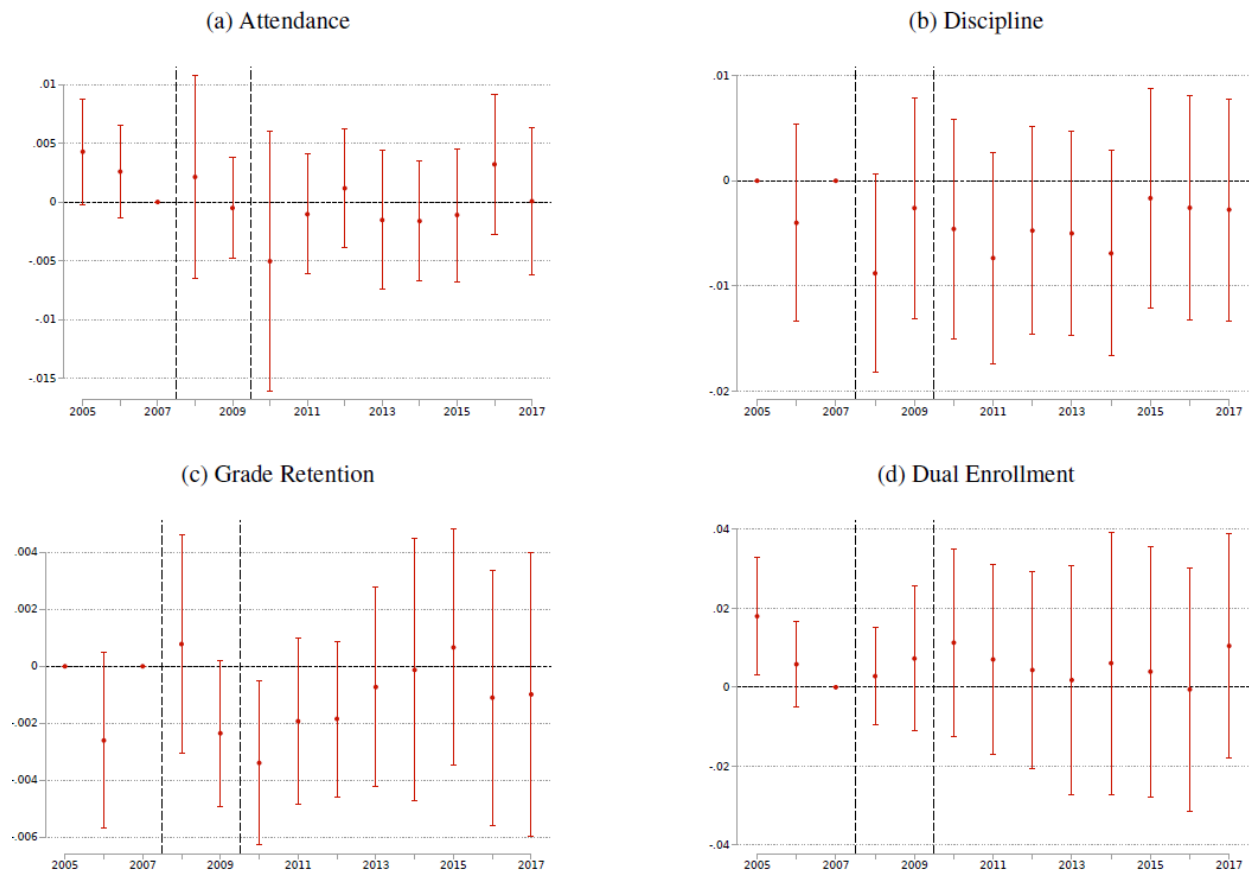
Notes: Panel A plots the average school-level math test scores in sparsity-eligible and ineligible districts across academic years 2002-2013 and grades 4, 8, and 10. Panel B repeats this plot measuring school-level reading test scores.

Figure A3.16. Event Study Estimates of Sparsity Aid Program on School-Level Test Scores



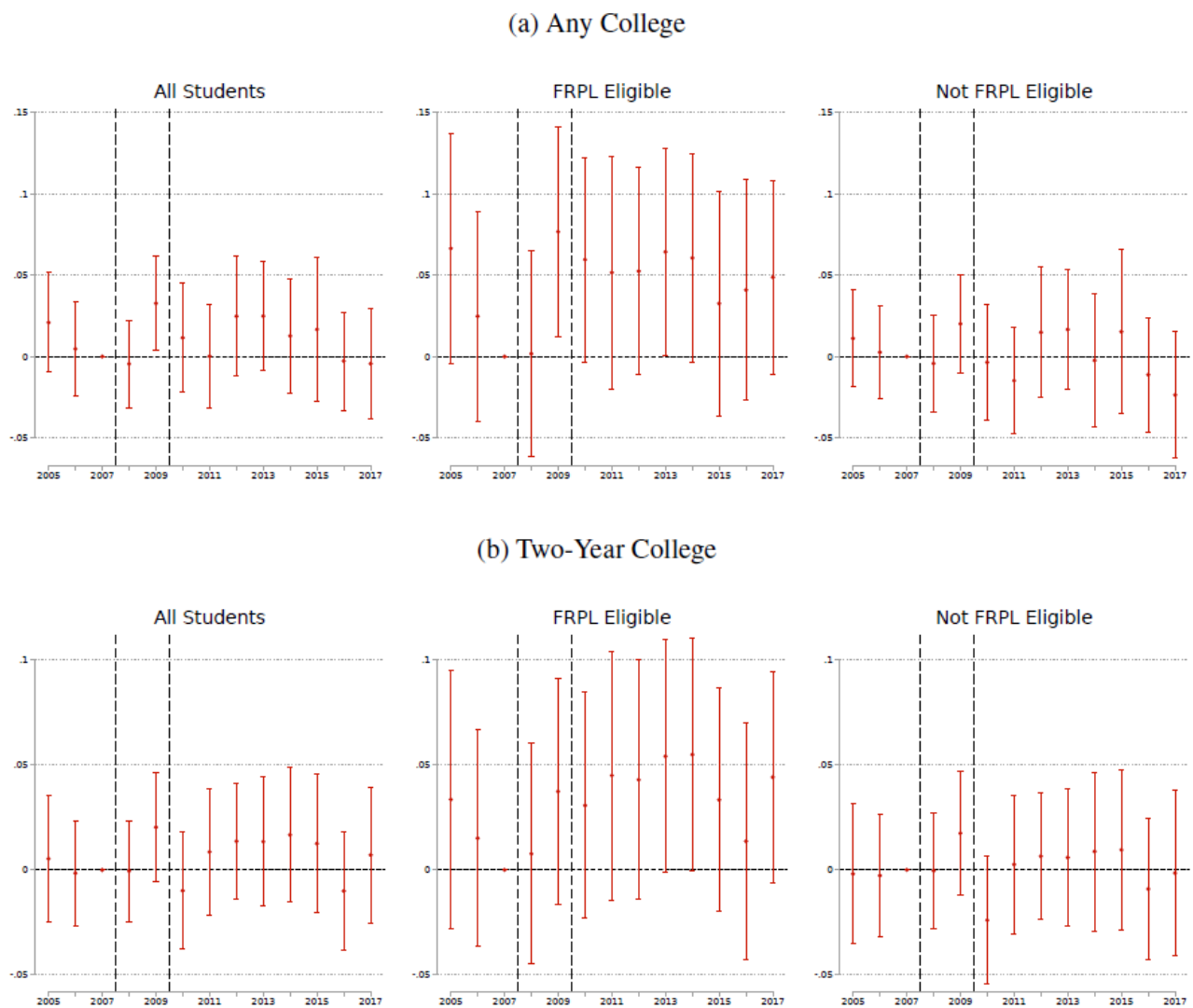
Notes: Each figure presents event study estimates for school-level outcomes. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level.

Figure A3.17. Event Study Estimates of Sparsity Aid Program on Behavioral Outcomes



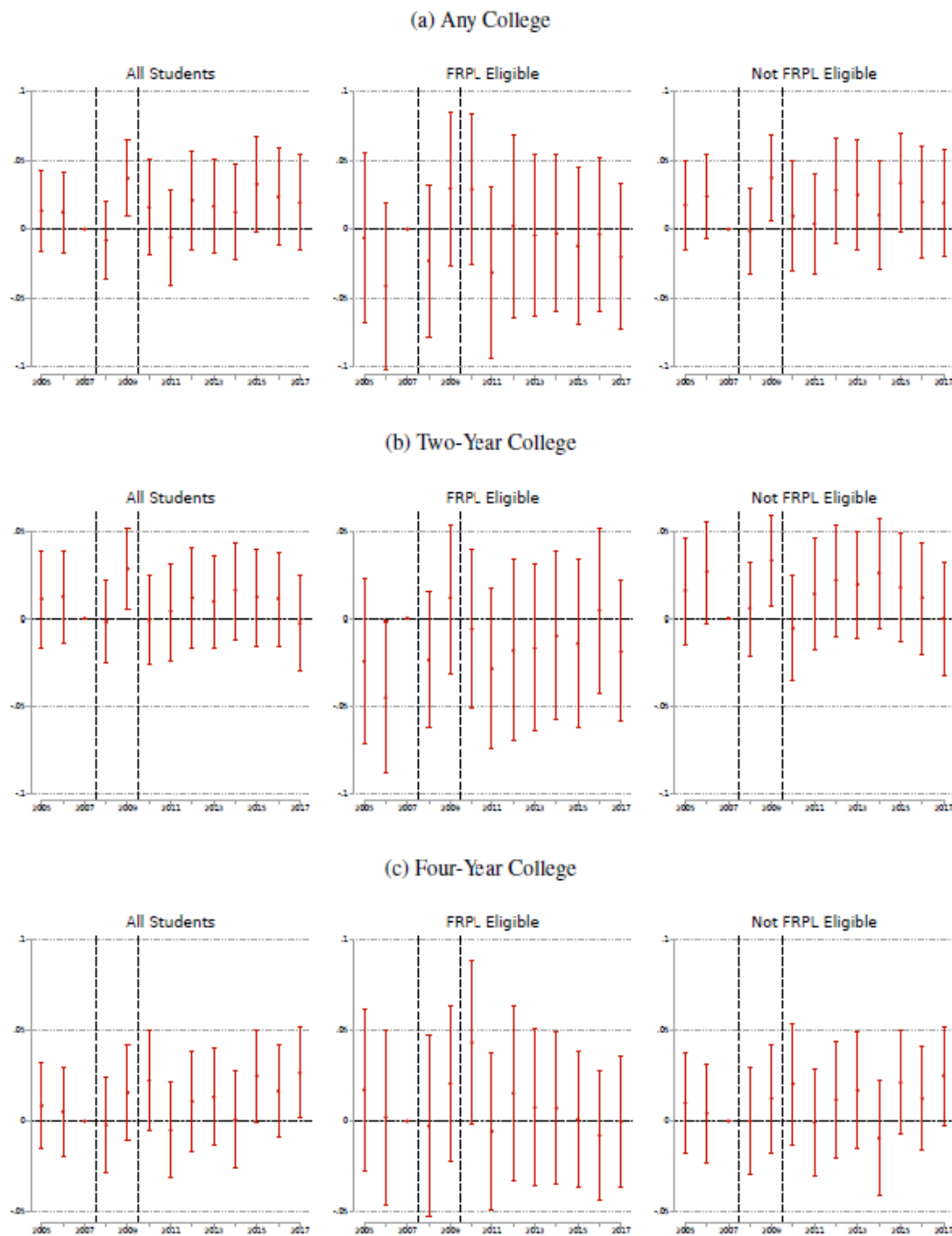
Notes: Each figure presents event study estimates for student-level outcomes. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level.

Figure A3.18. Event Study Estimates of Sparsity Aid Program on Postsecondary Enrollment



Notes: Each figure presents event study estimates for student-level outcomes. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level.

Figure A3.19. Event Study Estimates of Sparsity Aid Program on Postsecondary Completion



Notes: Each figure presents event study estimates for student-level outcomes. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, and the local child poverty rate, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level.

Table A3.1. Effect of Sparsity Aid on Placebo Finance Outcomes

	(1)	(2)	(3)	(4)
<i>Panel A. Revenue other than state non-formula</i>				
Received sparsity aid	7.91 (108.4)	4.13 (101.2)	27.72 (117.5)	35.31 (117.1)
Observations	2,820	2,820	2,820	2,820
<i>Panel B. Spending other than elementary & secondary education</i>				
Received sparsity aid	58.41 (161.01)	64.51 (163.38)	73.74 (183.64)	77.53 (188.18)
Observations	2,820	2,820	2,820	2,820
Membership control	X	X	X	X
Demographic controls		X	X	X
Transportation funding control				X
Year-by-CESA FEs			X	X

Notes: Each coefficient is estimated from a separate regression and represents β in equation (1), the effect of receiving sparsity aid funding. Column (1) controls for a district's log membership, column (2) adds controls for a district's log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, and the local child poverty rate, column (3) further controls for whether a district receives additional transportation funding, and column (4) adds year-by-region (CESA) fixed effects. All standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3.2. Effects of Sparsity Aid on Teacher Staffing

	FTEs per 100 Students	Avg. Salary	High Salary	Low Salary	Ratio of High/Low	Avg. Fringe	Avg. Local Experience	Avg. Total Experience
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Main Specification</i>								
Received sparsity aid	0.085 (0.082)	-337.4 (325.9)	98.16 (476.4)	-243.70 (461.4)	0.115 (0.158)	-220.6 (315.3)	0.080 (0.313)	0.103 (0.328)
Observations	2,820	2,820	2,820	2,820	2,820	2,820	2,820	2,820
<i>Panel B. Interaction with Baseline General Instruction Budget Share</i>								
Received sparsity aid	0.066 (0.080)	-338.2 (328.0)	28.29 (488.1)	-294.0 (462.7)	0.132 (0.160)	-226.9 (312.0)	0.125 (0.314)	0.162 (0.331)
Received sparsity aid x budget share	-0.041** (0.020)	-1.906 (68.45)	-149.8* (77.09)	-107.8 (91.28)	0.037 (0.032)	-13.48 (67.87)	0.095 (0.074)	0.126* (0.066)
Observations	2,820	2,820	2,820	2,820	2,820	2,820	2,820	2,820
Effect at 25th Percentile	0.160*	-333.8	373.03	-45.86	0.047	-195.8	-0.094	-0.129
Effect at 75th Percentile	-0.006	-341.5	-234.7	-483.2	0.198	-250.5	0.291	0.384

Notes: The coefficients in each column are estimated from a separate regression and represent β in equation (1), the effect of receiving sparsity aid funding, with interaction effects based on districts' pre-2008 budget shares. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects. All standard errors are clustered at the school district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3.3. Effect of Sparsity Aid on Behavioral Outcomes

	6th Grade (1)	7th Grade (2)	8th Grade (3)	9th Grade (4)	10th Grade (5)	11th Grade (6)	12th Grade (7)	All Grades (8)
Panel A. Attendance								
Received sparsity aid	0.001 (0.002)	-0.001 (0.002)	-0.000 (0.002)	-0.003 (0.003)	-0.004 (0.003)	-0.005 (0.003)	-0.005 (0.003)	-0.003 (0.002)
Observations	140,113	143,606	146,055	156,705	158,689	161,880	165,120	1,072,168
Mean	0.960	0.957	0.954	0.953	0.947	0.940	0.934	0.949
Panel B. Disciplinary Incidence								
Received sparsity aid	-0.005 (0.004)	-0.007 (0.005)	-0.006 (0.006)	-0.002 (0.005)	0.004 (0.005)	-0.003 (0.006)	-0.004 (0.005)	-0.003 (0.004)
Observations	128,662	131,674	133,521	142,583	144,734	148,048	151,400	980,622
Mean	0.018	0.027	0.033	0.039	0.042	0.039	0.030	0.033
Panel C. Grade Retention								
Received sparsity aid	0.001 (0.001)	-0.004 (0.002)	-0.000 (0.001)	0.001 (0.004)	0.003 (0.002)	-0.001 (0.002)	0.001 (0.004)	0.000 (0.001)
Observations	125,457	128,167	131,059	134,258	142,107	143,975	147,888	952,911
Mean	0.001	0.003	0.002	0.009	0.004	0.008	0.031	0.009
Panel D. Dual Enrollment								
Received sparsity aid					-0.005 (0.004)	-0.005 (0.014)	0.003 (0.013)	-0.002 (0.009)
Observations					158,779	162,010	165,442	486,231
Mean					0.008	0.155	0.115	0.042

Notes: Each coefficient is estimated from a separate regression and represents β in equation (3), the effect of receiving sparsity aid funding. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3.4. Effect of Sparsity Aid on Postsecondary Enrollment & Completion, 2005-2013 Sample

	Enrollment			Completion		
	Any (1)	Two-Year (2)	Four-Year (3)	Any (4)	Two-Year (5)	Four-Year (6)
<i>Panel A. All Students</i>						
Received sparsity aid	0.010 (0.010)	0.008 (0.008)	0.005 (0.007)	0.006 (0.009)	0.001 (0.007)	0.006 (0.008)
Observations	119,730	119,730	119,730	119,730	119,730	119,730
Mean	0.552	0.237	0.335	0.416	0.174	0.272
<i>Panel B. FRL Eligible Students</i>						
Received sparsity aid	0.029 (0.020)	0.021 (0.019)	0.011 (0.013)	0.022 (0.014)	0.016 (0.012)	0.007 (0.013)
Observations	24,435	24,435	24,435	24,435	24,435	24,435
Mean	0.374	0.204	0.182	0.256	0.142	0.129
<i>Panel C. FRL Ineligible Students</i>						
Received sparsity aid	0.004 (0.010)	0.005 (0.008)	0.002 (0.009)	0.003 (0.011)	0.000 (0.008)	0.006 (0.009)
Observations	95,291	95,291	95,291	95,291	95,291	95,291
Mean	0.598	0.246	0.374	0.456	0.183	0.309

Notes: Each coefficient is estimated from a separate regression and represents β in equation (3), the effect of receiving sparsity aid funding. All specifications control for a district's log membership, log house price index, number of school buildings, racial composition (% white, % Black, % Hispanic, and % Asian), % FRL, % special education, the local child poverty rate, a dummy variable indicating whether a district receives funding from the state's high-cost pupil transportation aid program, and year-by-region (CESA) fixed effects, along with student-level race, gender, FRL, special education, and limited English proficiency indicators. All standard errors are clustered at the school district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3.1. Email to Wisconsin Association of School District Administrators (WASDA)

Subject: Interview request for research on Wisconsin rural school districts

Dear [Insert Name],

I hope this email finds you well. My name is Dr. Riley Acton and I am an Assistant Professor of Economics at Miami University. I research the economics of education and education policy and am currently working with a team of researchers to study Wisconsin's sparsity aid program. Our work is funded by the Bill & Melinda Gates Foundation and aims to develop a richer understanding of the various challenges that rural school districts face, as well as the ways in which the sparsity aid program helped districts address these challenges. Our ultimate goal is to inform state and district policymakers on a national scale about how state investments in rural schools affect students, districts, and communities.

My research team and I would love to have an opportunity to speak with you or one of your WASDA colleagues about the history of the sparsity aid program and how it has been perceived among school district administrators. Is there a time in the coming weeks when you would be available for a phone or Zoom call? I am cc'ing my collaborator, Salem Rogers, who can coordinate with you or someone from your office to find a time that is most convenient.

Thank you for your consideration, and we look forward to hearing back from you!

All the best,

Riley Acton, Ph.D.

B.3.2. B.2 Survey Text

Sparsity Aid Research Survey Research Consent Information:

You are invited to participate in a research project being conducted by Dr. Riley Acton from Miami University. The purpose of this research is to examine the unique challenges of small, rural school districts and how Wisconsin's sparsity aid program has helped districts address these challenges. Participation in this research is restricted to persons 18 years of age or older.

Completing the survey should take about 10 minutes. Your participation is voluntary, you may skip any questions you do not want to answer, and you may stop at any time. Foreseeable risks and/or discomforts associated with your participation are minimal and you will receive no direct benefit from your participation. However, we hope our study will benefit Wisconsin students, district leaders, and national education policymakers by uncovering how school districts and policymakers can effectively support student success with increased funding, specifically in the rural context. To achieve this goal, we plan on broadly disseminating our findings to the academic community and interested parties in the general public.

Only the research team will have access to individual responses and we will not attribute the name of your school district to any of your answers in any presentations or publications without first receiving your permission, in writing, to do so. Unless you provide this permission, the results of the research will be presented publicly only as aggregate summaries. The research data will be retained until June 30, 2027.

Funding agencies or journal policies may require that individual participant data be made available to other researchers. Sharing data in this way advances the field by allowing the data to be used beyond this study. No personally identifying information will be included in the shared

data.

Our research team is not associated with Wisconsin's Department of Public Instruction, and we have no conflicts of interest to disclose. The research project is supported by a non-renewable grant from the Bill & Melinda Gates Foundation (INV-036567).

If you have any questions about this research or you feel you need more information to determine whether you would like to volunteer, you can contact the principal investigator (PI), Dr. Riley Acton, at actonr@miamioh.edu or at (513) 529-2865. If you have questions or concerns about the rights of research subjects, you may contact the Miami University Research Ethics and Integrity Office at (513) 529-3600 or humansubjects@miamioh.edu.

Please keep a copy of this information for future reference.

1. Do you consent to participate in this study?
 - I consent.
 - I do not consent.
2. Have you ever worked for a school district that received funding from Wisconsin's Sparsity Aid program?
 - Yes
 - No
 - I don't know

General Background: For these questions, please think about the most recent school district that you worked for that received sparsity aid funding. This can be your current school district.

3. What school district did you work for?
4. What title best described your highest position in this district
 - District Administrator/Superintendent

- District Treasurer/Business Official
- School Board Member
- Teacher
- Other [open-ended]

5. During what years did you work for this district? Please select all that apply.

- Before 2008
- 2008
- 2009
- ...
- 2020
- 2021
- 2022/Present

Sparsity Aid Funds: For these questions, please continue to think about the most recent school district that you worked for that received sparsity aid funding. This can be your current school district.

6. When your district received sparsity aid funding, how often were the funds set aside for a specific purpose?

- Never
- Rarely
- Often
- Always
- I don't know

7. To the best of your knowledge, for what purposes were sparsity aid funds used? Please select all that apply.

- Instruction in core academic subjects (e.g., math, reading)
- Supplemental instruction (e.g., tutoring)
- Electives (e.g., art, music) or co-curricular activities (e.g., athletics, speech & debate)
- Pupil Support (e.g., guidance, health, social work)
- Instructional Staff Support (e.g., curriculum development, training)
- Administration (e.g., general district administration, school building administration)
- Operations and/or maintenance (e.g., site and building repairs)
- Pupil Transportation
- Food Service
- I don't know

8. To the best of your knowledge, were sparsity aid funds spent on specific grade levels? Please select all that apply.

- Elementary
- Middle/junior high
- High school
- District-wide
- I don't know

Perceptions: For these questions, please continue to think about the most recent school district that you worked for that received sparsity aid funding. This can be your current school district.

9. Thinking back to the early years of the sparsity aid program (2008-2010), how likely did

you think it was that the program would continue long-term?

- Very unlikely
- Unlikely
- Likely
- Very likely
- Do not remember
- Was not aware of the sparsity aid program in 2008-2010

10. If all other funding sources were held constant, but your district no longer had access to sparsity aid funding, how likely do you think each of the following scenarios would be?

[Options: Very Unlikely, Unlikely, Likely, Very likely, No opinion]

- My district would employ fewer staff members.
- Staff retention in my district would worsen.
- Student achievement (e.g., test scores) in my district would decline.
- Graduation rates in my district would decline.
- Fewer students in my district would pursue postsecondary education.
- My district would consolidate with a neighboring district.
- My district would implement a four-day school week.

11. In your opinion, how has the sparsity aid program most affected your district, staff, and students since it began in 2008? Please include as much detail as you are able. [Open-ended response]

B.3.3. Respondent Characteristics

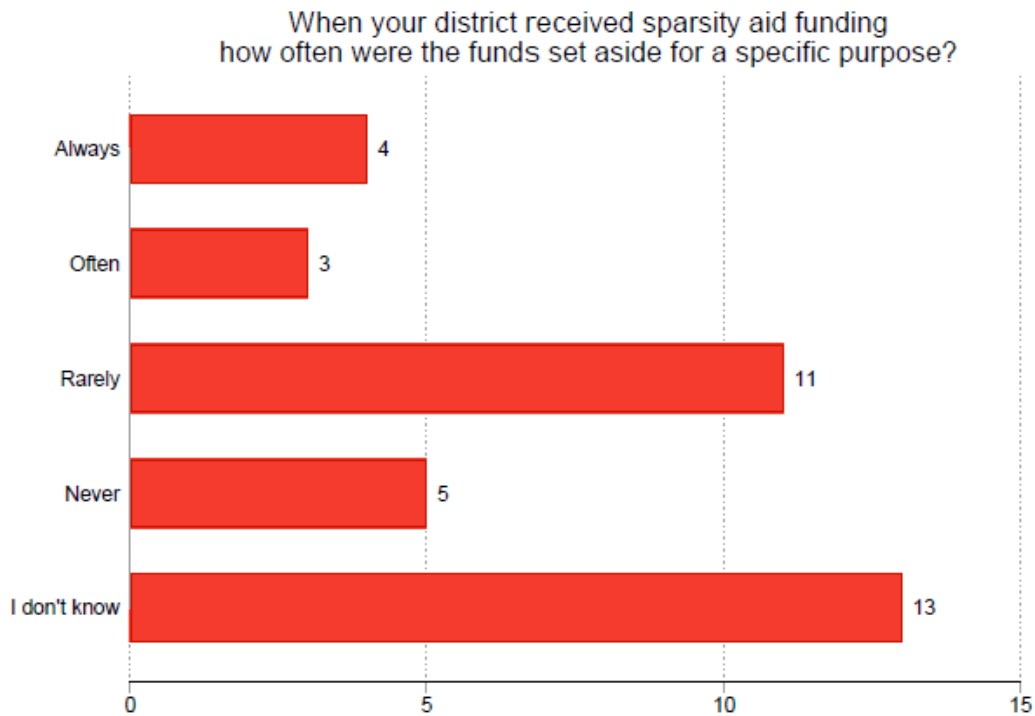
Table B.3.1. Survey Respondent Characteristics at Baseline

	Respondents (1)	Non-respondents (2)	p-value (3)
<i>Panel A. Size & Location</i>			
Membership	468.2	477.0	0.708
Membership per Square Mile	3.777	3.831	0.567
Number of Schools	2.553	2.465	0.135
Avg. Membership per School	195.1	207.7	0.237
NCES Rural Classification	1.000	0.968	0.024
<i>Panel B. Demographics</i>			
% White	0.918	0.952	0.000
% FRPL	0.326	0.345	0.563
% Special Education	0.153	0.159	0.027
Local Child Poverty Rate	0.133	0.148	0.012
District House Price Index	172.8	164.2	0.377
<i>Panel C. Finances</i>			
Revenue per student	12,638	12,517	0.435
Spending per student	10,460	10,392	0.980
% Instruction	0.685	0.668	0.561
% Support	0.064	0.065	0.890
% Administration	0.072	0.073	0.151
% Other	0.178	0.183	0.759
<i>Panel D. Staffing</i>			
Number of Teachers (FTE)	36.568	37.575	0.417
Teachers per 100 Members	8.099	8.053	0.558
Average Teacher Salary	40,610	40,446	0.932
Average Teacher Experience	15.97	15.92	0.729
Number of Administrators (FTE)	2.378	2.298	0.256
Administrators per 100 Members	0.542	0.496	0.042
<i>Panel E. Educational Outcomes</i>			
Math Proficiency Rate	0.403	0.389	0.673
Reading Proficiency Rate	0.342	0.321	0.182
College Enrollment Rate	0.532	0.515	0.759
College Completion Rate	0.398	0.386	0.989
Observations	170	310	465
Districts	34	62	96

Notes: Each column summarizes district-level characteristics over the 2003-2007 academic years. The test score and postsecondary outcomes are averaged over the 2005-2007 academic years. The college enrollment rate is defined as the share of high school seniors who enroll in a postsecondary institution within one year of graduating from high school and the college completion rate is defined as the share of high school seniors in the 2005-2007 cohorts who completed a postsecondary credential within the time frame of our data. The sample of respondent districts does not adhere to the same sample restrictions as in our analytic sample, so the total district count exceeds that from Table 3.1.

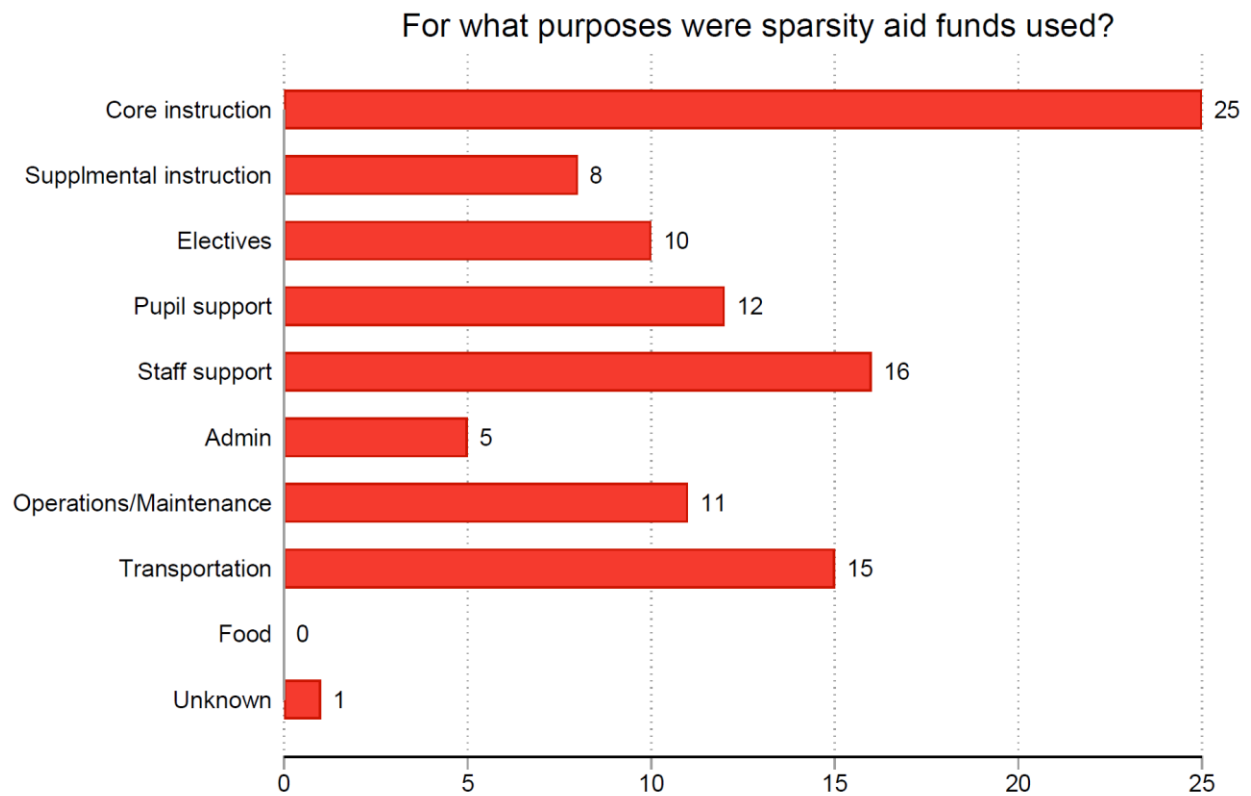
B.4. Survey Results

Figure B.3.1. Distribution of responses for Q6



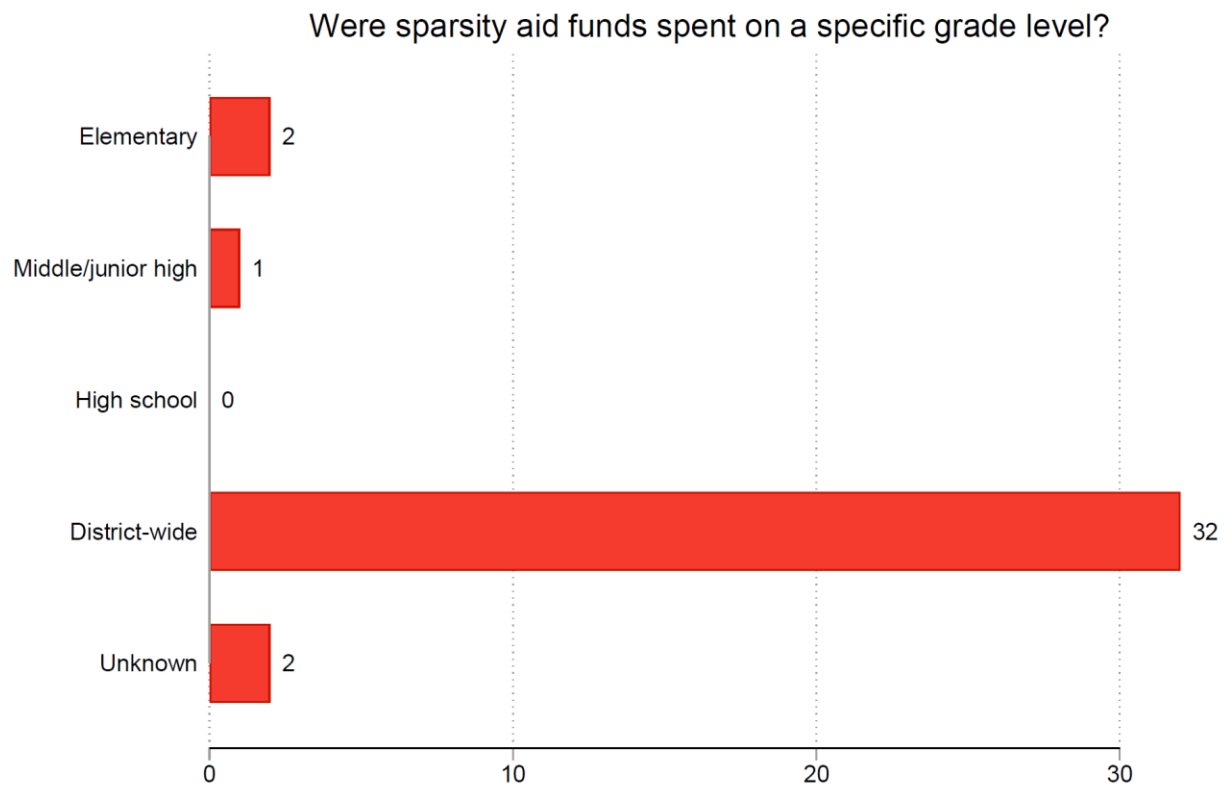
Notes: Each bar represents the number of survey respondents (out of 36) who selected the corresponding answer to, “When your district received sparsity aid funding, how often were the funds set aside for a specific purpose?”.

Figure B.3.2. Distribution of responses for Q7



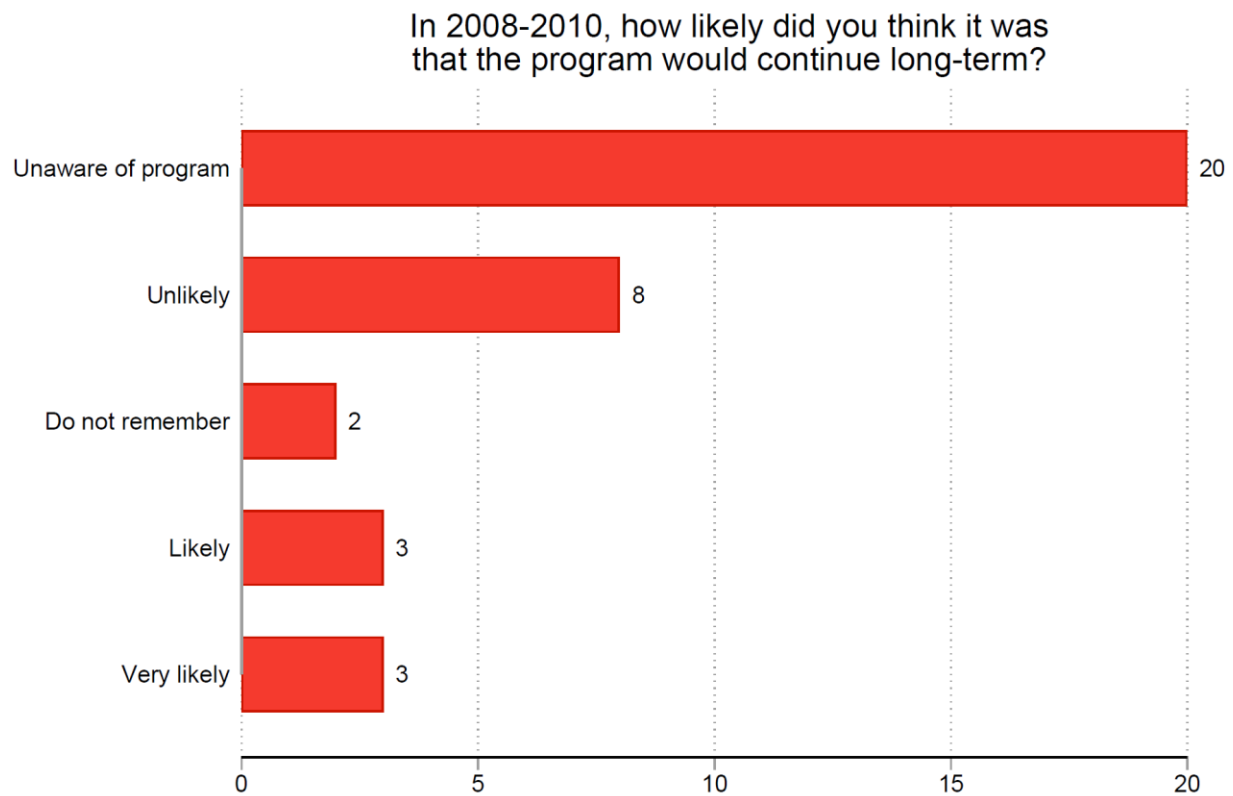
Notes: Each bar represents the number of survey respondents (out of 36) who selected the corresponding answer to, “To the best of your knowledge, for what purposes were sparsity aid funds used? Please select all that apply.”

Figure B.3.3. Distribution of responses for Q8



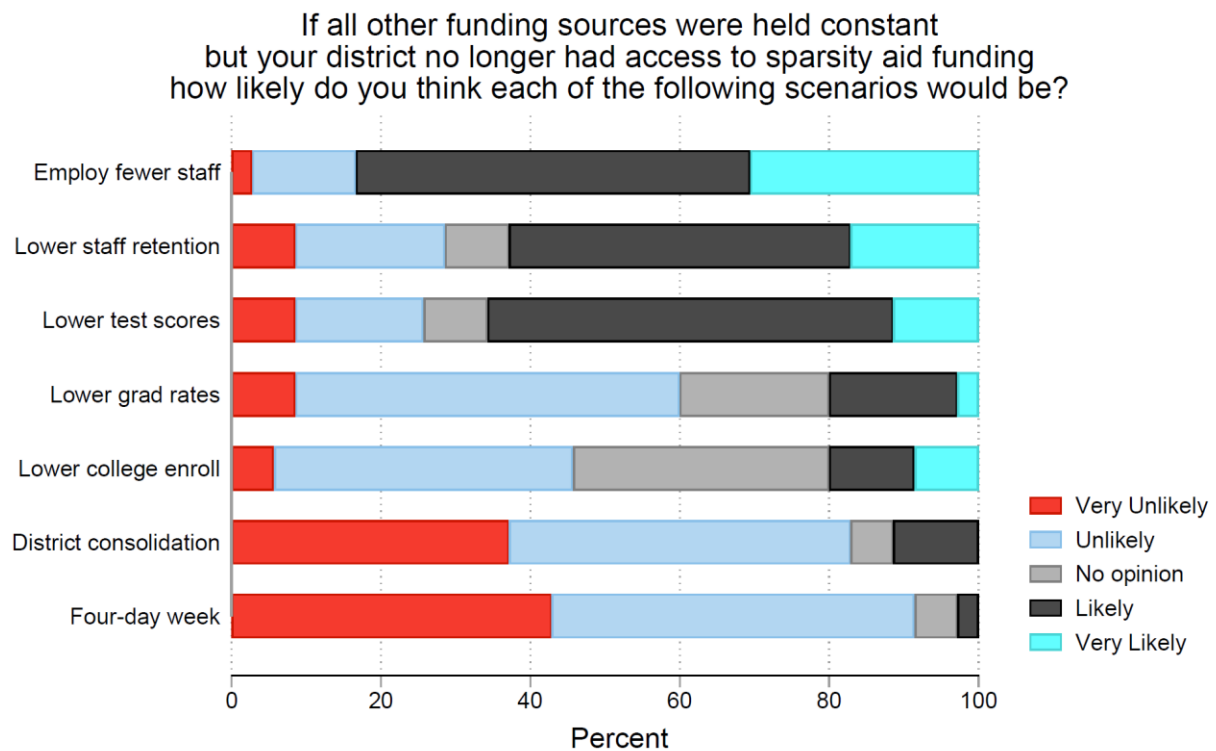
Notes: Each bar represents the number of survey respondents (out of 36) who selected the corresponding answer to, “To the best of your knowledge, were sparsity aid funds spent on specific grade levels? Please select all that apply. Note, the sum of responses equal 37 because one respondent answered with both Elementary and Middle/junior high.”

Figure B.3.4. Distribution of responses for Q9



Notes: Each bar represents the number of survey respondents (out of 36) who selected the corresponding answer to, “Thinking back to the early years of the sparsity aid program (2008-2010), how likely did you think it was that the program would continue long-term?”

Figure B.3.5. Distribution of responses for Q10



Notes: Each colored bar segment represents the proportion of survey respondents (out of 36) who selected the corresponding answer to, “If all other funding sources were held constant, but your district no longer had access to sparsity aid funding, how likely do you think [scenario] would be?”