Selection and Performance in Teachers' Unions

Morgan Foy*

October 2025

Abstract

This paper examines whether teachers' unions affect student achievement in Wisconsin. First, I establish several facts about which teachers are voluntary union members. In particular, I find that union members appear negatively selected on teacher value added. Second, using the staggered decertification of district unions over time, I find improvements in student outcomes after decertification. These effects are not driven by compositional changes within the teaching workforce; rather, I find evidence suggesting that teachers' productivity improved in decertified districts. Together, the results imply that union efforts to insulate workers may adversely affect the quality of public services.

*University of Illinois Urbana-Champaign (email: foy@illinois.edu). I thank Reed Walker, Guo Xu, and Jesse Rothstein for their extensive advice and support. For helpful comments and suggestions, I thank Sarah Baker, Sydnee Caldwell, Ernesto Dal Bó, Hilary Hoynes, Pat Kline, Alex Mas, Ricardo Perez-Truglia, Doug Staiger, Ini Umosen, Chris Walters, Danny Yagan, and many seminar participants. I also thank staff at the Wisconsin Department of Public Instruction for help in accessing the data. I gratefully acknowledge financial assistance from the Institute for Research on Labor and Employment and Xlab at UC-Berkeley. This research was approved by the UC-Berkeley IRB office under protocol number 2022-04-15229.

I. Introduction

Unions in the US are relatively common in the public sector, where about one-third of workers are union members — roughly five times the rate observed in the private sector (Bureau of Labor Statistics 2024). This is especially true for certain frontline occupations such as school teachers and police officers. Despite this, there is little evidence on whether public-sector unions enhance or hinder worker productivity. On the one hand, unions may provide workers with a "voice" in the workplace, improving morale or leading them to take greater agency in their roles (Freeman and Medoff 1984). Conversely, critics have expressed concerns that unions insulate underperforming workers from accountability (Moe 2011). The question of whether unions produce better workers or protect underperformers has important implications for the quality of public services like education or public safety.

In this paper, I study the relationship between public-sector unions and the quality of public services by analyzing whether teachers' unions affect student achievement. This question has been challenging to address due to the non-random nature of union existence (Hoxby 1996, Lovenheim 2009, Lovenheim and Willén 2019). Additionally, while there is extensive research on measuring worker performance for certain public-sector occupations (e.g., teachers), few studies are able to observe individual union membership status to examine the relationship between membership and performance. To overcome these limitations, I examine the education sector in Wisconsin, where the union environment has changed dramatically in recent years. In 2011, Wisconsin enacted legislation known as Act 10 that effectively made the state "right-to-work" in the public sector, allowing workers to choose whether to be union members or not. The legislation also curtailed public-sector unions' collective bargaining rights and required that they hold *annual* elections to remain certified as the exclusive bargaining representative.[1] Prior to Act 10, over 98 percent of teachers were union members, but after its implementation, teacher membership dropped below 50 percent (National Center for Education Statistics 2008, 2021). Likewise, roughly half of the state's 400 local teachers' unions lost all collective bargaining rights after failing to win support via the recertification election process. Importantly for my study, I observe individual teachers' union status linked to performance metrics, the universe of recertification election results, and a long panel of student achievement measures. This allows me to examine the relationship between membership and teacher performance and, concurrently, whether union "decertification" affects student outcomes.

My first contribution is presenting several novel descriptive facts on which teachers select into union membership. I do this by linking administrative personnel data on all public-school teachers

[1] The law also affected other aspects of education, such as teacher pay structures (Biasi 2021) and retirement benefits (Roth 2019).

with itemized campaign finance records that reveal which teachers pay dues to the union. On the surface, there exists an obvious free-rider problem since Act 10 allowed workers to stop paying dues even in union-covered workplaces. However, there are other reasons why workers voluntarily join unions, such as for social benefits and representation in disciplinary matters. In a conceptual framework, I outline why selection into membership may be positively or negatively related to worker quality depending on the correlation between performance and these individual benefits. To test this empirically, I construct a measure of teacher value added using student-teacher linked data (Kane and Staiger 2008). To my knowledge, this is the first paper to assess selection into unionization by linking individuals' union status with measures of productivity.

I find evidence of negative selection into membership: teachers with lower value-added scores are more likely to be members than higher value-added teachers holding constant teacher experience. My estimates imply that a teacher at the 5th percentile of the value-added distribution is 8 percent more likely to be a union member than a teacher at the 95th percentile. This pattern aligns with the idea that teachers join the union in part for representation benefits, which may be more valuable for lower performers. I can also observe which teachers vote in the recertification elections, but find no correlation between value added and supporting the union via the recertification process. This implies that all types of teachers want the union to exist, but low performers are more likely to value the individual benefits above the cost of membership. Finally, the negative selection patterns are driven by teachers in certified school districts rather than in decertified ones, suggesting that the representation benefits are diminished after the union decertifies.

Guided by the descriptive selection results, I next examine whether the diminution of a local teachers' union affects student outcomes, given that about half of the state's 400 local unions decertified in the ten years after Act 10. I use a staggered difference-in-differences research design, comparing student outcomes in decertified school districts relative to students in districts that remained certified.

I find suggestive evidence that decertification led to a significant decline in my measure of union membership. While I can only observe membership starting in 2016, event study estimates using the districts that decertified after 2016 indicate that membership declined by about 30 percent relative to certified districts. I then analyze how decertification affected student outcomes, given that I observe a long panel of student outcome data before and after the decertification events. I find no pre-trends in student outcomes, indicating that "decertifiers" were on similar trajectories to districts that remained certified. However, five years after a district decertified, I estimate that test scores increased by 7 percent of a standard deviation and average attendance rates rose by about 1 percent on average. This indicates that student achievement improved after a shock to the power and organization of the local teachers' union.

In the final part of the paper, I explore potential mechanisms as to what drove the improvement in student outcomes. While there are many potential inputs that affect student achievement, the fact that lower performers are more likely to be union members suggests one of two things at the teacher level. The first explanation is that a compositional change occurred if, for example, decertification allowed for the dismissal of low performers. The second explanation is that teachers increased their productivity in the absence of union coverage.

I find no evidence of large compositional changes in the teaching workforce, though I cannot distinguish between terminations and voluntary departures. There were no significant differences in exit or entry patterns, either on average or by worker quality metrics. This may not be surprising in this context as the state faced a wave of retirements before the Act 10 provisions took effect (Roth 2019), which I find did not differ by certification status. Additionally, Biasi (2021) finds that some districts adopted pay-for-performance type schemes that rewarded high-performing teachers after Act 10. However, I find no evidence that the adoption of performance pay differed between certified and decertified districts.[2]

Instead, the balance of evidence suggests that (1) teacher performance improved and (2) this improvement may be connected to the lack of union protections in decertified districts. To explore whether teacher performance changed, I construct a value-added metric at the grade-level, similar to Biasi (2021), to account for the fact that the state did not track classroom identifiers for students and teachers over the entire sample period. Holding the set of teachers fixed, I find that teachers in decertified districts were more likely to improve on their pre-decertification value-added scores relative to certified teachers over the same time frame. While these grade-level improvements could be potentially explained by some other change in district-level inputs, I find no evidence that average per-pupil spending differed post-decertification.

To further shed light on this, I also created and administered a qualitative survey, which I sent to the majority of teachers in the state. There are two findings that connect back to the descriptive selection results. First, 85 percent of recent union members chose "representation in case of conflict" as one of the reasons why they are a member, making it by far the most cited option. This reinforces the notion that individual representation is a key driver of union membership in a setting where dues are not mandatory. Second, teachers in decertified districts were significantly less likely to believe that they personally had the right to union representation in a disciplinary matter, regardless of their recent union status. Collectively, these results are suggestive that teachers in decertified districts improved their performance because they no longer believed they had the benefit of union protection.

[2]It does appear that compensation shifted in favor of less-experienced teachers, at the relative expense of more tenured teachers. This shift is consistent with the insider-outsider hypothesis of unions (Lindbeck and Snower 2001), as there is a strong positive relationship between years of experience and union membership.

Contribution to Recent Literature: This paper contributes to several literatures in labor economics. First, I add to research that has questioned why individuals join unions in "open-shop" settings. In a seminal work, Olson (1965) posited that, without compulsory mechanisms, rational workers will avoid paying the individual cost of membership because they still reap the collective benefits of unionization. However, other studies theorize that there are individual benefits which may sustain membership rates, such as social customs (Booth 1985, Naylor and Cripps 1993) or because of the representation benefits that unions provide (Blanchflower et al. 1990, Murphy 2020).[3]

I contribute to this research by examining whether the individual benefits to membership influence selection based on workers' underlying performance. This is theoretically ambiguous because we might expect that social customs will cause better teachers to join if being a good teacher correlates with pro-social behavior. Conversely, better teachers may be less likely to opt in if they derive less value over the individual representation benefits. Furthermore, I am able to examine the wedge between which teachers pay dues and which workers support the union in the recertification elections. On the surface, this gap reveals the "free-riders," though it is also consistent with some workers preferring that the union exists, but not valuing the individual benefits above the cost of union dues. More generally, this is among the first papers to use administrative data to examine which workers voluntarily select into union membership. A recent exception is Dodini et al. (2023), who examine how unions affect workers' careers in Norway.

Second, this paper builds on prior work examining whether public-sector unions, and especially teachers' unions, affect public services. One challenge in this literature is that there have been relatively few shocks to public-sector unionization over recent decades. As a result, several studies investigate how the establishment of collective bargaining in the 1960s–1980s affected unionization and student outcomes, though this time period lacks district-level proxies for student learning such as test scores (Kleiner and Petree 1988, Hoxby 1996, Lovenheim 2009, Lovenheim and Willén 2019). Likewise, recent papers explore how state-level, right-to-work legislation affected the education sector (Lyon 2021) and the interplay between state-level school finance reforms and unionization (Brunner, Hyman, and Ju 2020).[4] One concern with using state-level changes as identification is that the laws may be bundled policies containing both pro- and anti-labor components (Paglayan 2019). My paper contributes to this body of work by examining local-level shocks to teachers' unions in a modern setting.[5] This circumvents any state-level confounders and allows me

[3]Prior work also considers selection into union jobs or industries rather than selection into membership status within a workplace (e.g., see Farber 1983, Card 1996, Farber et al. 2021).

[4]For other papers examining the relationship between unionization and student achievement, see Eberts and Stone 1986, Peltzman 1996, Moe 2009, Strunk 2011, Lott and Kenny 2013, Cowen and Strunk 2015, Han 2020, Han and Maloney 2021, Han and Keefe 2022.

[5]An exception that conducts a sub-state analysis is Matsudaira and Patterson (2017) who examine how union certification among California charter schools affected student achievement. Outside of the education sector, several

to explore mechanisms at a more granular level. Additionally, while much of this work focuses on the establishment of unionization in a workplace, this paper examines a period of deunionization, which is particularly relevant given declining unionization rates and the recent *Janus v. AFSCME* Supreme Court case.

Finally, my findings complement the recent literature examining how Act 10 impacted education in Wisconsin. Beyond unionization, the bill affected other elements of public-sector employment such as employee benefits and the pay-setting process. Much of the existing research exploits the fact that the law went into effect at different times, depending on when districts' former collective bargaining agreements (CBAs) expired. For instance, Baron (2018) and Roth (2019) find significant increases in teachers exiting the profession before CBAs expired, but I find that this pattern was similar across decertified and certified districts. Likewise, Biasi (2021) documents how school districts used their new pay-setting flexibility to reward high-performing teachers, though this also did not differ between certified and decertified districts. More generally, Biasi and Sandholtz (2024) find that test scores and vote shares for the Republican governor increased in districts where Act 10 was implemented earlier.[6] While these aspects of Act 10 affected all districts eventually, my research focuses specifically on the differential effects of union decertification, which impacted roughly half of all teachers' unions in the state.[7] I therefore contribute to this research by focusing on the direct impact of deunionization, separate from other factors such as teacher labor supply responses.

II. Context and Setting

A. Teachers' Unions in the US

Teachers unions largely do three things: they negotiate CBAs with local school districts; they conduct political advocacy at the local, state, and national level; and they represent workers in disciplinary and other matters. Representation can take several forms. For instance, members of the National Education Association (NEA) have access to legal liability insurance, an excludable benefit for members. A second form of representation is known as Weingarten Rights, which allows a union-covered employee to request that a representative attend a disciplinary proceeding

recent papers have studied how police unions affect public safety (Goncalves 2021; Dharmapala, McAdams, and Rappaport 2022; Cunningham, Feir, Gillezeau 2021). For evidence on how private-sector unions affect productivity, see Lee and Mas 2012; Sojourner et al. 2015; Dube, Kaplan, and Thompson 2016; Barth, Bryson, and Dale-Olsen 2020; Kini et al. 2022.

[6]See also Litten (2016) who shows that average teacher compensation declined and Biasi and Sarsons (2022) who document a widening in the gender pay gap.

[7]In a cross-sectional policy analysis, Flanders and Tunney (2019) compare districts that voted to decertify against those that remained certified in 2019.

(NLRB v. J. Weingarten, Inc. 1975).[8]

Historically, about half of US states had "right-to-work" laws, which prohibited the compulsory payment of membership dues in union-covered workplaces. In states without right-to-work laws, union-covered workers were required to pay union dues or a close equivalent known as "fair-share" or "agency" fees. However, a landmark Supreme Court case, Janus v. AFSCME (2018), ruled that fair-share fees in the public sector were an unconstitutional restriction on the First Amendment rights of workers who are not union supporters, effectively extending right-to-work to all teachers' unions across the country.

B. Wisconsin Act 10

In February 2011, Wisconsin Governor Scott Walker proposed legislation known as Act 10, a standard budget bill that became controversial because it took several steps to weaken the state's public-sector unions. The bill curtailed collective bargaining rights, restricting unions to only being able to legally bargain over base wage increases capped by the consumer price index (Wisconsin Legislative Council 2011). For teachers, this meant that they had no legal right to bargain over salary schedules, benefits, or non-wage amenities such as class size. Act 10 also made fair-share fees illegal, allowing workers to choose whether to pay union dues or not. Finally, it mandated that public-sector unions must hold *annual* recertification elections to retain their right to collectively bargain over wage increases.[9]

For school districts, the Act 10 provisions took effect once the CBAs between each district and its union expired (CBAs were usually two-year agreements). Prior research has used the staggered expiration of districts' CBAs to identify the short-run effects of Act 10 on various outcomes (Litten 2016, Baron 2018, Biasi 2021, Biasi and Sarsons 2022, Biasi and Sandholtz 2024). This approach leverages the fact that, while Act 10 eventually affected all of public education, the timing varied across districts. In this paper, identification comes from the fact that, although Act 10 greatly affected unionization statewide, some districts lost their legal right to bargain completely, while others have maintained their legal certification status.

Act 10 had a dramatic effect on union membership in the state. Prior to 2011, survey data indicates that 98 percent of teachers were union members, one of the highest state averages in the country (Appendix Figure A1, National Center for Education Statistics 2008). A decade after Act

[8]Technically, all individuals in a union-covered workplace have Weingarten Rights regardless of whether they are dues-paying members or not. An open question, however, is whether the value of this representation is equivalent for members and non-members. I examine the perception of representation when turning to survey evidence in Section VII.

[9]Act 10 also mandated that workers pay a greater share of their health insurance and state retirement contributions. Additionally, teachers often received district-specific supplementary retirement contributions, which were subject to elimination under Act 10. Roth (2019) shows that this led to a spike in teacher retirements before the 2011–12 school year, as districts were required to honor these supplementary contributions before Act 10 went into effect.

10, fewer than 50 percent of teachers were members of a union (National Center for Education Statistics 2021).

C. Union Recertification Elections

Act 10 requires that all state and municipal unions, except public-safety unions, hold annual recertification elections to keep their legal right to bargain over wage increases. To retain legal status, a union needs to get 51 percent of all workers to vote in favor of recertification. This grants the union a year of bargaining rights before another annual vote is required. Importantly, this absolute majority requirement is over all employees, not just those who cast votes. As a result, there is no meaningful difference between a "no" vote and not voting at all.

The recertification election provisions were unprecedented for several reasons. For instance, unions rarely need to hold recognition elections after they become certified, let alone repeating this process every year. Moreover, union officials felt that the voting requirements were overly burdensome given that they needed 51 percent of all employees to vote in favor, rather than just 51 percent of those who cast votes (Schneider 2014). Despite the onerous voting requirements, the vast majority of districts recertify conditional on an election being held, with the modal election vote share around 75 percent (Appendix Figure A2 panel A).

Strictly speaking, the decertification of a local union means that the school district is not legally obligated to negotiate salary increases. However, teachers can still be union members regardless of the union's legal certification status because most local unions are still affiliated with the state chapter of the NEA. It is not clear whether members can still access union representation in decertified districts, a question I address when assessing the survey evidence in Section VII.

The annual recertification elections were slowly implemented due to the enforcement of previously negotiated CBAs and legal challenges to Act 10. While some districts held elections immediately after Act 10 was passed, most districts had CBAs that extended beyond the 2011–2012 school year. Legal challenges to Act 10 then halted the elections completely during the 2012–13 school year, with the process resuming during the 2013–14 school year. Appendix Figure A3 plots the number of decertifications by year, which shows that the majority occurred in 2013–14, the year that the legal challenges were resolved.

III. Data

In this section, I provide a brief overview of the several datasets used in the analysis. For a more detailed description, see Appendix Section B.

A. Wisconsin Department of Public Instruction

Staff Administrative Data: I use individual-level data from the Wisconsin Department of Public Instruction (DPI) spanning the years 2006–2022.[10] The DPI maintains a publicly available dataset of all staff members employed in the state's public schools. This dataset includes an employee's name, annual salary, and demographic characteristics. It also includes information on their specific teaching assignments, such as the school where they worked, and the subjects and grades they taught.

Student Administrative Data: Under a data-use agreement with the DPI, I have access to deidentified data on all public-school students from 2006–2022. This dataset includes students' basic demographic characteristics, attendance and disciplinary records, and standardized exam scores. As is common across states, students in Wisconsin take exams in grades 3–8 and 10th grade.

State testing underwent changes in 2015 and again in 2016, which changed the numerical scale of the exams. To account for this, and to be consistent with prior work, I normalize exam scores to have a mean of 0 and a standard deviation of 1 within each year, subject, and grade. The primary concern for identification is if the changing test regime somehow favored students in certified or decertified districts at the time of decertification.[11] While this is untestable, I also find improvements in attendance, an outcome that is unaffected by the testing changes.

Linking Students to Teachers: For both datasets, I observe the school and grade where each individual worked or attended. However, before 2017, the DPI did not track classroom identifiers, preventing me from observing the exact students an instructor taught in a particular grade.[12] From 2016–17 onwards, the DPI began collecting "staff links," which connects each student to their respective teachers for each enrolled course. I use the staff links to create a classroom-level teacher value-added measure for the years where this data is available. I discuss this further in Section IV.

B. Wisconsin Campaign Finance Data

As is typical across states, Wisconsin requires that political action committees (PACs) disclose individual-level documentation on contributions received. The state makes this data publicly ac-

[10] Henceforth, I refer to a school year using the date that a school year ends. E.g., 2022 is the 2021–2022 school year.

[11] Note that this concern is not unique to Wisconsin, as other states changed testing over this time period to align with the "Common Core" standards. For example, California changed exams in 2014–15 as well (Warren and Murphy 2014). Backes et al. (2018) find that teacher value added measures are generally persistent across test regimes, especially in math.

[12] To circumvent this issue, Biasi (2021) constructs a value-added metric at the grade level using the teacher-student-grade assignments. I discuss this further in Section VII.

cessible online, where users can view the names of individual donors and often other information such as the donor's address or occupation. Most states implement an annual small donor threshold, under which a PAC can lump together "unitemized" donations into one total. Prior to 2016, this threshold in Wisconsin was set at \$20, meaning that a PAC could combine annual donations of \$19.99 or smaller without providing detailed contributor-level information. However, starting in January 2016, Wisconsin required PACs to itemize every contributor's donations, regardless of the amount.

This is relevant because the state chapter of the NEA automatically deducts \$19.99 from each member's annual dues and routes the funds to its PAC. These contributions are reported several times per year and the comments note that the contributions are from "eDues." Likewise, there are 13 regional chapters of the state union, each with its own PAC, that automatically receive \$5 from member dues. This allows me to identify which teachers are paying membership dues for the local unions affiliated with the NEA.[13] Due to the 2016 removal of the itemization threshold, the PACs were required to list the name and address of every union member who, by a condition of paying dues, is indirectly contributing to the union's PAC.[14] In contrast, before 2016, the PACs only list a sum of unitemized member fees. Appendix Figure A4 illustrates the annual distribution of individual contributions, revealing that the majority of contributions are bunched at the \$20 mark for the state PAC (panel A) and at the \$5 mark for the regional PACs (panel B).

C. Certification Elections

The Wisconsin Employment Relations Commission publicly lists the results of the annual recertification elections, which I use to determine when a union decertified. The election results include the number of votes in favor of and against recertification for each union that submits a petition for recertification. The employment relations agency also maintains annual lists of the recertification election voters. Per Wisconsin law, information on individual voter turnout is public record, though ballot choices are not. I submitted an open records request for the universe of recertification voters for each school district and year since the elections started in 2012. While this individual-level data does not indicate whether someone voted in favor of or against recertification, the median district-level share of votes in favor was 99 percent, given that there was no incentive to cast an "against" vote (see Appendix Figure A2 panel B). I therefore use this information as a proxy for who supported the union in the recertification process.

[13] About 97 percent of districts are affiliated with the Wisconsin Education Association Council, the state chapter of the NEA. While other major unions also have PACs, such as the American Federation of Teachers, they do not automatically deduct a small share of membership dues for campaign finance purposes.

[14] A member can submit a written request for a refund of the voluntary PAC contribution (see an example membership form here https://weac.org/wp-content/uploads/2020/08/20-21-Membership-form_Main_Teacher_FINAL.pdf). However, I find that very few people actually receive this refund, as the PAC is also required to itemize disbursements.

IV. Sample Construction

A. Measuring Union Membership

In this section, I provide a brief outline of how I determine which teachers are union members and recertification election voters. See Appendix Section B for a more detailed discussion. First, I merge the campaign contributions to the DPI administrative salary data to construct a measure of union membership. The campaign contributions data serves as a reasonable proxy to a list of NEA members given how the state union routes dues to its PAC. The assumption is that anyone who appears in the PAC receipts is actually a union member and not someone who independently donated to the union PAC. This seems reasonable because (1) it is hard to imagine a non-member donating to the union PAC and (2) empirically, the distribution of donations is centered around the PAC thresholds as shown in Appendix Figure A4.

Since each of the 13 regional chapters has its own PAC, the identifying information that is common across the two datasets is a person's full name and region. Approximately 99 percent of teachers are uniquely identified by name and regional chapter. Therefore, I first match exactly on name within each region, after which I use a fuzzy merge to correct for minor misspellings.

I denote a person as a union member if the DPI salary data and campaign contributions data match in a particular year. The total matched membership counts are similar in size to publicly reported figures of unionized teachers in Wisconsin. Appendix Figure A5 compares the number of members in NEA financial reports (gray) against the number of union members from the matching procedure (blue) and the total unique individuals in the union campaign contributions data (red). There is a gap of around 5,000 people between my matched count and the NEA-cited counts, which is reasonable because the NEA numbers include support professionals who may not be included in the DPI teacher dataset. The red line of Appendix Figure A5 shows the number of unique small donors in the union campaign contributions data, which closely resembles the NEA-cited numbers.[15] Additionally, panel B of Appendix Figure A5 compares the share of teachers who I denote as union members relative to survey evidence on teacher union membership in Wisconsin. Reassuringly, my matched sample (blue) falls in between estimates from the Current Population Survey (red) and the National Teacher and Principal Survey (gray).

Finally, I merge in the individual-level union elections voter data. This data consists of a list of names for each school district. I follow the same procedure of matching on full names, but this match is relatively easier because both datasets include the exact school district rather than the coarser WEAC region (99.9 percent of names are unique by school district). Again, I correct for

[15] Note that there appears to be an under-count in 2016 since the itemization requirement started mid-school year in January 2016. This can occur as some people are listed as giving $19.99 at the beginning of the school year, while others give around $5 every quarter. I detail how I address this issue in Appendix Section B.

misspellings using a fuzzy matching procedure.

B. Measuring Teacher Performance

As previously discussed, the DPI only started tracking classroom student-staff links in the 2017 school year. This allows for the construction of a value-added measure for the 2017 to 2019 school years, before standardized testing was disrupted due to the COVID-19 pandemic.[16]

When measuring teacher value added, the econometrician is interested in predicting the portion of a student's standardized test score explained by the student's teacher (See Bacher-Hicks and Koedel 2023 for a review). To estimate this, consider an equation describing how a student's test scores in a given year are a function of observable characteristics and an unobserved term. Specifically, for student i taught by teacher j in year t, the equation of interest is

$$A_{ijt} = \mathbf{X_{ijt}}\beta + v_{ijt},$$

where A_{ijt} is a student's test score in a given year, $\mathbf{X_{ijt}}$ is a vector of observable covariates, and v_{ijt} is the unobserved residual. The residual can be expanded as

$$A_{ijt} = \mathbf{X_{ijt}}\beta + \mu_j + \theta_{jt} + \epsilon_{ijt},$$

where μ_j is the teacher effect, θ_{jt} is an unobserved classroom effect in year t (e.g., if there are peer effects), and ϵ_{ijt} is an idiosyncratic student-level effect.

My primary method for estimating the μ_j parameters uses the Kane and Staiger (2008) two-step approach, which includes an empirical Bayes term to shrink less reliable estimates toward the mean of 0. In robustness exercises, I examine other value-added measures, including a one-step estimator and using the unshrunken estimates directly.[17]

Here, I give a brief overview of the procedure; see Appendix Section C for a more detailed description. In the two-step approach, I first capture the residuals from the following regression:

$$A_{ijt} = \alpha + \mathbf{X_{ijt}}\beta + v_{ijt},$$

where A_{ijt} is a student's math or English/language arts (ELA) test score in a given year and $\mathbf{X_{ijt}}$ is a vector of control variables. Following the literature, I include as controls a cubic polyno-

[16]There was no standardized testing in 2020 and limited testing in 2021. Universal testing resumed in 2022, but value added can only be measured for those who took the test in 2021 (to include a lagged score control) opening up the possibility of bias resulting from selection into taking the test in the 2021 school year.

[17]One could also estimate a time-varying value-added measure as in Chetty, Friedman, and Rockoff (2014). However, the "CFR" approach is less suitable in this setting because it uses a jackknifing procedure from surrounding years to apply an estimate to the leave-out year. This excludes important information as I am interested in incorporating a teacher's "effect" over the time frame for which I have membership data.

mial of lagged test scores in both ELA and math interacted with a student's grade, as well as classroom-level averages of these lagged test score polynomials. I also control for lagged disciplinary measures (grade repetition, log days disciplined, and log days absent) in the same fashion at the student and class levels, as well as grade-by-year fixed effects, class size, and student and classroom average demographics (special education status, limited English proficiency, FRPL status, gender, and race indicators).[18] Controlling for lagged test scores limits the estimation of value added to math or ELA instructors who teach grades 4–8.

In the second step, the residuals from the above regression are then regressed on teacher-by-year indicators. To obtain a time-invariant measure of value added, a teacher's average annual residuals are weighted by their precision, where years with larger classrooms are given more weight. Finally, this unshrunken value-added measure is multiplied by a Bayes shrinkage term, where less reliable estimates (e.g., teachers with fewer years in the data or fewer students) are shrunk towards the mean of 0.

To my knowledge, this is the first paper to analyze classroom-level value-added measures in Wisconsin. As observed in other settings (e.g., Chetty, Friedman, and Rockoff 2014), I find that teacher effects are more pronounced in math than in reading, and slightly more pronounced in elementary school than in middle school (see Appendix Figure A6 for the distribution of estimates). In general, the two-step model implies that a one standard deviation increase in teacher value-added raises student test scores by 0.14 standard deviations in math and 0.11 standard deviations in ELA.

V. Descriptive Evidence on Selection into Union Membership

In this section, I consider whether union members are positively or negatively selected on value added. This is informative because it provides descriptive evidence on whether there are possible incentives to being part of the union that may affect students. For instance, if members are negatively selected, decertification may improve student achievement; conversely, if members are positively selected, decertification may harm student performance. I first discuss the possible reasons why teachers join the union when there is no compulsory mechanism in place forcing people to pay dues. Then, I explore the empirical relationship between membership and value added using the linked teacher data.

In Appendix Section D, I lay out a conceptual framework that considers whether selection into union membership is positive or negative as it relates to workers' performance. The simplified model builds on prior work that suggests that unions overcome the free-rider problem due to either excludable representation benefits (Blanchflower et al. 1990, Murphy 2020) or due to social

[18]Observations are dropped when missing either the test score of interest or the lagged score. If a different variable is missing, the observation is included with an indicator variable denoting missing status.

reasons (Booth 1985, Naylor and Cripps 1993). The sign of selection is theoretically ambiguous because it depends on whether representation benefits and/or social reasons are correlated with workers' type or performance. If representation is the dominant factor, the framework predicts that selection will be negative as lower performers will benefit relatively more from union coverage. However, if selection is positive, it indicates that there exists a positive relationship between social motivations and performance that is relatively more important than representation benefits.

With this in mind, I document several facts regarding which workers voluntarily choose to be union members. In particular, I estimate a series of OLS regressions examining the predictors of union membership to assess whether membership is positively or negatively associated with performance. Recall that this can only be examined for a subsample of the teacher workforce for whom value added can be estimated (i.e., math or reading teachers in grades 4–8). Table 1 column 1 reveals a strong positive relationship between membership and years of experience, which can be seen visually in the red dots in Figure 1. Column 2 adds the shrunken value-added estimates, which shows that membership is negatively correlated with performance, conditional on controlling for years of experience.[19] The negative selection pattern aligns with the case in the conceptual framework where the protection benefits are negatively correlated with performance and the social benefits are either not correlated with performance or are less important than representation. Columns 3–4 of Table 1 repeat this analysis, but split the sample by whether a teacher was in a certified or a decertified district. The coefficients are negative in both cases, but the average effect seen in column 2 is driven by the certified districts. This may suggest that the protection benefits, such as Weingarten rights, are weakened or non-existent in decertified districts.

Columns 5 and 6 compare selection patterns into membership versus whether someone voted in the recertification elections, which I use as a proxy for whether someone supported the union.[20] In the conceptual framework, I outline why there may exist a negative relationship for membership but not supporting the union because membership confers individual representation benefits, whereas voting in the recertification elections does not. In column 5, I find a similarly sized correlation between membership and value-added using the subsample of districts where a recertification election took place.[21] However, for the exact same set of teachers, there is no correlation between voting and value added in column 6, as the coefficient is close to 0 and not statistically significant.

[19]Appendix Figure A7 shows that there exists a positive relationship between value added and experience, as has been found in other settings (e.g., Rockoff 2004). Conditioning on experience profiles is important because there are many potential reasons why membership increases with experience unrelated to quality such as the relative cost of membership (dues are a fixed fee for everyone), potential familiarity with unions pre-Act 10, or if unions uphold the experience-driven salary schedules.

[20]As with membership, there exists a strong relationship between years of experience and voting in the recertification elections (gold dots in Figure 1).

[21]Columns 3 and 5 use a similar sample but are not identical because there are a few cases where a decertified district was voting to regain certification status.

This implies that the negative correlation is tied not to a measure of union support, but to the individual benefits of membership. Again, this aligns with the concept that some teachers join the union for the representation benefits, which are only attainable by paying dues.[22]

Figure 2 visualizes the differences in Table 1 columns 3–4 and 5–6. Panel A shows that the significant correlation between membership and value added is driven by the certified districts (blue) rather than the decertified ones (gray). Likewise, Panel B of Figure 2 illustrates that there is a negative correlation between membership and value added (in blue) but not between voting and value added (in gray).

In terms of magnitude, the point estimate in Table 1 column 2 implies that moving from 0 to 1 on the value-added measure is associated with a 15 percentage point reduction in the likelihood of being a union member. However, "1" corresponds to an extremely large value-added score; the standard deviation is 0.074, implying that a 1 standard deviation increase in value added is associated with a 1.1 percentage point decrease in the likelihood of being a union member (0.146×0.074).[23]

To summarize, the descriptive selection analysis reveals that lower value-added teachers are more likely to be union members than higher value-added teachers within the same experience profile, consistent with union representation being more valuable for lower performers. This is especially true in certified districts, but similar patterns are not observed in terms of supporting the union in the recertification election process. Guided by these descriptive facts, I now examine whether union decertification affects student achievement.

VI. Do Unions Affect Student Achievement?

When considering an education production function, there are many potential "inputs" that determine student outcomes ranging from district resources, to teacher impacts, to peer effects (for a review, see Hanushek 2020). Prior work on how unions affect students often focuses on whether unions affect the distribution of spending at the district level (Hoxby 1996, Lovenheim 2009, Lovenheim and Willén 2019). However, the descriptive analysis suggests additional potential channels at the teacher level through which unions may affect student achievement. For example, perhaps school districts have more discretion to dismiss lower performers when unions are not present. Alternatively, perhaps teachers improve in the absence of union coverage due to increased accountability. To investigate whether teachers' unions affect student achievement in a

[22]This is not explained by the measurement error inherent in assuming that voters are casting "yes" votes in favor of the union; Appendix Table A1 finds a similar pattern when only examining district-years where greater than 99 percent of people voted in favor of the union.

[23]The descriptive selection results are robust to using either the unshrunken value added or a one-step value added measure as seen in Appendix Tables A2–A3, respectively. Additionally, the results are robust to using an errors-in-variables model, given the inherent measurement error of using value added as a regressor (Appendix Table A4).

causal manner, the next section examines whether union decertification affected student outcomes using a staggered difference-in-differences research design. I first discuss the sample of school districts used in the analysis, I then outline the research design in more detail, and finally I present the main results.

A. Analysis Sample

In what follows, I detail which of the 421 school districts are included in the analysis sample. First, I drop two sets of districts that consolidated after Act 10, as the pre-consolidation set of districts may have decertified in different years.[24] Second, I exclude seven districts (2 percent of districts in the state) for which I could not find any information on a union relationship with either NEA or the American Federation of Teachers (the other national teachers' union). I also drop all non-institutional charter schools, as the teachers are not school district employees.[25]

For the remaining 412 districts, there are 160 districts that have maintained certification status over the entire sample period and 205 districts who decertified and never regained their certification status. Additionally, there are 47 districts that decertified and later recertified, which I refer to as the "recertifiers." Appendix Figure A8 illustrates the geographic distribution of decertifications across Wisconsin. Decertifications occurred in all regions but were more prevalent in the northern, more conservative parts of the state.

There is no central source of which districts are certified at any given time. In most cases, this information can be ascertained according to the local union's voting behavior. For instance, it is clear for the 160 certified unions because they consistently vote above 51 percent over the time frame. For the remaining districts, there are two possibilities. First, some districts voted above 51 percent until either losing an election or opting not to hold one. This makes clear the "event year" as the decertification occurred either when the district failed to reach the 51 percent threshold or the year after the string of successful votes stopped. Second, some districts never held an election and therefore decertified when their previous CBA expired. There are 89 districts meeting this criteria, which means they decertified in either 2012 or 2014.[26] Biasi and Sarsons (2022) identified the timing of the CBA expirations for 46 of these districts, leaving 43 districts where the precise year of decertification is unknown.

In the primary analysis, I drop the 43 districts with unknown decertification dates, but examine robustness to conservatively assigning 2012 as the event year in section VI. Initially, I also ex-

[24]The two consolidations are the Herman-Neosho-Rubicon district (consolidated in 2016) and the Holy Hill district (consolidated in 2018). There were five reorganizations after 2006 but before Act 10. For these districts, I collapse the teacher and student data into the district that remained as of the 2011 school year.

[25]There are other "institutional" charter schools, in which the teachers are school district employees. I keep these schools in the sample as the teachers are often union members and recertification election voters.

[26]Recall that there were no elections in 2013 due to the legal challenges.

clude the recertifiers given that these districts later regained their legal status, making it less clear whether the decertification truly represented a shock to the local union. In additional specifications, I examine results including the recertifiers in what can be considered an "intent-to-treat" (ITT) design.

B. Research Design

The ideal experiment would randomly assign unions to decertification status and compare student outcomes. The ideal quasi-experimental approach might involve using a regression discontinuity design, given that districts vote each year to remain certified. However, few district-year observations fall right below the 51 percent threshold (see Appendix Figure A2 panel A). This would also discard valuable information because some districts decertify without holding a vote. Presumably, these districts would fall below the 51 percent threshold if they were to hold an election but, unfortunately, this is unobserved.

Instead, I take advantage of the longitudinal nature of the data and estimate difference-in-difference regressions and associated event studies. The primary identifying assumption is that the timing of decertification is unrelated to student outcomes. This assumption would be violated if, for instance, districts decertified after a downwards trajectory in student achievement. However, two features of the setting help to mitigate this concern. First, most decertifications occurred in districts' first year of eligibility after the law was passed. As discussed in Section II, this period was marked by uncertainty about the legality of the decertification process. This uncertainty included an actual pause in recertification elections in 2013 before the state Supreme Court ruled that the process was legal shortly before the 2014 elections. Second, given the panel data setup, I can assess event study evidence to see if the two sets of districts were on parallel trends prior to decertification.

Identification also relies on the assumption that no confounding shocks occurred in decertified districts at the time of decertification. Although I cannot rule out every possibility, I consider the most likely confounders (in particular, the other parts of Act 10) when assessing robustness.

Finally, the estimation is potentially susceptible to making poor comparisons due to the staggered rollout of decertification (Goodman-Bacon 2021). To address this I follow the recent literature on staggered difference-in-differences and re-estimate results with robust estimators (de Chaisemartin and d'Haultfoeuille 2020, Callaway and Sant'Anna 2021, Sun and Abraham 2021). I also present simple difference-in-difference results (i.e., not staggered) using only the 2014 decertification cohort as the "treatment" group, given that this is when the majority of decertifications occurred.

Table 2 examines differences in baseline characteristics between certified and decertified districts. On average, certified districts are larger as they are more likely to be located in urban areas. Accordingly, decertified districts have a smaller tax base, fewer students enrolled, and teachers are paid less. In terms of student characteristics, decertified districts are less diverse, with fewer minority students and limited English proficiency students. More importantly for the research design, the two sets of districts exhibit similar baseline trends. Column 3 examines the three-year trends in outcomes leading up to Act 10, finding no statistically significant differences in district or teacher characteristics. For students, nine of ten outcomes are balanced with the exception of the share of Hispanic students, which is growing in certified districts but stable in decertified ones. I probe how this affects results by incorporating inverse propensity weighting and separately estimating results by various subgroups.

The baseline regression model is as follows:

$$Y_{idt} = \alpha + \beta D_{dt} + \gamma_d + \theta_t + \epsilon_{idt}, \tag{1}$$

where Y_{idt} represents an outcome for either teacher or student i in district d and year t and D_{dt} is an indicator variable denoting whether a district's union was decertified in a given year. For certified districts, D_{dt} is always equal to 0; for decertified districts, $D_{dt} = 1$ once the union decertifies and $D_{dt} = 0$ before the decertification event. All specifications include school district γ_d and year θ_t fixed effects. The parameter of interest is β, which captures the effect of decertification on a particular outcome.

Recall that the primary specification uses a sample of districts that were always certified or decertified once and never regained certification. In this case, D_{dt} is an absorbing state. However, in some specifications, I use the ITT sample, which includes the districts that recertified. For this specification, D_{dt} denotes all years after the initial decertification, regardless of when or if the district recertified. I cluster standard errors at the district level, which corresponds to the level of treatment.

The corresponding event study used throughout the paper is

$$Y_{idt} = \alpha + \sum_{\tau \neq -1} \beta_\tau \mathbb{1}(t - e_d = \tau) + \gamma_d + \theta_t + \epsilon_{idt}, \tag{2}$$

where the terms are as before, except that the decertified indicator D_{dt} is replaced with indicator variables for each event time year before or after the decertification event year e_d. This allows for the examination of trends in the treatment effect relative to the year before decertification occurred ($\tau = -1$). Unless otherwise noted, I use a sample of data from 2006–2019, excluding data from 2020–2022 that was potentially affected by the COVID-19 pandemic.

C. Main Results

Union Membership: Ideally, I would first trace out the full event time path of union membership before and after decertification as a measure of the "first stage." Unfortunately, this is not feasible because the itemized campaign finance data starts in 2016, which is after most districts had already decertified. However, I show two pieces of evidence that decertification significantly affected membership rates. First, I estimate event study results using a small number of districts that have membership data both before and after decertification. Then, I present average differences between the early decertifiers and the certified districts, which are roughly equivalent to estimating 10-year first difference regressions.

In Figure 3, I plot event study estimates of how decertification affected membership for the districts that decertified after the campaign finance data is available. The blue dots show results using the baseline two-way fixed effects model and the red dots show results using the Callaway and Sant'Anna (2021) estimator (hereafter, I refer to this as the CS estimator). The figure displays estimates for three years before and three years after decertification as few districts have data outside of this range. With the caveat that this figure only uses twenty "treated" districts, the results indicate that membership declined significantly after the union decertified. The point estimates suggest that membership declined by 12–14 percentage points, or about 30 percent relative to the baseline mean of 0.41. This implies that while decertification may have meant relatively little from a legal perspective, it significantly affected the union's ability to organize members. Appendix Figure A9 demonstrates the robustness of this result to using other two-way fixed effect estimators (panel A), including individual controls for education and experience (panel B), including controls for the recertification election vote share (panel C), and including event years that have no pre- or post-decertification data (panel D).[27] Results are similar though there may be a slight pre-trend when including all event years (panel D, blue dots), owing to the fact that the 2022 decertifiers, who are outside the student data sample, seem to be trending downwards before decertifying. However, this pre-trend is flat when adding controls as in panels B and C (red dots).

There are two possible reasons for the decline in membership. The first is that former members dropped their membership after decertification. The second is that the union's sign-up rate declined. Appendix Figure A10 examines these two possibilities, plotting the share of former members who left the union in a given year against the share of new sign-ups. While noisy, the figure suggests that the effect is driven by the first explanation as there are significant point estimates for former members exiting (gold), but no difference in the rate of new sign-ups (red).

The preceding analysis suggests that decertification affected membership but does not incorporate the districts that decertified before 2016, which represents the majority of the treatment

[27] When controlling for the election vote share, I input a value of 0 for the vote share in district-years where an election was not held following Cellini, Ferreira, and Rothstein (2010).

group. As an additional piece of evidence, Appendix Figure A11 presents membership trends over time between certified districts and the early decertifiers. While the data does not allow for pre-decertification estimates, it is important to recall that union membership was nearly universal before Act 10 (National Center for Education Statistics 2008, Appendix Figure A1). In other words, Appendix Figure A11 is equivalent to estimating a series of 10-year first-difference regressions, assuming universal membership prior to 2012. The results of this exercise indicate that, by 2016, membership in decertified districts declined by an additional 25 percentage points compared to certified districts, a difference of nearly 50 percent.

Overall, the evidence suggests that decertification hampered unions' ability to organize as former members left at a faster rate. Next, I examine how the decertification of a local union affected student outcomes, for which I have a complete panel of data before and after decertification.

Student Test Scores: I first examine how decertification affected students' standardized exam scores, averaging across all subjects (math, ELA, science, and social studies). I present the main findings in Figure 4 panel A, which plots event study coefficients for six years before decertification and five years after decertification over the 2006–2019 time period.[28] This allows for a wide window around decertification, but avoids confounding the estimates with any COVID-related disruptions to state testing. While the point estimates are near-zero and not significant prior to decertification, the coefficients in years 2–5 are significantly different from 0 at the 95 percent level. By three years after decertification, student test scores have increased by over 5 percent of a standard deviation. This indicates that the decertification of the local teachers' union led to a significant gain in student test scores.

Panel B of Figure 4 repeats the prior analysis, but focuses solely on the 2014 decertifiers and the certifiers. This approach has two advantages: First, it shows robustness to the fact that, while the main sample is balanced in calendar time, it remains unbalanced in event time given that more recent decertifications may not have five post-treatment years. Second, it allows for the inclusion of a wider window around decertification. In particular, I plot eight years pre-decertification and also incorporate the post-pandemic years to examine whether the pandemic reduced or increased the estimated impacts. As with panel A, there is no pre-trend leading up to 2014. The point estimates also exhibit a similar upward pattern after decertification, remaining elevated after the 2020 school year when no testing occurred.

I summarize the test score impacts in Table 3 panel A. Columns 1 and 2 present the point estimates for the baseline two-way fixed effects and CS estimators, respectively. The average post-decertification impact is 0.049σ in the baseline model and 0.041σ in the CS model. The next row

[28] All event year groups have six years of data pre-decertification; most of the decertifiers have five post-decertification years given that most events occurred before 2015.

summarizes the 5+ year impact as seen in Figure 4, which is about 0.07σ. Columns 3–4 repeat this analysis using the ITT sample which includes the recertifier districts. The coefficients are somewhat attenuated compared to columns 1–2, likely because these additional districts maintained a higher level of union support post-decertification relative to the always-decertified districts.[29]

To settle on a baseline estimate, I follow Armstrong, Kline, and Sun (2025), which finds the optimal weighted average between the two-way fixed effects and CS models. Using this procedure, I find a single DiD parameter estimate of 0.045σ. To benchmark this baseline effect, it is useful to compare it to estimates from the school finance and teacher value-added literatures. In a meta-analysis of the effects of school spending, Jackson and Mackevicius (2024) find that, on average, a $1,000 per student increase in spending (in 2018 dollars) causes average test scores to increase by 0.045σ.[30] This suggests that the decertification impact on test scores is roughly equivalent to a $1,000 per pupil spending increase. In my context, average spending in 2018 was roughly $14,000 per student, suggesting that the effect sizes are equivalent to increasing spending by around 7 percent. Regarding teacher quality, Bacher-Hicks and Koedel (2023) note that teacher value-added studies typically find that a 1 standard deviation better teacher improves test scores by around 0.075 – 0.15σ. Therefore, my estimated decertification impacts are about one-third to two-thirds of the effect of a 1 standard deviation improvement in teacher performance.

Appendix Figure A13 plots the event studies by each individual subject. I find the largest impacts in math, with significant, but smaller, increases in ELA. This pattern again aligns with the teacher value-added literature, which consistently finds that teachers have larger impacts in math than in reading (Bacher-Hicks and Koedel 2023). Decertification also led to significant gains in science (column 3) and social studies (column 4), though fewer grades take exams in these subjects each year.

Student Behavior: Recent research suggests that teachers also influence behavioral outcomes such as attendance and suspensions (Jackson 2018; Petek and Pope 2023; Rose, Schellenberg, and Shem-Tov 2022). To investigate this in my context, I begin by examining how decertification affected students' average attendance rate, defined as the percentage of scheduled school days that a student attended. Unlike with testing, this provides an outcome for students in all K–12 grades.

Figure 5 displays the second main finding: student attendance rates significantly increased following the decertification of the local union. Panel A plots event study estimates using the full sample over 2006–2019 and panel B compares the 2014 decertifiers to the certifiers as in Figure 4. Both panels demonstrate a significant increase in attendance rates following decertification, mirror-

[29] Appendix Figure A12 plots the corresponding event studies using the ITT sample.

[30] Their precision-weighted point estimate is 0.032σ. It is likely more informative to compare my estimates to the 0.045σ figure, as my estimate would also be shrunk when applying their precision weights.

ing the findings observed with test scores. Notably, panel B indicates that attendance effects were particularly pronounced during the COVID-19 pandemic.[31] Table 3 panel B summarizes the attendance rate impacts. Across specifications, there is an average increase of 0.4 percentage points, which increases to a 0.9 percentage point impact five years after decertification. This translates to roughly a 1 percent increase given the high average baseline attendance rate (0.95). The positive effects on attendance help alleviate concerns that the effects were specific to state testing, such as if decertified teachers taught to the test or if confounders arose due to the changing test regime.

Turning to other behavioral outcomes, I find no significant effect on out-of-school suspensions or expulsions (Appendix Table A5 columns 1–2). These outcomes are very rare as about 5 percent of kids are suspended and 0.1 percent are expelled in a given year. I also find no evidence that decertification affected high-school dropout (column 3), an outcome that prior work focused on given the lack of student data in the earlier time periods of study (e.g., Hoxby 1996, Lovenheim 2009). This highlights the importance of incorporating other achievement measures, such as test scores and attendance, that are relevant for all students.

To summarize, I find that union decertification, on average, improved student test scores and attendance rates. Next, I examine heterogeneity by different subgroups and achievement levels both within and across school districts.

Heterogeneity: One identification concern is that the share of Hispanic students was trending differently prior to decertification (Table 2). As a first step to understanding whether this affects the results, I re-estimate the main specification separately by race. Appendix Figure A14 panels A–C plot test score impacts and panels D–F plot attendance impacts for the three largest racial groups: white students (72 percent of all students), Hispanic students (10 percent), and Black students (10 percent). Results are similar across racial groups, with relatively larger point estimates for Black students.

In addition to race, Figure 6 summarizes the base difference-in-differences results for other subgroups, such as student gender, FRPL status, and grade level. In panel A, I find that the standardized test results are especially pronounced for elementary and middle school students, with more muted effects for high school students. This pattern again aligns with the value-added literature that finds that teachers have the biggest impacts in earlier grades (Bacher-Hicks and Koedel 2023). Results appear similar by gender or FRPL status. Panel B repeats this analysis using attendance rate as the outcome variable. The post-decertification effect on attendance is especially strong for FRPL and non-white students.

[31] Decertified districts were significantly more likely to be holding classes in-person during 2021. In unreported results, I find that decertified districts spent 72 percent of 2021 in person, while certified districts spent 50 percent of the year in person (data from the COVID-19 School Data Hub, see Jack et al. 2023).

Next, I examine differences across district-level characteristics. In Appendix Table A6, I report results on test scores where different average district-level variables are interacted with the post-decertification term. For ease of interpretation, I interact the post-decertification term with an indicator for whether a certain district-level characteristic was above median for all districts.

Formally, the estimating equation is

$$Y_{idt} = \alpha + \beta_1 D_{dt} + \beta_2 D_{dt} \cdot \mathbb{1}(X_d > p(50)) + \gamma_d + \theta_t + \epsilon_{idt}, \tag{3}$$

where $\mathbb{1}(X_d > p(50))$ denotes that the average district-level characteristic is above the median for all districts. Note that X_d is not included in the specification as it is collinear with the district fixed effects γ_d. Here, β_2 signifies whether the average test score effect was distinct in treated districts that were above median in the given X variable relative to below-median districts.

First, I interact the post-decertification indicator with two characteristics of the teachers: the average years of experience (column 1) and the average turnover rate (column 2) prior to the recertification elections. This illustrates whether effects were more pronounced in either districts that had more tenured teachers or districts that had more turnover before Act 10 was implemented. For instance, Roth (2019) shows that there was a large spike in exits in 2011 as teachers retired before any changes to their retirement compensation took effect. In both cases, however, the second row of Appendix Table A6 shows that there was no differential impact for districts that had above-median experience or teacher exits. This suggests that the large turnover caused by Act 10 was not related to the estimated impacts of decertification.

Next, I incorporate two baseline characteristics of the students. In columns 3 and 4, I interact the post-decertification term with whether the district had an above-median rate of FRPL students or above-median enrollment. In both cases, the interaction term is insignificant, suggesting that effects were not more or less pronounced in lower-income or larger districts.

Distributional Effects: Next, I examine whether the impacts varied across the student achievement distribution. First, I test whether the variance of test scores or attendance changed by collapsing the data to the district level in Appendix Figure A15. For test scores, the analysis reveals no significant change in the variance of student exam scores following decertification (panel A). However, panel B shows a decrease in the variance of student attendance, implying that attendance patterns became more uniform post-decertification.

I also estimate treatment effects at different quantiles of the student achievement distribution to investigate where the impacts are concentrated more closely. Figure 7 summarizes results from unconditional quantile regressions, where each point represents a different "RIF" regression following Firpo, Fortin, and Lemieux (2009).[32] This exercise indicates how decertification affected different quantiles of the student achievement distribution. For standardized test scores (panel A), the estimated treatment effect remains fairly stable across the student outcome distribution, coinciding with the null effect on the variance. If anything, there are larger impacts for the lowest achievers, though event study estimates in Appendix Figure A16 shows that there are slight pre-trends at the lowest percentiles.[33] For attendance (panel B), the quantile treatment effects are largest at the lower end of the distribution and get progressively smaller moving towards the top percentiles, which again corresponds with the decrease in variance (see event study figures in Appendix Figure A17). For students at the top of the distribution, there is little room for improvement as their attendance records are already close to perfect.

Overall, the heterogeneity and distributional analyses indicate that the achievement results are consistently positive for different subgroups and achievement levels. Additionally, the results are especially pronounced for students who qualify for FRPL and those with higher baseline absentee rates.

Robustness: I probe the robustness of the effects on standardized test scores and attendance in Appendix Figures A18 and A19, respectively. One identification concern is that the share of Hispanic students was trending differently prior to decertification. To address this further, I re-weight the sample using the doubly-robust inverse propensity estimator (Sant'Anna and Zhao 2020, Callaway and Sant'Anna 2021). The propensity score model includes student demographics such as race, FRPL status, limited English status, and gender. In practice, this procedure gives more weight to control students who are more similar to treated students. I find similar results using the re-weighted estimates (panel A).

Panel B demonstrates that the results are almost identical when using alternative staggered two-way fixed effects estimators, such as those by de Chaisemartin and d'Haultfoeuille (2020) or Sun and Abraham (2021).

Next, I incorporate the 43 districts for which the precise year of decertification is unknown. Panel C adds in these districts, conservatively assigning 2012 as the event year. The results remain very similar to the estimates using the primary sample.

An additional concern is that the results are picking up some other aspect of Act 10 unrelated to the recertification elections, such as the timing of when Act 10 went into effect or districts' new ability to set pay more flexibly. To address this, Panels D and E control for either the year that a

[32]The unconditional quantile regression method allows for the estimation of marginal effects at a given quantile. This in general cannot be achieved using conditional quantile regression because the law of iterated expectations does not hold for quantiles.

[33]I superimpose a linear trend in event time within these figures to more clearly observe the deviation from the pre-decertification trend line.

district's CBA expired or controlling for whether the district moved to a more flexible pay scheme. Specifically, I add CBA-by-year or flexible pay-by-year fixed effects as categorized by Biasi and Sarsons (2022) and Biasi (2021), respectively. The results are robust to the addition of CBA fixed effects (panel D), which implies that the findings are not driven by the broader implementation of Act 10. Furthermore, this essentially restricts comparisons to districts that were required to hold an initial recertification election at the same time. I also find similar results when adding fixed effects for whether a district moved to a more flexible teacher compensation scheme (panel E), despite the fact that this drops about half of the districts from the sample (as Biasi (2021) could not find compensation scheme information for all districts). This provides strong evidence that the results are not merely reflecting the shift to flexible pay schemes.

Panel F controls for the recertification election vote share. This effectively compares districts that were on like trajectories in terms of their recertification election voting behavior. The results are similar, though slightly more muted for test scores in the post-decertification period.

Lastly in panel G, I probe for outliers by successively dropping each district from the sample. This exercise plots 322 different leave-one-out point estimates and confirms that the estimated magnitudes and significance are robust to excluding each individual district.

VII. Mechanisms and Survey Evidence

Thus far, the causal evidence shows that student achievement improved following the decertification of the local teachers' union. Pinning down precise mechanisms for these improvements is challenging because, in an education production function, there are many potential teacher- and district-level "inputs" that contribute to student achievement. However, at the teacher-level, the negative selection patterns from Section V suggest that decertification may have led to either a change in the composition of the workforce or teacher improvement in decertified districts. For instance, a compositional explanation would be that union decertification enables administrators to dismiss underperforming teachers or that teachers have a certain preference for union status and move districts accordingly. Conversely, teachers may improve their performance in the absence of union coverage.

I consider these possibilities by grouping the following analysis into compositional vs. direct effects. Compositional impacts include whether the characteristics of teachers or students changed. I consider other explanations as direct effects, which can be thought of as mechanisms that hold the type of teachers and students fixed. This might range from changes in teacher performance to broader district-wide changes. I go through each of these items in turn, starting with an examination of compositional effects.

A. Composition

Teacher Composition: One potential explanation for the achievement increases is that teachers left the district following union decertification. One caveat is that I cannot observe *why* a teacher left and therefore cannot distinguish terminations from voluntary departures. However, I find no evidence of exit or entry differences post-decertification on average. Appendix Table A7 examines the rate at which teachers left the district (column 1) or moved into the district (column 3). For both measures, the post-decertification coefficient is near zero and not significant, indicating that teachers did not move in response to decertification. This also provides supporting evidence that the decertification impact was separate from the large churn in the workforce that immediately preceded Act 10 (Roth 2019).

The previous analysis focused on average differences; however, it remains possible that there was a compositional change that altered the quality of the local workforce. For instance, perhaps poor performers were dismissed without the value of union protection. The ideal test would be to examine whether low- or high-value-added teachers exited or entered a district in response to the union decertification. Unfortunately, the classroom value-added measure is inherently endogenous given that it is measured over 2017–2019. As an example of this problem, a teacher for whom I can estimate value added is in the sample by construction and therefore cannot have left teaching. As a coarser alternative, I construct a grade-level value-added measure that predates Act 10, similar to Biasi (2021). I follow the two-step value-added procedure described in Section IV but replace the classroom-level residuals with grade-level average residuals. Appendix Section C explains this procedure in more detail.[34]

I first estimate grade-level value added from 2006–2011 to assign teachers a pre-"treatment" value-added score. I then repeat the exit and entry analysis but interacting an individual's pre-2012 value added score with the post-decertification indicator. Formally, this specification is similar to equation 3 except that the interaction variable X is at the individual level:

$$Y_{idt} = \alpha + \beta_1 D_{dt} + \beta_2 D_{dt} \cdot \mathbb{1}(X_i > p(50)) + \beta_3 \mathbb{1}(X_i > p(50)) + \gamma_d + \theta_t + \epsilon_{idt}. \tag{4}$$

However, in line with the average differences, the interaction terms in Appendix Table A7 columns 2 and 4 suggest that there was no differential exit post-decertification between low- and high-value added teachers.

In a similar exercise, I examine whether a district's average teacher quality improved after decertification by estimating the primary difference-in-differences but with 2006–11 grade-level

[34] I validate the grade-level measure against the class-level measure in Appendix Figure A20. While the grade-level value added is a strong predictor of class value added, there exists a small degree of bias that is statistically different than 0.

value added as the outcome. If there is an improvement in this measure mirroring the results in Figure 4, then this would provide evidence that the district either shed low-performers or gained high-performers. However, in Appendix Table A8 panel A column 1, the estimated post-decertification effect is a very precise zero, in line with the null average effects from Table A7.

Lastly, Appendix Table A8 panel A summarizes whether other average characteristics of teachers changed post-decertification. Columns 2–4 find no differences in the average experience level, age, or education level of the workforce. Column 5 finds some evidence that the teacher workforce was less likely to be male, a difference of about 4 percent.[35] Likewise, Appendix Table A8 panels B–C repeat this analysis but focusing on people who left the district or moved into the district, respectively. Again, the characteristics of the exiters and entrants do not appear to significantly change post-decertification. In all, there appears to be limited workforce composition changes, both on average and with respect to teacher quality metrics.

Student Composition: A similar explanation would be that the characteristics of the students changed if, for instance, families moved districts in response to decertification. I examine this in Appendix Table A9 and Appendix Figure A22, but find little evidence suggesting that student demographics changed following decertification. There is no difference in the share of students who are FRPL status, Black, Asian, limited English proficiency, or special education status. The Hispanic coefficient is negative but, as seen in Table 2, there are pre-trends in the share of Hispanic students prior to decertification. Appendix Figure A22 panel C shows that this was smoothly changing before and after decertification and therefore does not follow the abrupt change following decertification that is seen for the student outcomes.

B. Direct Effects

The preceding analysis found little evidence of compositional changes. This implies that there was a direct effect of union decertification that affected student outcomes. While I cannot capture all factors that may have influenced student achievement, I consider the leading possibilities such as teacher performance, the distribution of resources, and teacher perceptions.

Teacher Performance: Quantifying whether teacher productivity changed post-decertification is challenging due to the limitations of measuring value added before 2017. However, I examine the closest substitute by creating pre- and post-decertification value added at the grade level, holding the composition of teachers fixed. Specifically, I take the 2014 decertifiers (the majority of the

[35] On average, I find that female teachers have higher value added than male teachers. However, female teachers are especially better at ELA, whereas the test score impacts are concentrated in math. See Appendix Figure A21 for the corresponding event study plots of the specifications in Appendix Table A8.

treatment group) and the never-decertifiers and estimate grade-level value-added separately from 2008–2013 and again from 2014–2019 over a sample of teachers who are present in both periods.

Figure 8 plots the distribution of teacher "gains" between 2008–13 and 2014–19 separated by certified districts (gray) and decertified districts (blue). A teacher with identical scores between the two periods is at 0 on the figure. A teacher who improved is to the right of 0 and a teacher who did worse is to the left of 0. The figure shows that, while both distributions are centered around 0, the decertified teachers are more likely to have positive gains in the right tail of the distribution and less likely to have negative gains in the left tail. A Kolmogorov–Smirnov test of equality rejects that the two distributions are the same (p-value < 0.001).[36] This implies that, for teachers who stayed in the same union regime, the decertified teachers were more likely to improve on their pre-2014 scores relative to certified teachers. It is important to emphasize that that this may be mechanical if some other input positively influenced test scores, though in the next section I consider the most probable alternative explanations (e.g., spending changes). Nonetheless, it suggests that performance improved as it shows evidence that decertified teachers had higher gains on average relative to certified teachers. It also further shuts down the compositional channel as this analysis focuses on teachers present in both time periods.

Distribution of Resources: Another possibility is that the district altered the distribution of resources post-decertification. In Table A10, I examine this by looking at whether teacher compensation and overall district per-pupil expenditures changed. In column 1, I find a positive but insignificant effect on teacher salaries post decertification. This suggests that losing the right to bargain over base wage increases did not hurt teacher pay for decertified districts. I also find no relationship between decertification and per-pupil spending at the district level (column 2). This may not be surprising given the context as Wisconsin has school district revenue limits that can only be exceeded via local referendums (e.g., see Baron 2022).

A similar explanation would be that union decertification allowed administrators to implement pay-for-performance type schemes that reward high-performing teachers, which would be consistent with Biasi (2021). To test this, I interact the post-decertification indicator with teachers' value added scores. Appendix Table A11 column 1 examines this using the classroom value added measure and controlling for one's experience level. The coefficient on the interaction between decertification and value added is negative and insignificant, implying that high-performing teachers were not paid more post-decertification. Likewise, Appendix Figure A24 plots the coefficient of variation within experience levels by event time and fails to find evidence that the variance in

[36]Appendix Figure A23 plots the empirical distribution functions separately by the 2008–2013 and 2014–2019 periods. The pre-2014 distributions are very similar (panel A). In contrast, the post-decertification distribution for the decertified teachers is shifted to the right of the certified teachers (panel B).

salaries within a particular experience bin expanded post-decertification. This suggests that districts' new ability to set pay in accordance with performance as shown by Biasi (2021) did not differ between decertified and certified districts.

While there is no evidence that pay-for-performance expanded, I do find evidence that the salary structure changed with regard to experience. Appendix Table A11 column 2 repeats the previous analysis but interacting the post-decertification coefficient with teachers' total years of experience. Here, I find a strong negative effect on the interaction term. This implies that districts raised the floor of teacher pay for inexperienced teachers at the relative expense of high-tenured teachers. This is also consistent with the idea of the insider-outsider hypothesis of unions (Lindbeck and Snower 2001), as the low-experienced teachers are also the ones who are the least likely to be in the union as shown in Figure 1.

Overall, I find no evidence that decertification affected total spending or that it led to districts implementing pay-for-performance type systems. However, the positive interaction with experience suggests that districts were able to increase starting salaries for new teachers. This may suggest that decertification allowed districts to become relatively more flexible organizations.

Teacher Perceptions: One explanation for the performance gains is that teachers improve when they no longer have the benefit of union representation against management. This theory would align with the descriptive evidence in Section V that showed an inverse relationship between membership and value added specifically in certified districts. While I cannot directly tie this to the performance gains, I turn to survey evidence to understand whether teachers believe that they lost individual union representation benefits after the union decertified.

In the spring of 2024, I conducted a survey to understand Wisconsin teachers' perceptions of their jobs and unions. I sent an online Qualtrics survey to public-school staff in the state, inviting the majority of the teaching workforce to take part in a short survey. In total, I received responses from nearly 3,000 individuals from 343 districts, a response rate of about 6 percent. The benefit of the survey is that it provides more context on what teachers felt about decertification, Act 10, and their jobs in general. However, a limitation is that it is a point-in-time survey years after most districts decertified.[37]

To examine the mechanism of representation protections, I asked respondents about their perceptions of union coverage in their school district. In Table 4 column 1, I first present results regarding whether respondents felt that workers should have the right to form unions at all. Teachers in decertified districts were four percentage points less likely to believe that workers should be

[37] Appendix Table A12 shows that the survey respondents were slightly selected on various characteristics. To account for this, I re-weight the responses to match the joint distribution of the covariates in the full sample, but results are nearly identical without re-weighting. See Appendix Section E for more details on the survey administration.

able to form unions relative to certified districts. However, even in decertified districts, 90 percent of teachers believed in the right to form unions. This suggests that there may be less favorability of unions in general, though the size of the gap is relatively small. In comparison, column 2 shows a substantial gap in whether teachers believed that they personally had union representation benefits. Specifically, teachers in decertified districts were 34 percentage points less likely to agree with the statement "Do you have the right to union representation if being investigated for disciplinary action?" This finding is not strictly because decertified districts have fewer union members: both recent members and non-members in decertified districts perceive a lack of representation relative to their certified counterparts (columns 3–4). This indicates that, while decertification legally affected collective bargaining, the on-the-ground perception is that it reduced union powers more generally. This suggests that one channel for the improvements in student outcomes may have been the perceived reduction in work protections in decertified districts.

Additionally, I asked teachers who had recently been members of the union to list reasons why they were members, which I summarize in Figure 9. Overwhelmingly, the most cited reason for being a union member is "representation in case of conflict" with about 85 percent selecting this option. Other popular selections were that people "felt it was the right thing to do" and because they wanted a "voice in the decision-making process." This provides further support for the idea that people joined for representation benefits as outlined in the conceptual framework.

I also asked teachers about their perceptions of their job and work environment. I summarize these workplace perceptions in Appendix Table A13, regressing various survey answers on an indicator for whether the district ever decertified. In column 1, teachers in decertified districts reported working slightly more hours per week, a mean difference of about 2 percent. Teachers reported similar levels of satisfaction with their work environment (column 2), satisfaction (or lack thereof) with their salaries (column 3), and feelings of security from being laid off (column 4). However, teachers in decertified districts were more likely to express that they had discretion in setting their curriculum (column 5). The fact that teachers felt they have more say over their classroom may accord with the idea that districts became more flexible organizations after the union decertified. In column 6, both sets of teachers expressed having similar levels of agency over their assigned work duties. Finally, I asked teachers to rank what they felt administrators cared about most regarding their job performance. In Appendix Figure A25 panel A, I show that both certified and decertified teachers felt that administrators cared about state standardized testing the most, but there was no average difference between the two groups. This suggests that the test score effects are not due to increased pressure on teachers to raise examination scores. Panel B summarizes what teachers care about most; again, there are no significant differences by union status.

To summarize, I find limited evidence that decertification led to a significant change in the composition of the teaching workforce both on average and with respect to worker performance metrics. Teachers were not more likely to enter or leave the school district. Instead, I find that teachers who stayed present across the pre- and post-decertification periods were more likely to improve in decertified districts relative to certified ones. This implies that the student achievement effects were due to a direct treatment effect of the union. Pinning down the precise mechanisms for these gains is challenging due to the limited range of outcomes in the administrative data. However, I find no evidence that there were changes to total district spending or that districts were more likely to implement pay-for-performance type schemes. Instead, the complementary survey evidence points to the idea that workers lost the ability to access union representation rights in issues with administration. This aligns with the descriptive membership patterns where a negative relationship emerged between membership and value-added, specifically in certified districts. Additionally, there is evidence that the district may have become more flexible in setting policy, as there is evidence of a redistribution of pay towards less-experienced workers and teachers report having more say over their curriculum.

VIII. Conclusion

This paper explores why teachers opt into union membership and the implications of declining membership on student outcomes. In terms of selection, I find that higher-performing teachers, as measured by value added, are less likely to be union members relative to lower-performing teachers. Conceptually, this is consistent with the idea that lower-performing teachers value the representation benefits that unions provide relatively more than higher-performing teachers. Supporting this notion, I find no value-added differential when examining union voting behavior, which alone does not confer any representation benefits.

After a union decertifies, I find that students' standardized test scores and attendance increase relative to districts that retained some level of union power. The results are most pronounced for lower performers at baseline, non-white students, and those who qualify for free- or reduced-price lunch. These improvements are not explained by compositional changes, as there is no evidence of differential teacher turnover in response to decertification. Instead, descriptive survey evidence shows that most teachers in decertified districts do not perceive having protection or representation, consistent with the negative selection patterns being strongest in certified districts that retained some union power.

A couple of caveats should be noted. First, this paper analyzes student outcomes in the short run; deunionization may have different long-run impacts if it affects state spending or the composition of who chooses to become an educator (Baron 2021). Second, this is not the ideal setting to analyze the benefits to unionization for employees given that collective bargaining was largely curtailed statewide. More generally, this paper concludes that the loss of union status benefited students, but whether public-sector unions are welfare-improving may depend on their relative benefits and costs to both employees and the public. A trade-off may exist between improving workplaces through unionization and the spillover effect on public services.

References

Anderson, Michael. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484): 1481–95.

Armstrong, Timothy B., Patrick Kline, and Liyang Sun. 2025. "Adapting to Misspecification." Working Paper.

Bacher-Hicks, Andrew, and Cory Koedel. 2023. "Estimation and Interpretation of Teacher Value Added in Research Applications." *Handbook of the Economics of Education* 6(2): 93–134.

Backes, Ben, James Cowan, Dan Goldhaber, Cory Koedel, Luke C. Miller, and Zeyu Xu. 2018. "The Common Core Conundrum: To What Extent Should We Worry that Changes to Assessments Will Affect Test-Based Measures of Teacher Performance?" *Economics of Education Review* 62: 48–65.

Baron, E. Jason. 2018. "The Effect of Teachers' Unions on Student Achievement in the Short Run: Evidence from Wisconsin's Act 10." *Economics of Education Review* 67: 40–57.

Baron, E. Jason. 2021. "Union Reform, Performance Pay, and New Teacher Supply: Evidence from Wisconsin's Act 10." *AEA Papers and Proceedings* 111: 445–9.

Baron, E. Jason. 2022. "School Spending and Student Outcomes: Evidence from Revenue Limit Elections in Wisconsin." *American Economic Journal: Economic Policy* 14(1): 1–39.

Barth, Erling, Alex Bryson, and Harald Dale-Olsen. 2020. "Union Density Effects on Productivity and Wages." *Economic Journal* 130: 1898–936.

Biasi, Barbara. 2021. "The Labor Market for Teachers under Different Pay Schemes." *American Economic Journal: Economic Policy* 13(3): 63–102.

Biasi, Barbara, and Wayne Sandholtz. 2024. "The Political Consequences of Controversial Education Reform: Lessons from Wisconsin's Act 10." Working Paper.

Biasi, Barbara, and Heather Sarsons. 2022. "Flexible Wages, Bargaining, and the Gender Gap." *Quarterly Journal of Economics* 137(1): 215–66.

Blanchflower, David G., Robert Crouchley, Saul Estrin, and Andrew J. Oswald. 1990. "Unemployment and the Demand for Unions." Working Paper.

Booth, Alison L. 1985. "The Free Rider Problem and a Social Custom Model of Trade Union Membership." *Quarterly Journal of Economics* 100(1): 253–61.

Brunner, Eric, Joshua Hyman, and Andrew Ju. 2020. "School Finance Reforms, Teachers' Unions, and the Allocation of School Resources." *Review of Economic Statistics* 102(3): 473–89.

Bureau of Labor Statistics. 2024. "Economic News Release: Union Members Survey" https://www.bls.gov/news.release/union2.nr0.htm

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225(2): 200–30.

Card, David. 1996. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica* 64(4): 957–79.

Cellini, Stephanie R., Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125(1): 215–61.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104(9): 2593–632.

Cowen, Joshua M., and Katharine O. Shrunk. 2015. "The Impact of Teachers' Unions on Educational Outcomes: What We Know and What We Need to Learn." *Economics of Education Review* 48: 208–23.

Cunningham, Jamein, Donna Feir, and Rob Gillezeau. 2021. "Collective Bargaining Rights, Policing, and Civilian Deaths." Working Paper.

Dal Bó, Ernesto, Frederico Finan, and Martín A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128(3): 1169–218.

de Chaisemartin, Clément, Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110(9): 2964–96.

Dharmapala, Dhammika, Richard H. McAdams, and John Rappaport. 2022. "Collective Bargaining Rights and Police Misconduct: Evidence from Florida." *Journal of Law, Economics, and Organization* 38(1): 1–41.

Dodini, Samuel, Kjell G. Salvanes, and Alexander Willén. "The Career Effects of Union Membership." Working Paper.

Dube, Arindrajit, Ethan Kaplan, and Owen Thompson. 2016. "Nurse Unions and Patient Outcomes" *ILR Review* 69(4): 803–33.

Eberts, Randall W., and Joe A. Stone. 1986. "Teacher Unions and the Cost of Public Education" *Economic Inquiry* 24(4): 631–43.

Farber, Henry S. 1983. "The Determination of the Union Status of Workers." *Econometrica* 51(5): 1417–37.

Farber, Henry S., Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu. 2021. "Unions and Inequality over the Twentieth Century: New Evidence from Survey Data." *Quarterly Journal of Economics* 136(3): 1325–85.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77(3): 953–73.

Flanders, Will, and Lauren Tunney. 2019. "Democracy in the Workplace: Examining Union Recertification in Wisconsin under Act 10." Wisconsin Institute for Law and Liberty report.

Freeman, Richard B., and James L. Medoff. 1984. *What Do Unions Do?* New York City: Basic Books.

Goncalves, Felipe. 2021. "Do Police Unions Increase Misconduct?" Working Paper.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225(2): 254–77.

Hainmueller, Jens, and Yiqing Xu. 2013. "ebalance: A Stata Package for Entropy Balancing." *Journal of Statistical Software* 54(7): 1–18.

Hamlin, Alan, and Colin Jennings. 2019. "Expressive Voting." *Oxford Handbook of Public Choice* 1: 333–50.

Han, Eunice S. 2020. "The Myth of Unions' Overprotection of Bad Teachers: Evidence from the District–Teacher Matched Data on Teacher Turnover." *Industrial Relations* 59(2): 316–52.

Han, Eunice S., and Thomas N. Maloney. 2021. "Teacher Unionization and Student Academic Performance: Looking beyond Collective Bargaining." *Labor Studies Journal* 46(1): 43–74.

Han, Eunice S., and Jeffrey Keefe. 2022. "Teachers' Unions: Engaging Teachers and Improving Student Achievement in States that Prohibit Collective Bargaining." *Educational Policy* 36(2): 407–39.

Hanushek, Eric. 2020. "Education Production Functions." Chapter 13 in *The Economics of Education (Second Edition)* 161–70.

Hoxby, Caroline. 1996. "How Teachers' Unions Affect Education Production." *Quarterly Journal of Economics* 111(3): 671–718.

Jack, Rebecca, Clare Halloran, James Okun, and Emily Oster. 2023. "Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts." *AER: Insights* 5(2): 173–90.

Jackson, Kirabo C. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy* 126(5): 2072–107.

Jackson, Kirabo C., and Claire L. Mackevicius. 2024. "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States." *American Economic Journal: Applied Economics* 16(1): 412–46.

Janus v AFSCME. 2018. No. 16–1466. https://www.supremecourt.gov/opinions/17pdf/16-1466_2b3j.pdf.

Kane, Thomas J., and Douglas O. Staiger. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." NBER Working Paper No. 14607.

Kini, Omesh, Mo Shen, Jaideep Shenoy, and Venkat Subramaniam. 2022. "Labor Unions and Product Quality Failures." *Management Science* 68(7): 5403–5440.

Kleiner, Morris M., and Daniel L. Petree. 1988. "Unionism and Licensing of Public School Teachers: Impact on Wages and Educational Output." In *When Public Sector Workers Unionize*, edited by Richard B. Freeman and Casey Ichniowski, 305–22. Chicago: University of Chicago Press.

Lee, David S., and Alexandre Mas. 2012. "Long-Run Impacts of Unions on Firms: New Evidence from Financial Markets, 1961–1999." *Quarterly Journal of Economics* 127(1): 333–378.

Lindbeck, Assar, and Dennis J. Snower. 2001. "Insiders versus Outsiders." *Journal of Economic Perspectives* 15(1): 165–88.

Litten, Andrew. 2016. "The Effects of Public Unions on Compensation: Evidence from Wisconsin." Unpublished Paper.

Lott, Jonathan, and Lawrence W. Kenny. 2013. "State Teacher Union Strength and Student Achievement." *Economics of Education Review* 35: 93–103.

Lovenheim, Michael. 2009. "The Effect of Teachers' Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States." *Journal of Labor Economics* 27(4): 525–87.

Lovenheim, Michael, and Alexander Willén. 2019. "The Long-Run Effects of Teacher Collective Bargaining." *American Economic Journal: Economic Policy* 11(3): 292–324.

Lyon, Melissa Arnold. 2021. "Heroes, Villains, or Something in Between? How "Right to Work" Policies Affect Teachers, Students, and Education Policymaking." *Economics of Education Review* 82: 1–13.

Matsudaira, Jordan D., and Richard W. Patterson. 2017. "Teachers' Unions and School Performance: Evidence from California Charter Schools." *Economics of Education Review* 61: 35–50.

Moe, Terry M. 2009. "Collective Bargaining and the Performance of the Public Schools." *American Journal of Political Science* 53(1): 156–74.

Moe, Terry M. 2011. *Special Interest: Teachers Unions and America's Public Schools*. Washington, DC: Brookings Institution Press.

Murphy, Richard. 2020. "Why Unions Survive: Understanding How Unions Overcome the Free-Rider Problem." *Journal of Labor Economics* 38(4): 1141–88.

National Center for Education Statistics. 2008. "Schools and Staffing Survey, 2008." Accessed December 2023. https://nces.ed.gov/surveys/sass/questionnaire.asp.

National Center for Education Statistics. 2021. "National Teacher and Principal Survey, 2021." Accessed December 2023. https://nces.ed.gov/surveys/ntps/.

Naylor, Robin, and Martin Cripps. 1993. "An Economic Theory of the Open Shop Trade Union." *European Economic Review* 37: 1599–620.

NLRB v. J. Weingarten, Inc. 1975. No. 73-1363.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge: Harvard University Press.

Paglayan, Agustina S. 2019. "Public-Sector Unions and the Size of Government." *American Journal of Political Science* 63(1): 21–36.

Parsons, Eric, Cory Koedel, and Li Tan. 2019. "Accounting for Student Disadvantage in Value-Added Models." *Journal of Educational and Behavioral Statistics* 44(2): 144–79.

Peltzman, Sam. 1996. "Political Economy of Public Education: Non-College-Bound Students." *Journal of Law and Economics* 39(1): 73–120.

Petek, Nathan, and Nolan G. Pope. 2023. "The Multidimensional Impact of Teachers on Students." *Journal of Political Economy* 131(4): 1057–107.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94(2): 247–52.

Rose, Evan K, Jonathan T. Schellenberg, and Yotam Shem-Tov. "The Effects of Teacher Quality on Adult Criminal Justice Contact." NBER Working Paper No. 30274.

Roth, Jonathan. 2019. "Union Reform and Teacher Turnover: Evidence from Wisconsin's Act 10." Working Paper.

Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 135–46.

Sant'Anna, Pedro H.C., and Jun Zhao. 2020. "Doubly Robust Difference-in-Differences Estimators." *Journal of Econometrics* 219(1): 101–22.

Schneider, Pat. 2014. "MTI Gears Up for First-Ever Union Recertification Vote." *Capitol Times*, Oct 13, 2014. https://captimes.com/news/local/writers/pat_schneider/mti-gears-up-for-first-ever-union-recertification-vote/article_477d4c3c-ee77-5441-87ee-154a5099a742.html

Sojourner, Aaron J., Brigham R. Frandsen, Robert J. Town, David C. Grambowski, and Min M. Chen. 2015. "Impacts of Unionization on Quality and Productivity: Regression Discontinuity Evidence from Nursing Homes." *ILR Review* 68(4): 771–806.

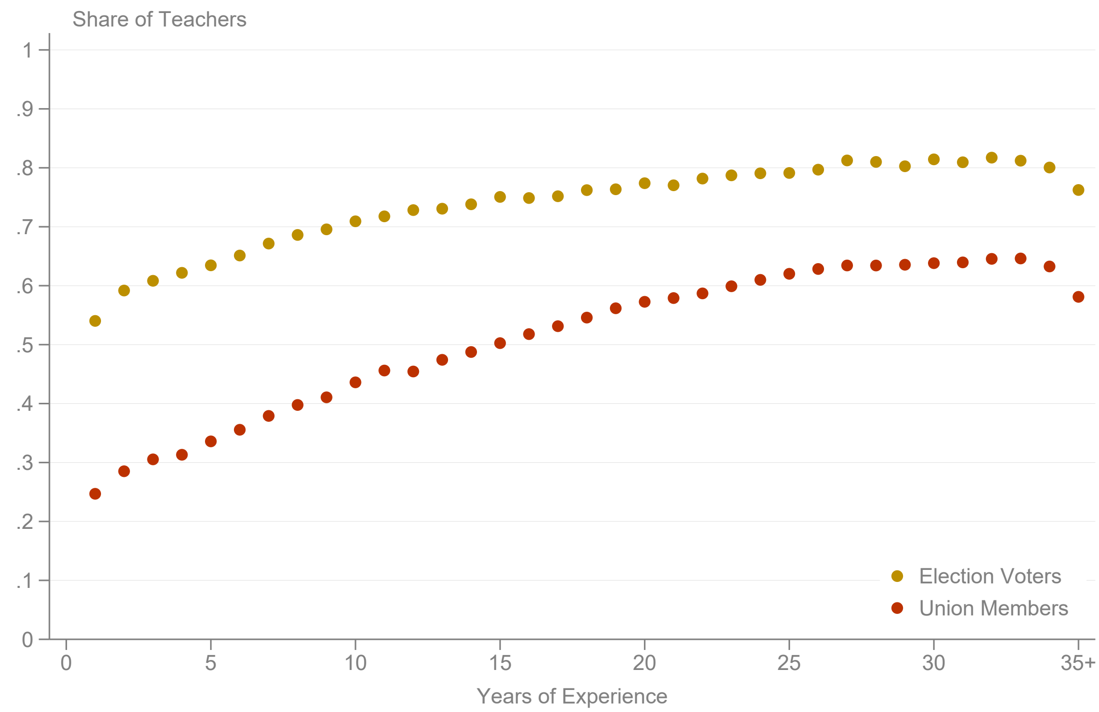
Strunk, Katharine O. 2011. "Are Teachers' Unions Really to Blame? Collective Bargaining Agreements and Their Relationships with District Resource Allocation and Student Performance in California." *Education Finance and Policy* 6(3): 354–98.

Sun and Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225(2): 175–99.

Warren, Paul, and Patrick Murphy. 2014. "Implementing the Common Core State Standards in California." Public Policy of Institute of California Report.

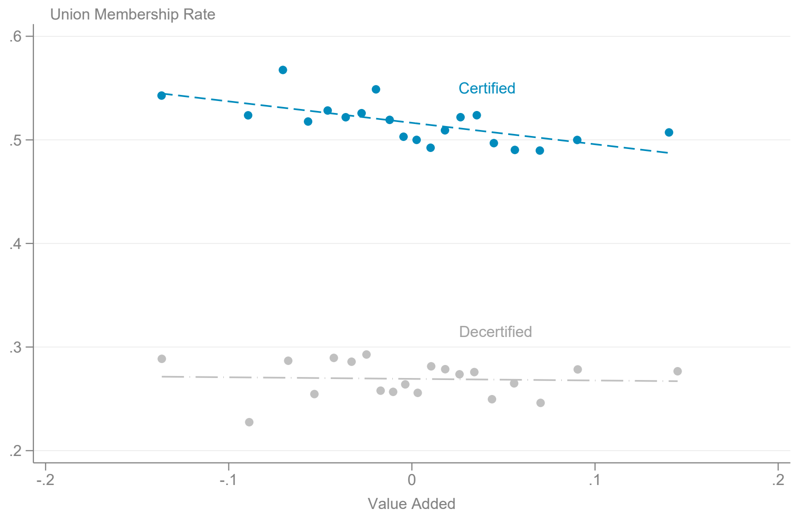
Wisconsin Legislative Council. 2011. "2011 Wisconsin Act 10 [January 2011 Special Session Assembly Bill 11]." https://docs.legis.wisconsin.gov/2011/related/lcactmemo/act010.

Figure 1: Union Membership and Union Vote Share by Experience

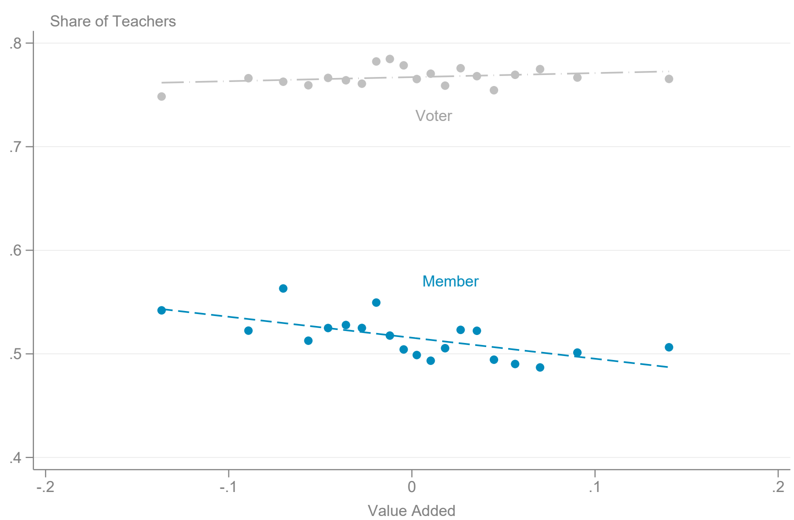


Note: The red dots show the estimated average union membership rate by each year of experience. The gold dots show the estimated average recertification election voting rate by each year of experience. The data in the plot is winsorized at 35 years of experience for clarity.

Figure 2: Union Membership and Union Vote Share by Value Added



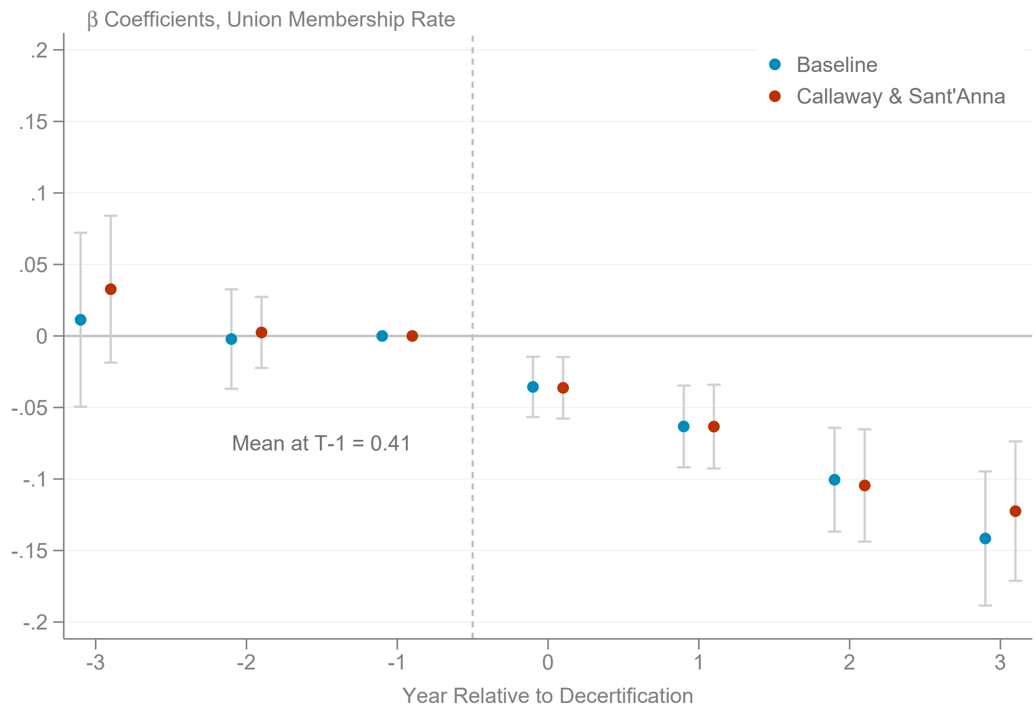
(a) Membership by Certification Status



(b) Union Support versus Membership

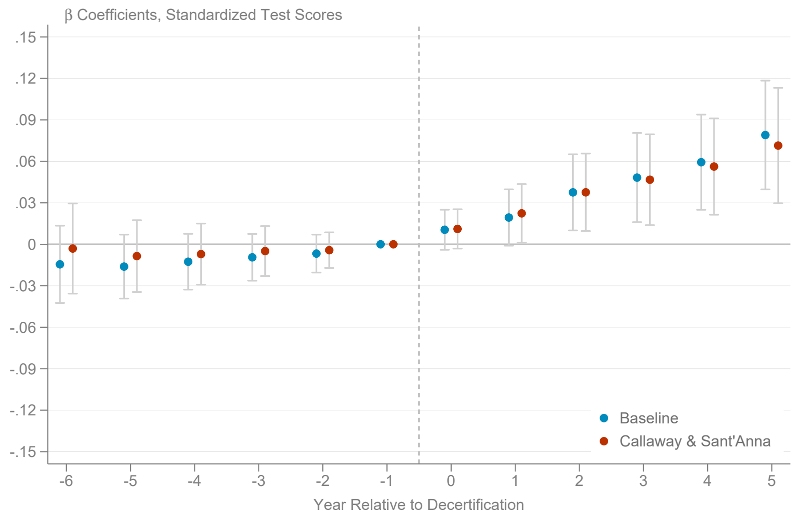
Note: The figures show binscatter estimates controlling for years of experience, education, school fixed effects, and district-by-year fixed effects as in Table 1. Panel A shows the average membership rate by value added, separated by whether a teacher was in a certified or a decertified district (as in Table 1 columns 3–4). Panel B contrasts membership vs. election voting by value added for the same set of teacher-by-year observations (as in Table 1 columns 5–6).

Figure 3: Did Decertification Affect Membership?

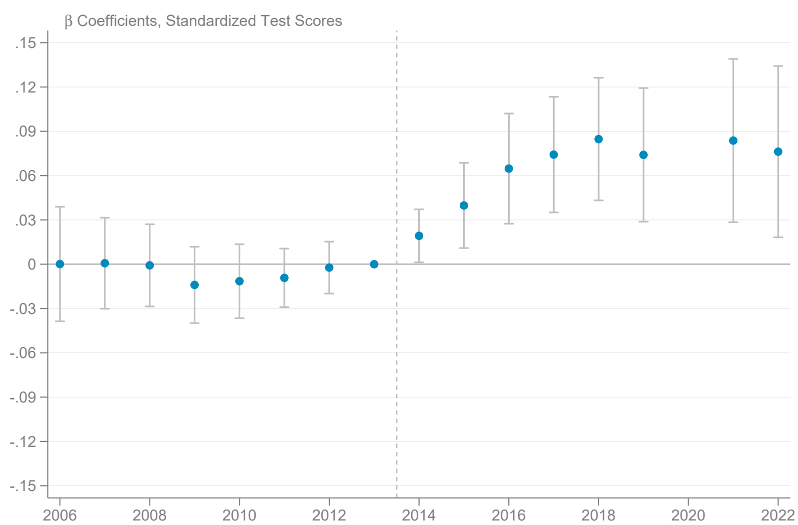


Note: The figure plots event study estimates of how decertification affected union membership. The figure uses data from 2016–2022 and using a treatment group of districts that decertified between 2017 and 2021. The coefficients are normalized relative to t-1, the school year before the union decertification event. For reference, the t+2 coefficients indicate that membership is roughly 10 percentage points lower than the year before decertification. The blue dots represent the baseline two-way fixed effects specification (binned at event time -3 and +3) and the red dots use the Callaway and Sant'Anna (2021) estimator. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure 4: Event Study of Standardized Test Scores



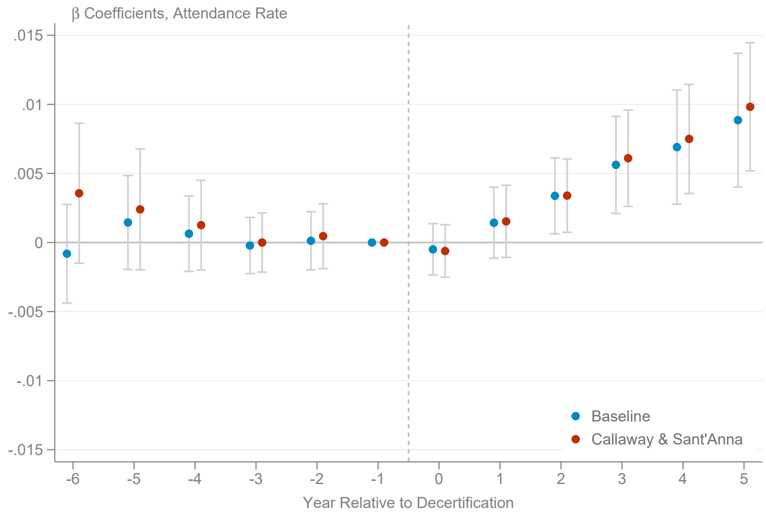
(a) All districts



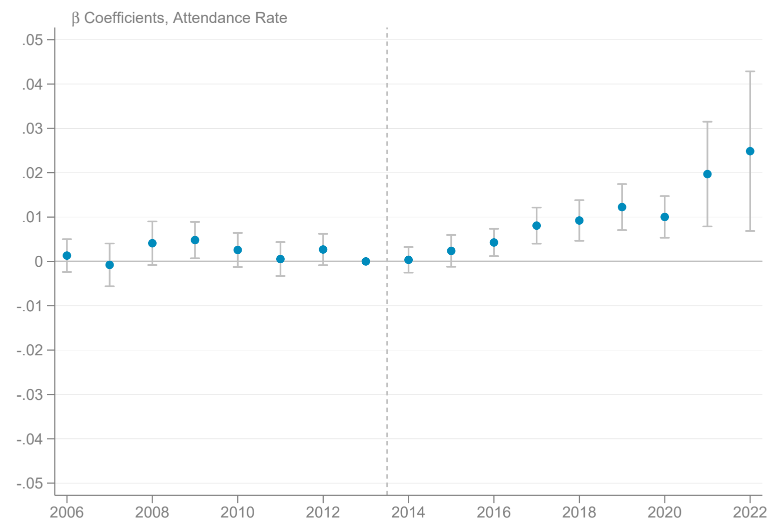
(b) 2014 Decertifiers v. Certifiers

Note: The figure plots event study estimates of how decertification affected average standardized test scores across all subjects. The coefficients are normalized relative to t-1, the school year before the union decertification event. Panel A shows results from the staggered difference-in-differences research design using all districts, where the blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. For reference, the t+4 coefficients indicate that standardized test scores are 6 percent of a standard deviation higher relative to the year before decertification. Panel B shows results from a simple difference-in-differences research design (i.e., not staggered), comparing the 2014 decertifiers to all certified districts. Panel A uses data from 2006–2019; panel B uses data from 2006–2022. There is no estimate in panel B for 2020 due to the COVID-19 pandemic. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure 5: Event Study of Attendance



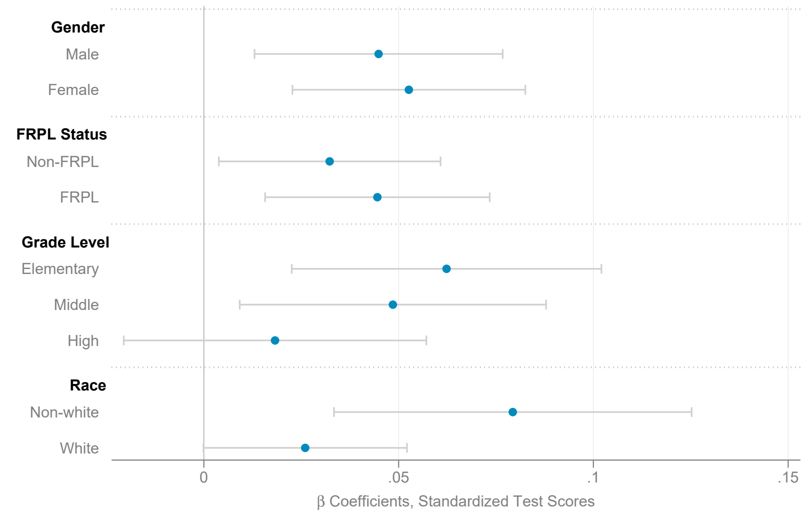
(a) All Districts



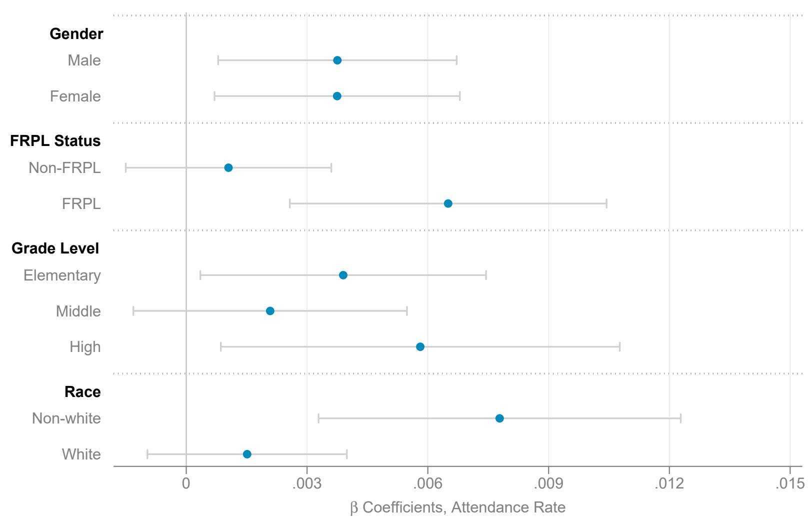
(b) 2014 Decertifiers v. Certifiers

Note: The figure plots event study estimates of how decertification affected student attendance rates. The coefficients are normalized relative to t-1, the school year before the union decertification event. Panel A shows results from the staggered difference-in-differences research design using all districts, where the blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. For reference, the t+5 coefficients indicate that attendance rates are roughly 1 percentage point higher relative to the year before decertification. Panel B shows results from a simple difference-in-differences research design (i.e., not staggered), comparing the 2014 decertifiers to all certified districts. Panel A uses data from 2006–2019; panel B uses data from 2006–2022. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure 6: Subgroup Heterogeneity



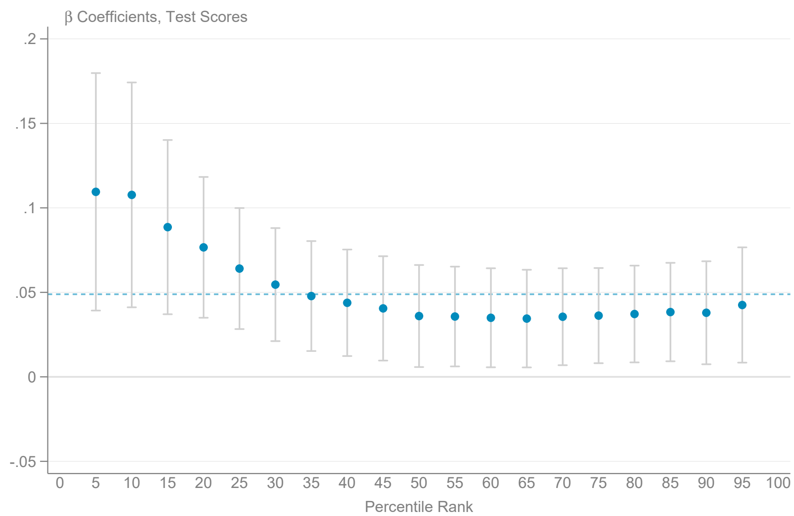
(a) Standardized Test Scores



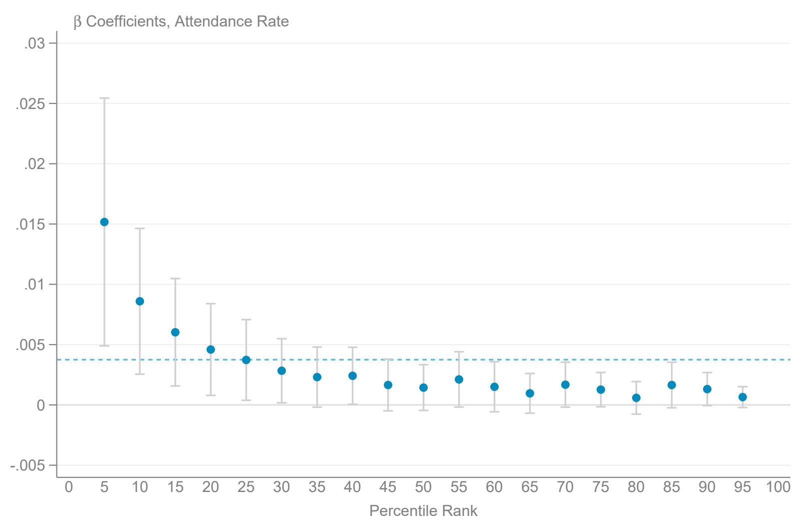
(b) Attendance Rate

Note: The figure presents difference-in-differences estimates of how decertification affected student outcomes by various subgroups. The outcome variable in panel A is average standardized test scores across all subjects; the outcome variable in panel B is student attendance rate. Both panels use data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure 7: Unconditional Quantile Treatment Effects



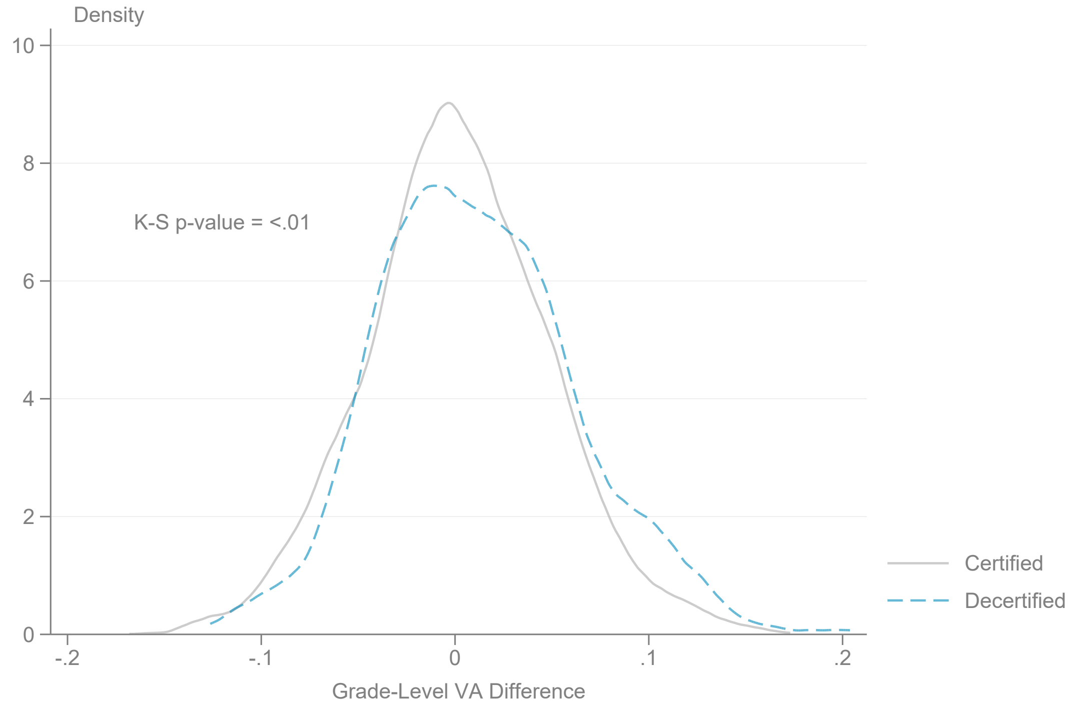
(a) Standardized Test Scores



(b) Attendance Rate

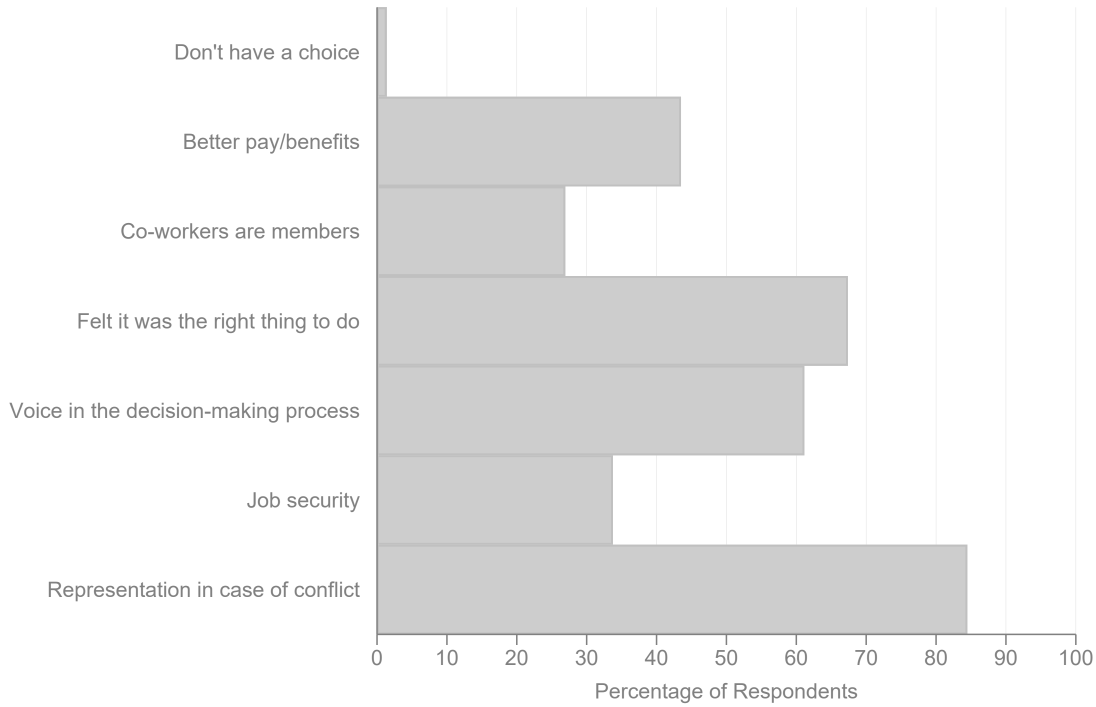
Note: The figure summarizes quantile treatment effect estimates at five percentile intervals. Each point represents a different "RIF" regression following Firpo, Fortin, and Lemieux (2009). The outcome variable in panel A is average standardized test scores across all subjects; the outcome variable in panel B is student attendance rates. The dashed horizontal lines represent the mean coefficient from Table 3 column 1. Both panels use data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure 8: Distribution of Teachers' Grade-Level Value-Added Gains



Note: The figure plots the distribution of grade-level value-added gains from 2008–2013 to 2014–2019 by certified (gray) or decertified (blue) status. For example, a teacher with a difference of 0.1 has a 2014–19 value added that is 0.1 units greater than their 2008–2013 score. The sample includes districts that decertified in 2014 and districts that always remained certified. Only teachers who taught in both the pre- and post-2014 period are included. In the top left corner, I report the p-value from a Kolmogorov–Smirnov test of equality between the two distributions.

Figure 9: Survey Evidence: Why Are You a Union Member?



Note: Survey respondents were asked "Why are/were you a member of the teachers' union? (Check all that apply)." The figure shows the percentage of respondents who checked each answer. The means are weighted to match the distribution of covariates observed in the full teacher sample. See Appendix Section E for survey details. N=1,368.

Table 1: Correlates of Union Membership and Recertification Election Voting

	(1) Member	(2) Member	(3) Member, Certified	(4) Member, Decertified	(5) Member	(6) Voter
Total Experience	0.013***	0.013***	0.013***	0.012***	0.013***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0004)
Advanced Degree	0.027***	0.027***	0.029**	0.023	0.028**	0.038***
	(0.010)	(0.010)	(0.012)	(0.017)	(0.012)	(0.007)
Value Added		-0.146**	-0.203***	-0.020	-0.197***	0.033
		(0.062)	(0.075)	(0.088)	(0.074)	(0.042)
Mean	0.44	0.44	0.52	0.27	0.52	0.77
Number of Teachers	11,941	11,941	9,142	4,764	9,196	9,196
Number of Districts	385	385	217	234	219	219
Observations	70,229	70,229	48,578	21,651	49,479	49,479

Note: The dependent variable in columns 1–4 is union membership. Columns 1–2 use a sample of all teachers with non-missing years of experience, education, and value added. The value-added variable uses the Bayes shrinkage estimator following Kane and Staiger (2008). One standard deviation of the distribution of value added scores is 0.074. Columns 3–4 use a sample of districts that were certified or decertified in year t, respectively. Columns 5–6 restrict the sample to district-years when a recertification election was held. The dependent variable in columns 5–6 are union membership and recertification election voting, respectively. All columns use data from 2016–2022. All regressions include school and district-by-year fixed effects. Standard errors are two-way clustered by school district and individual. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 2: Baseline Differences, Certified vs. Decertified Districts

	(1) Certified Mean	(2) Raw Difference	(3) Trend Difference
District			
Log Property Tax Revenue	16.01	-0.601***	-0.007
		(0.110)	(0.011)
Log Student Enrollment	7.39	-0.606***	-0.007
		(0.111)	(0.008)
Teacher			
Student-Staff Ratio	8.60	-0.016	0.179
		(0.192)	(0.135)
Log Salary	10.86	-0.088***	0.012
		(0.015)	(0.013)
Total Experience	14.77	0.486	-0.113
		(0.356)	(0.132)
Student			
Standardized Math Score	-0.06	0.152	-0.017
		(0.123)	(0.015)
Standardized ELA Score	-0.06	0.158	-0.003
		(0.113)	(0.009)
Attendance Rate	0.94	0.011	-0.003
		(0.008)	(0.003)
Suspension Rate	0.07	-0.042*	0.008
		(0.023)	(0.008)
Share English Proficient	0.07	-0.047***	-0.004
		(0.010)	(0.002)
Share Special Education	0.14	-0.008	0.001
		(0.011)	(0.008)
Share Free/Reduced Price Lunch	0.42	-0.095	0.015
		(0.059)	(0.011)
Share Black	0.14	-0.127*	0.001
		(0.068)	(0.003)
Share Hispanic	0.12	-0.071***	-0.008***
		(0.020)	(0.002)
Share Female	0.48	-0.002*	0.001
		(0.001)	(0.002)

Note: Each row shows a separate baseline variable. Column 1 presents the mean value as of 2011 in certified districts. Column 2 shows the point estimate from regressing each variable on an indicator for whether the school district's union decertified at some point from 2012–2022. Column 3 shows point estimates from an analogous regression, but where the outcome variable is the change in each variable from 2008 to 2011. Standard errors are clustered by school district. *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 3: Decertification Effect on Student Outcomes

	(1)	(2)	(3)	(4)
Panel A: Average Test Scores				
Post-Decertification	0.049***	0.041***	0.039***	0.036***
	(0.015)	(0.013)	(0.014)	(0.012)
5+ Years Post-Decertification	0.079***	0.071***	0.061***	0.059***
	(0.020)	(0.021)	(0.018)	(0.019)
Observations	5,004,965	5,004,965	5,594,469	5,594,469
Panel B: Attendance Rate				
Post-Decertification	0.004**	0.004***	0.004***	0.004***
	(0.002)	(0.001)	(0.001)	(0.001)
5+ Years Post-Decertification	0.009***	0.010***	0.007***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
CS Estimator	No	Yes	No	Yes
ITT Sample	No	No	Yes	Yes
Number of Districts	322	322	357	357
Observations	10,470,810	10,470,810	11,690,199	11,690,199

Note: This table summarizes how decertification affected either student standardized test scores across all subjects (Panel A) or student attendance rates (Panel B). The "post-decertification" term signifies a regression of the outcome variable on the indicator for whether a district was decertified in a given year. The second row "5+ Years Post-Decertification" presents results from a separate regression that replaces the post-decertification term with the full set of event time coefficients and indicates the coefficient on the t+5 indicator as in Figures 4 and 5. Columns 1 and 3 use the baseline two-way fixed effects estimator; columns 2 and 4 use the Callaway and Sant'Anna (2021) estimator. Columns 1–2 use the primary set of districts that either never decertified or decertified once over the sample period. Columns 3–4 use the ITT sample, which includes the recertifier districts. All specifications include district and year fixed effects and use data from 2006–2019. Standard errors are clustered by school district. *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table 4: Survey Evidence: Teacher Perceptions on Unions

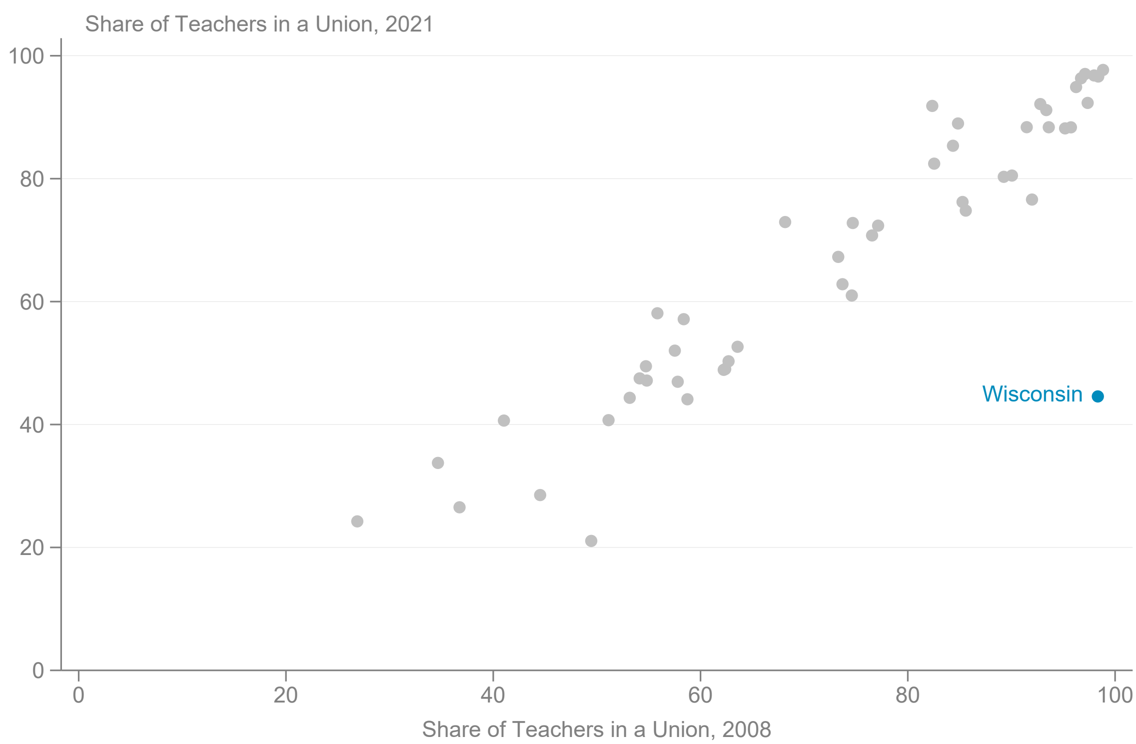
	(1)	(2)	(3)	(4)
	Workers Should Have the Right to Form Unions	Right to Representation		
		All	Members	Non-Members
Decertified	-0.045***	-0.340***	-0.170***	-0.157***
	(0.016)	(0.045)	(0.042)	(0.028)
Observations	2,280	2,273	1,153	1,120
Certified mean	0.94	0.67	0.90	0.31
FDR q-values	0.006	0.001	0.001	0.001
Number of Districts	264	264	199	234

Note: The column 1 dependent variable is whether the respondent believes that workers should have the right to form unions. The dependent variable in columns 2–4 is whether the respondent believes that they personally have the right to union representation in case of a disciplinary proceeding. Column 2 uses the entire sample, column 3 is a sample of those individuals who reported being members in the past five years, column 4 is a sample of those who did not report being members in the past five years. The regressions are weighted to match the distribution of covariates observed in the full teacher sample. See Appendix Section E for survey details. False-discovery rate q-values (Anderson 2008) are reported. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels according to the adjusted q-values, respectively.

Appendix

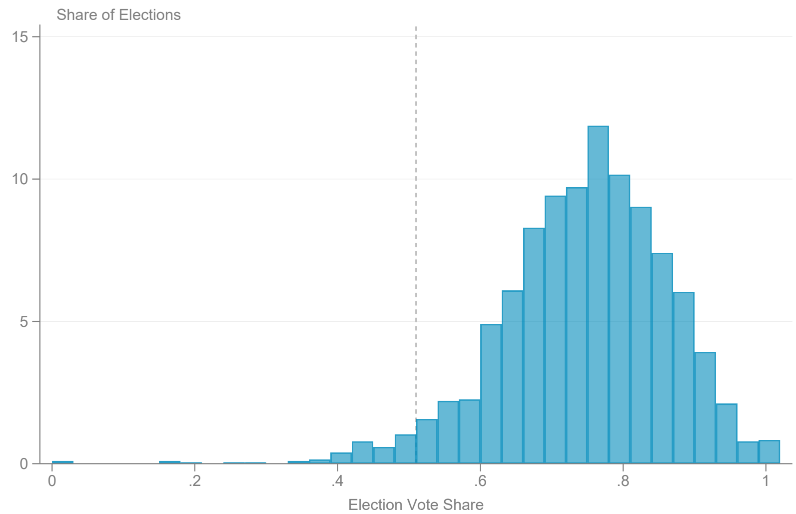
A. Additional Figures/Tables

Figure A1: Teachers in a Union by State, 2008 vs. 2021

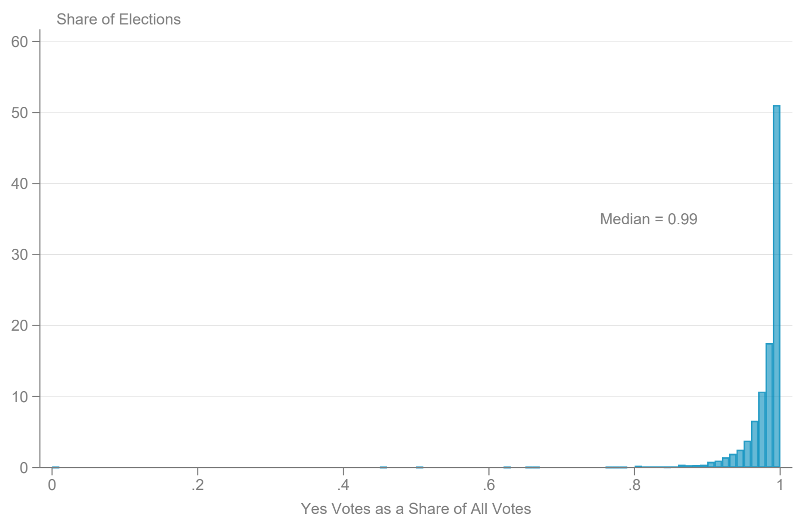


Note: The x-axis is the percentage of public school teachers in a union or employees' association from the 2007–08 Schools and Staffing Survey (National Center for Education Statistics 2008). The y-axis is the percentage of public school teachers in a union or employees' association from the 2020–21 National Teachers and Principal Survey (National Center for Education Statistics 2021). Each point represents a different US state.

Figure A2: Histogram of Recertification Election Vote Shares



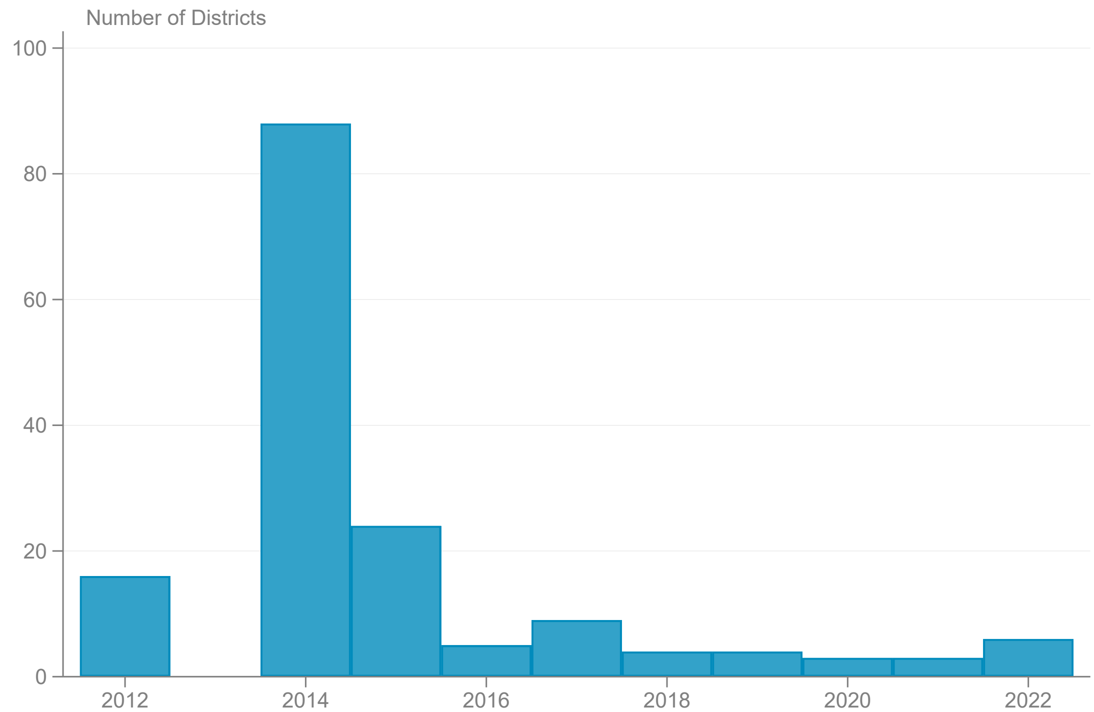
(a) Election Vote Shares



(b) Yes Votes as a Percentage of All Votes

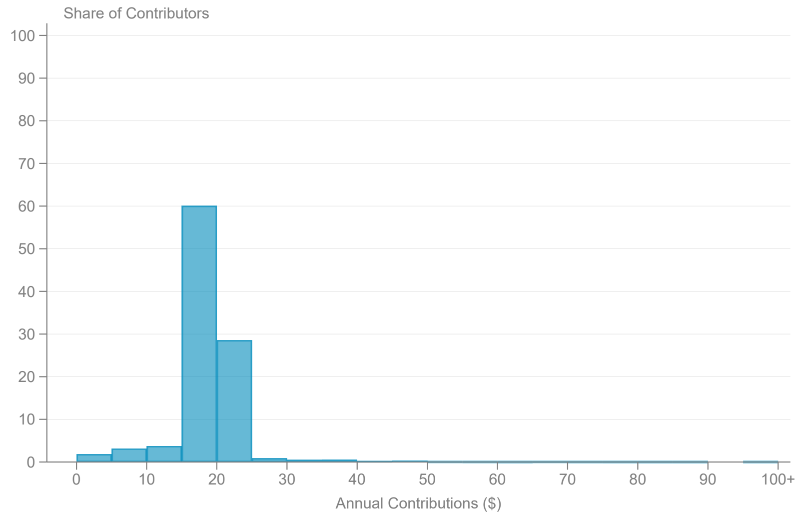
Note: Panel A plots the distribution of recertification election vote shares broken into 3 percentage point bins. The dashed line indicates the 51 percent threshold for recertification. Panel B plots the yes votes as a share of all votes broken into one percentage point bins.

Figure A3: Number of Decertification Events by Year

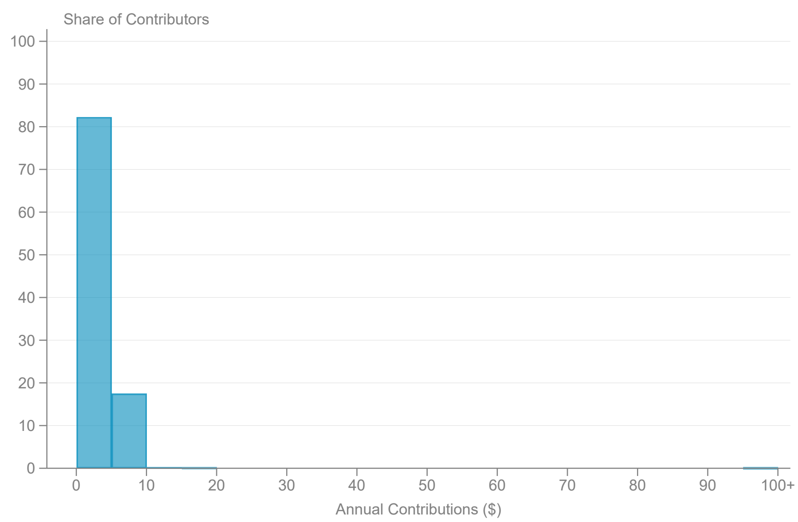


Note: This figure plots the number of decertification events by year for those districts that decertified and never recertified. There were no elections in 2013 due to the legal challenges described in Section III.

Figure A4: Distribution of Individual PAC Campaign Contributions



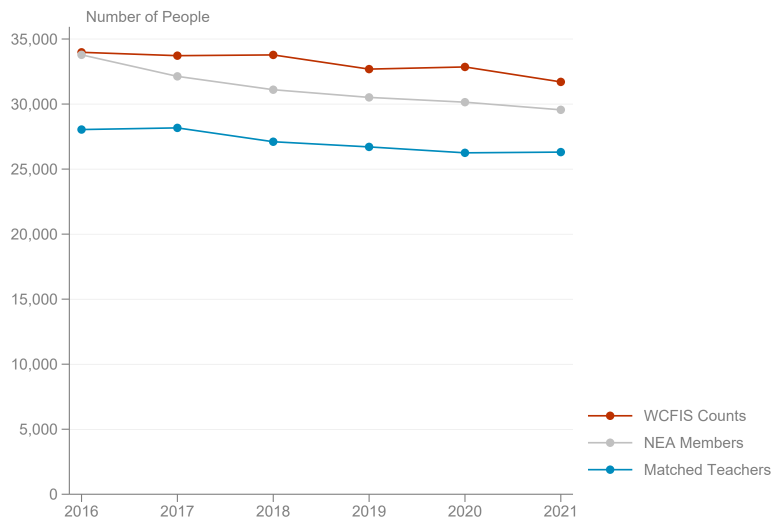
(a) State PAC Contributions



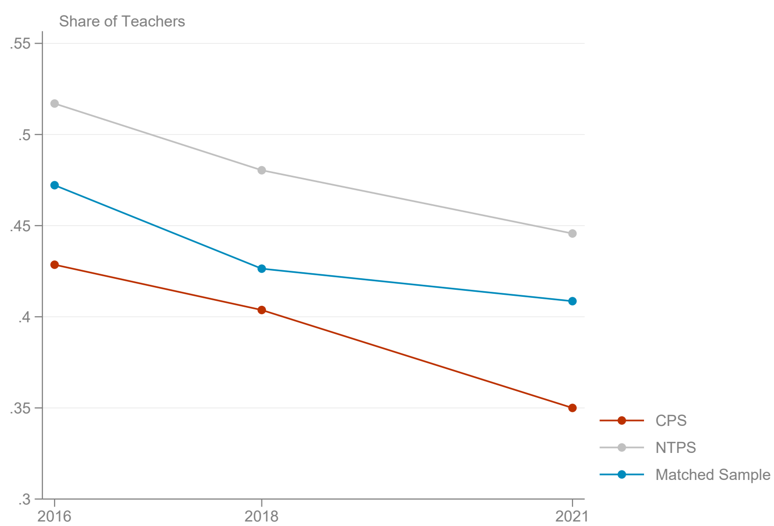
(b) Regional PAC Contributions

Note: The figure plots the distribution of individual annual contributions to the state union's PAC broken into $5 bins. Panel A shows contributions to the state PAC; panel B shows contributions to the regional chapter PACs. For clarity, this plot does not use data from Madison or Milwaukee because they sometimes use a different contribution amount than the rest of the state.

Figure A5: Estimated Membership Data Relative to Other Sources



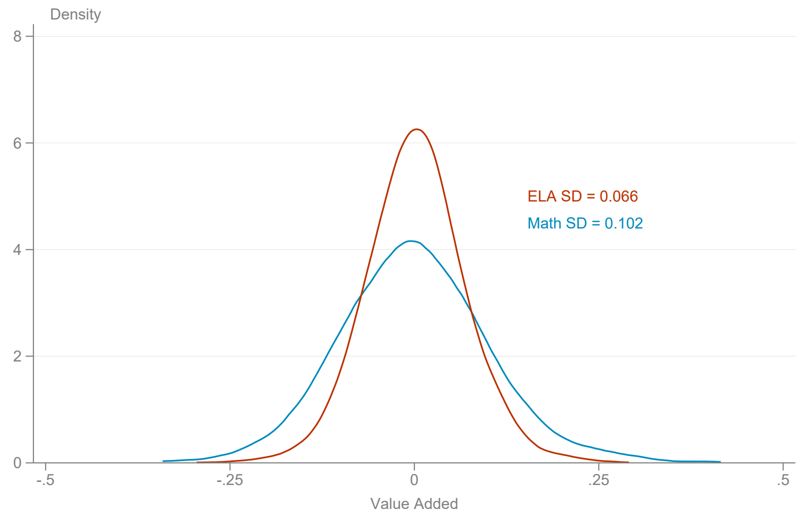
(a) Comparison to NEA Counts



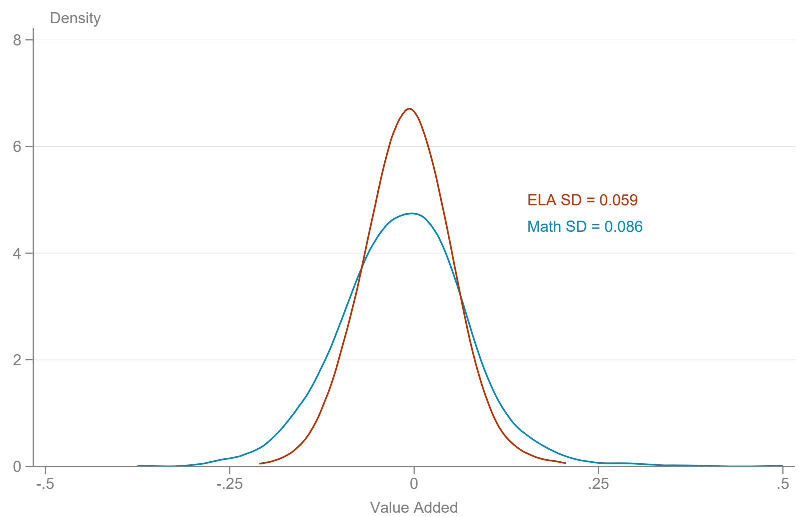
(b) Comparison to Survey Data

Note: Panel A plots the number of National Education Association members in Wisconsin. The gray line shows counts of active members from annual NEA financial reports. The blue line shows the number of people in the DPI administrative data who were matched to the campaign contributions data. The red line shows the number of unique people in the campaign contributions data, regardless of whether they were matched or not. Panel B plots the share of teachers who are estimated to be union members (blue line), relative to the estimate from the National Teachers and Principals Survey from the National Center for Education Statistics (gray line) and my calculations using the Current Population Survey (red line). The estimates from the Current Population Survey group the current and previous survey year together to more closely resemble the school year schedule. For instance, the 2016 estimate in the figure uses 2015 and 2016 responses.

Figure A6: Distribution of Teacher Value Added by Subject



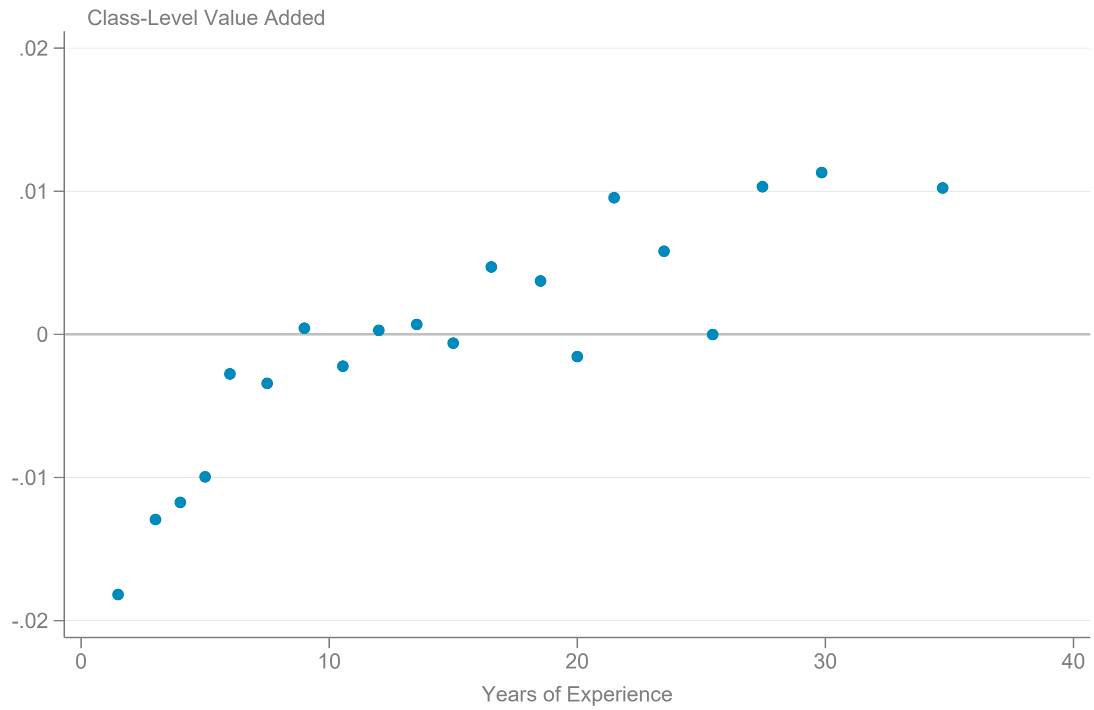
(a) Elementary School Teachers



(b) Middle School Teachers

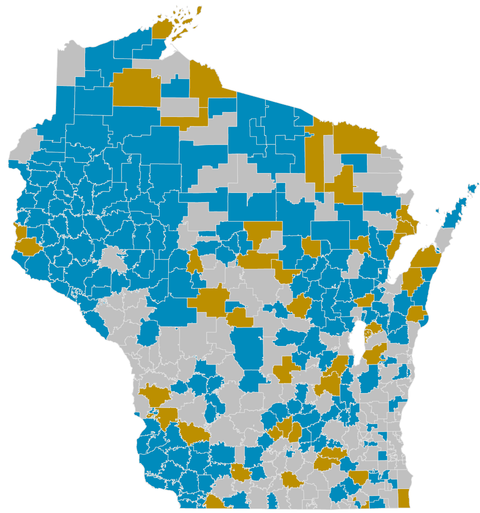
Note: The figures plot the empirical distribution of value-added estimates for elementary school teachers (panel A) and middle school teachers (panel B). The blue lines represent the distribution of math value added; the red lines are the distribution of ELA value added. The densities use a bandwidth of 0.02. Note that the reported standard deviations are smaller than those cited in the text because they reflect the standard deviation of the empirical distribution after the shrinkage procedure.

Figure A7: Classroom Value Added by Experience



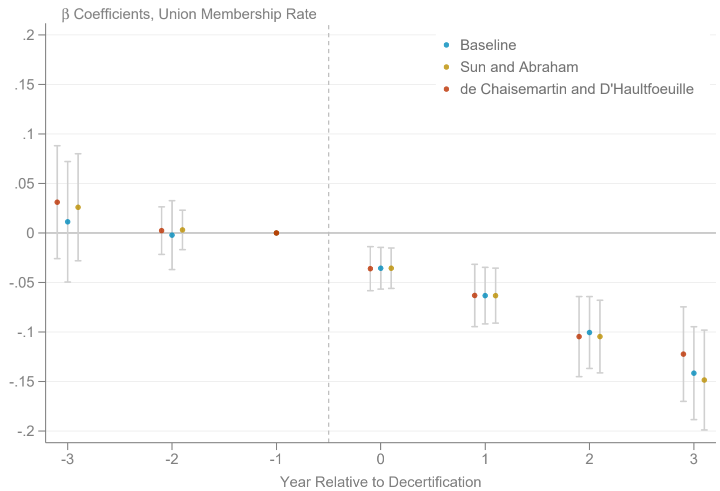
Note: This figure plots a binned scatterplot of classroom value added (y-axis) by total years of experience (x-axis). Classroom value added is calculated over 2017–19 and using the shrunken estimator as described in Section IV. Years of experience is the highest level of experience observed from 2017–19.

Figure A8: Geographic Distribution of Decertifications

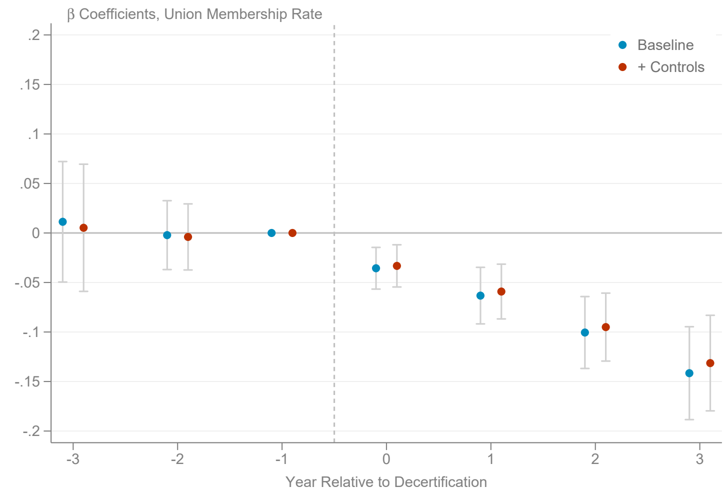


Note: Each polygon is a school district. Districts with decertified unions are in blue; districts with certified unions are in gray. Districts that decertified and later recertified are in gold.

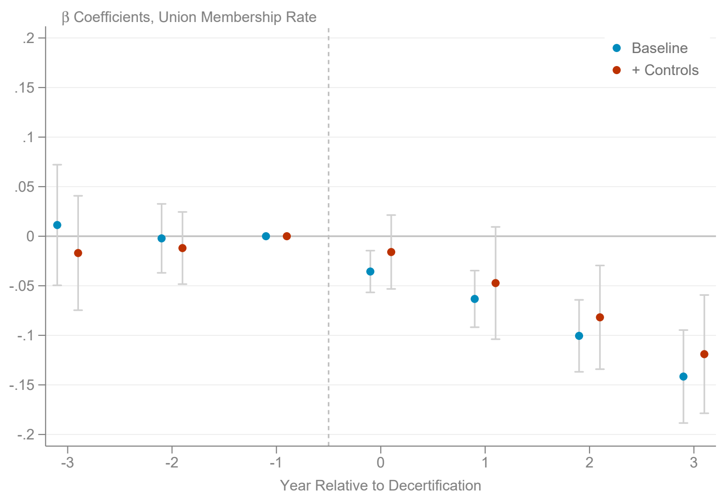
Figure A9: Event Study of Membership, Robustness



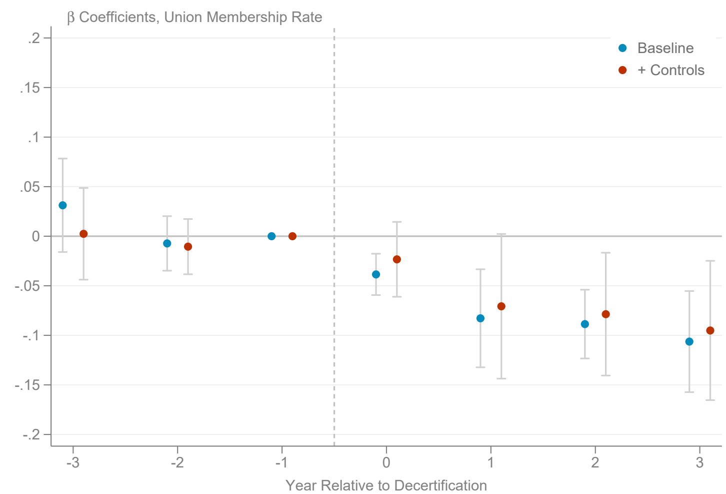
(a) Robust TWFE Estimators



(b) + Individual Controls



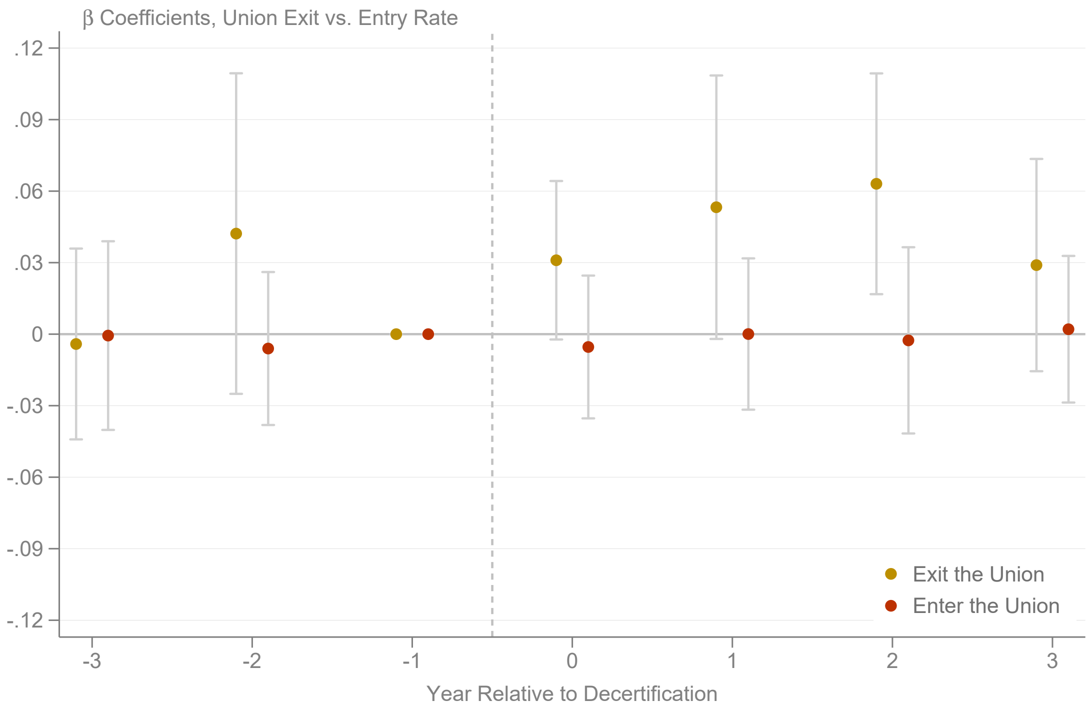
(c) + Vote Share Controls



(d) All Districts

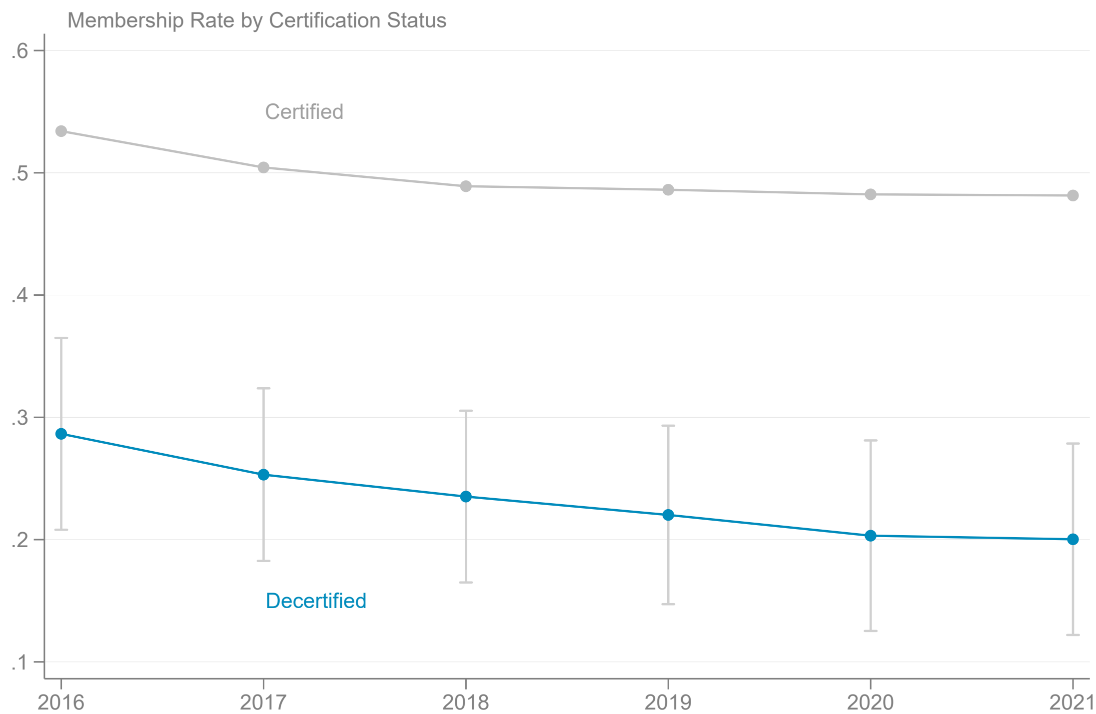
Note: All panels show an event study of how decertification affected union membership as in Figure 3. Panel A adds the Sun and Abraham (2021) and the de Chaisemartin and D'Haultfoeuille (2020) estimators. Panel B shows the baseline estimates (blue) relative to a specification that adds controls for teacher experience and education. Panel C adds controls for a cubic polynomial in the recertification election vote share. Panel D uses all districts, including those that have no pre- or post-decertification data. All panels use data from 2016–2022. The coefficients are normalized relative to the school year before the union decertification event. All specifications are binned at event time -3 and +3. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A10: Union Membership: Exit vs. Entry Rates



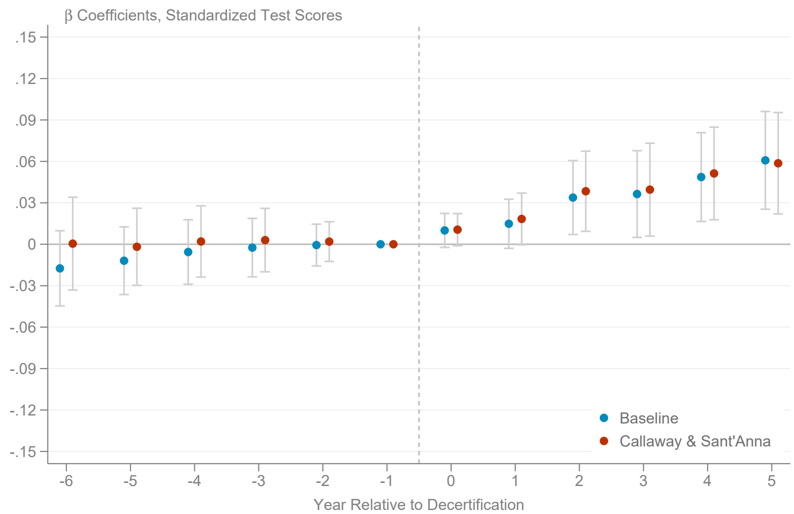
Note: The yellow dots show the share of former members (as of t-1) who exited the union (in year t). The red dots show the rate of new sign-ups in year t, where the sample is either new hires or people who were not members in year t-1. Event time is binned at three years before the event and three years after the event. The figure uses data from 2017–2022 (i.e., excluding the 2016 data as it is unknown whether someone was a member in 2015). The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A11: Trends in Membership Rates by Certification Status

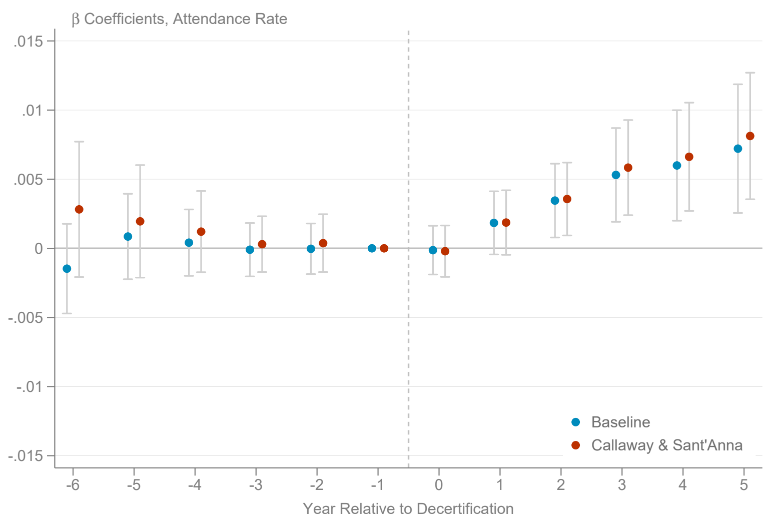


Note: This figure presents the results of a regression of union membership on decertification status. The gray line is the constant in the regression, equivalent to the average membership rate in certified districts. The blue line is the constant plus the regression coefficient, equivalent to the average membership rate in decertified districts. The sample of districts includes all certified districts and districts that decertified before 2016. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A12: Event Study Results, ITT Sample



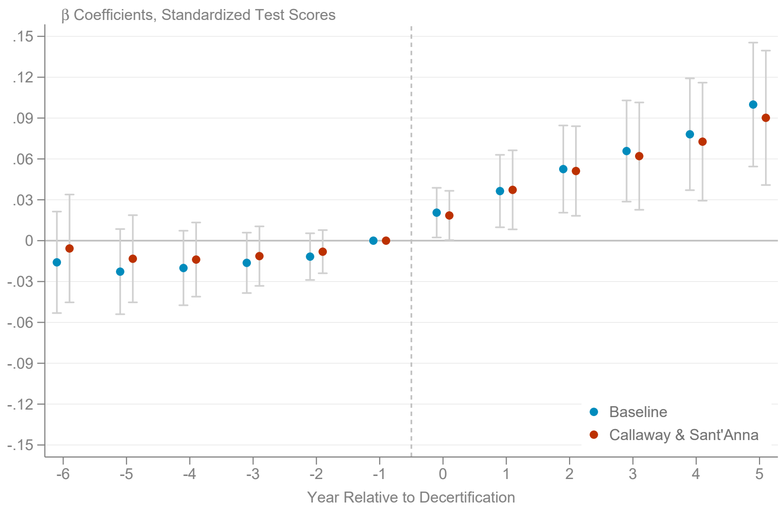
(a) Test Scores



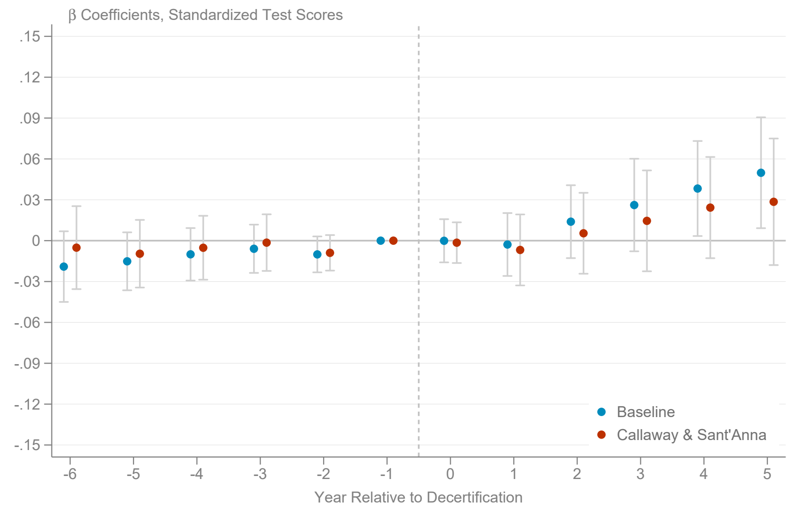
(b) Attendance Rate

Note: This figure replicates Figures 4–5 but using the ITT sample that includes districts that decertified and later recertified. The coefficients are normalized relative to the school year before the union decertification event. The blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. The figure uses data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

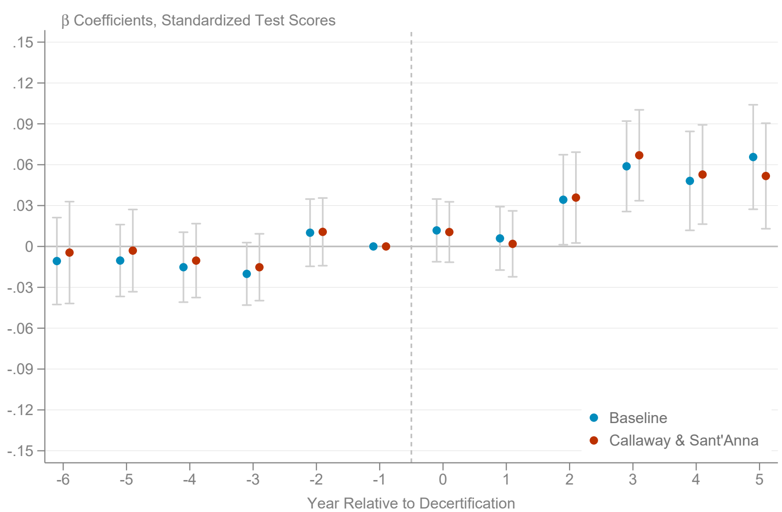
Figure A13: Event Study of Test Scores by Subject



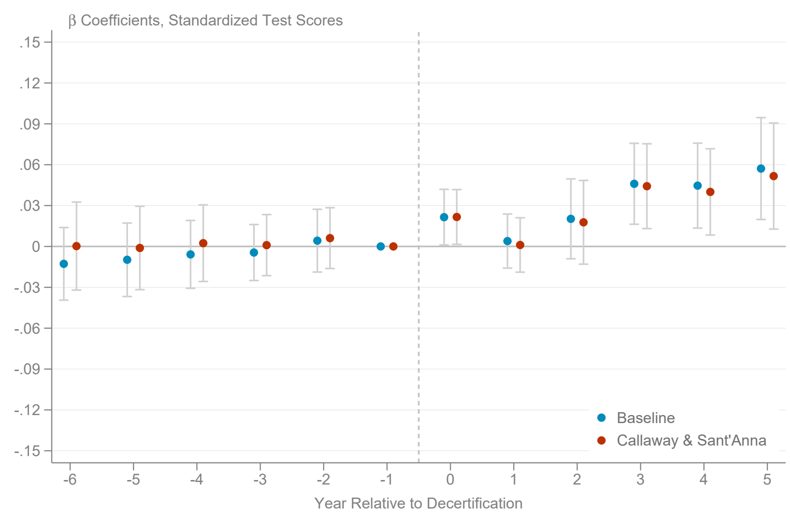
(a) Math



(b) ELA



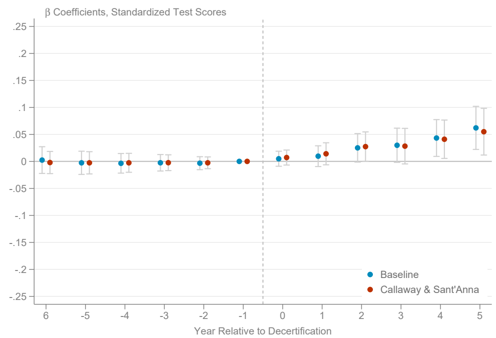
(c) Science



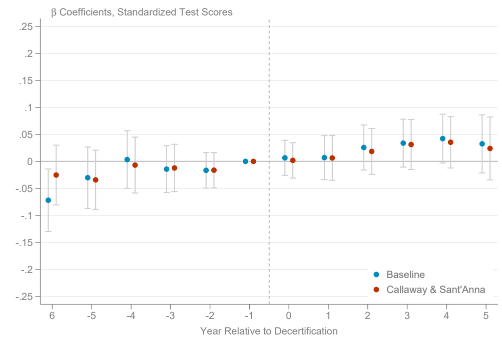
(d) Social Studies

Note: The figure reproduces Figure 4 but breaking out the estimates by test subject. The subject in Panels A–D are math, ELA, science, and social studies, respectively. The coefficients are normalized relative to the school year before the union decertification event. The blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. All panels use data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

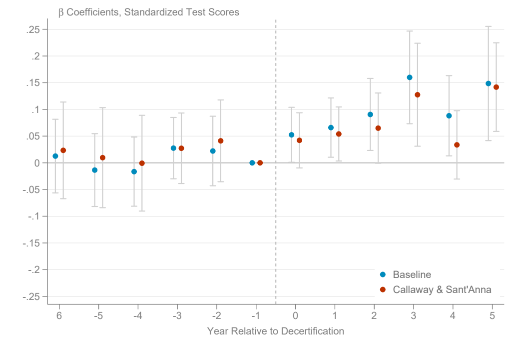
Figure A14: Main Results Disaggregated by Race



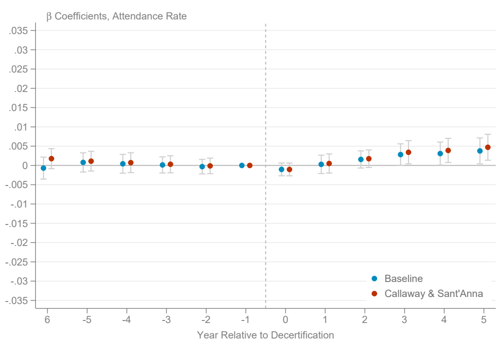
(a) Test Scores, White Students



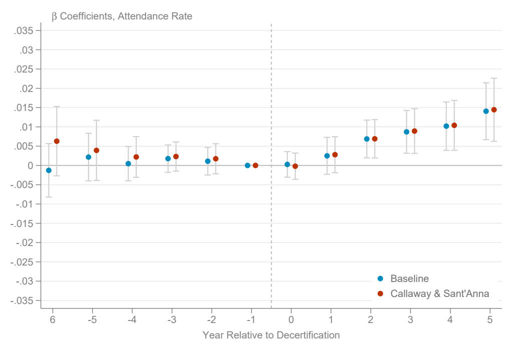
(b) Test Scores, Hispanic Students



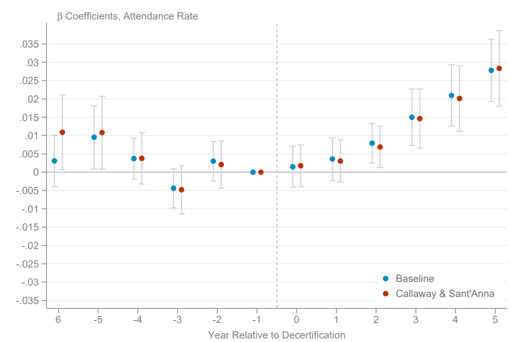
(c) Test Scores, Black Students



(d) Attendance Rates, White Students



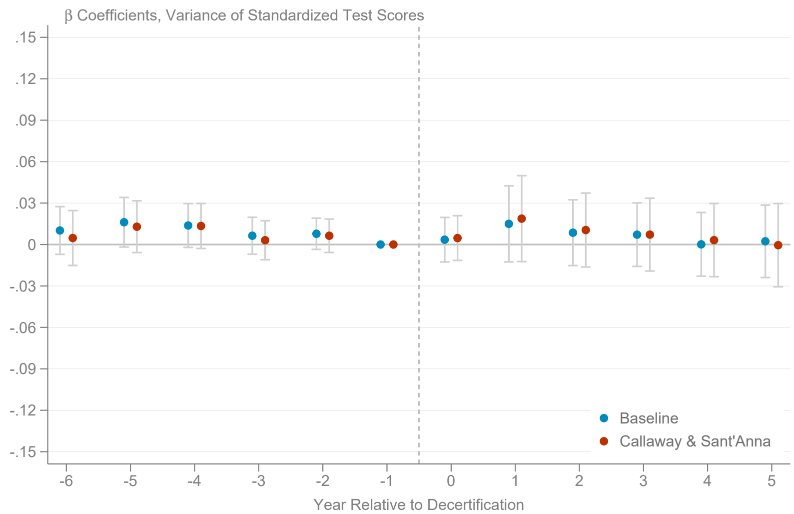
(e) Attendance Rates, Hispanic Students



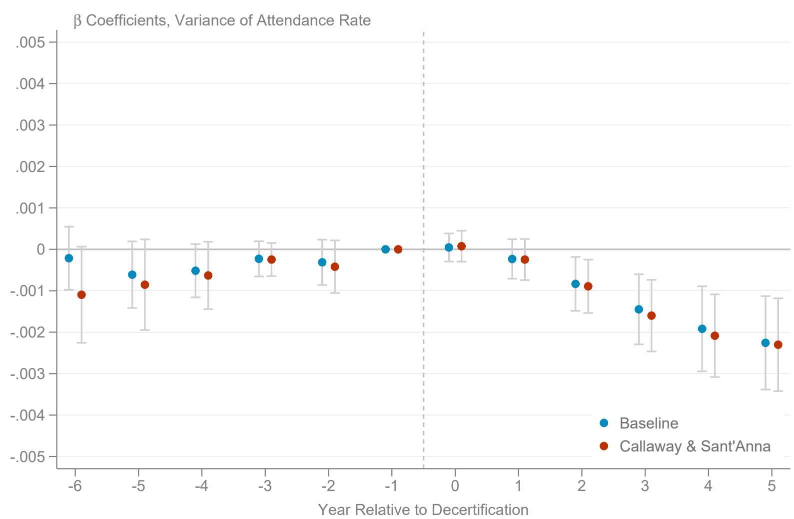
(f) Attendance Rates, Black Students

Note: The dependent variable in panels A–C is student test scores across all subjects. The dependent variable in panels D–F is student attendance rates. Panels A and D use a sample of white students; panels B and E use a sample of Hispanic students; panels C and F use a sample of Black students. The coefficients are normalized relative to the school year before the union decertification event. The y-axes are consistent across panels but wider than in Figures 4–5 to accommodate the relatively wider confidence intervals. The blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. All panels use data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A15: Event Study of Outcome Variance



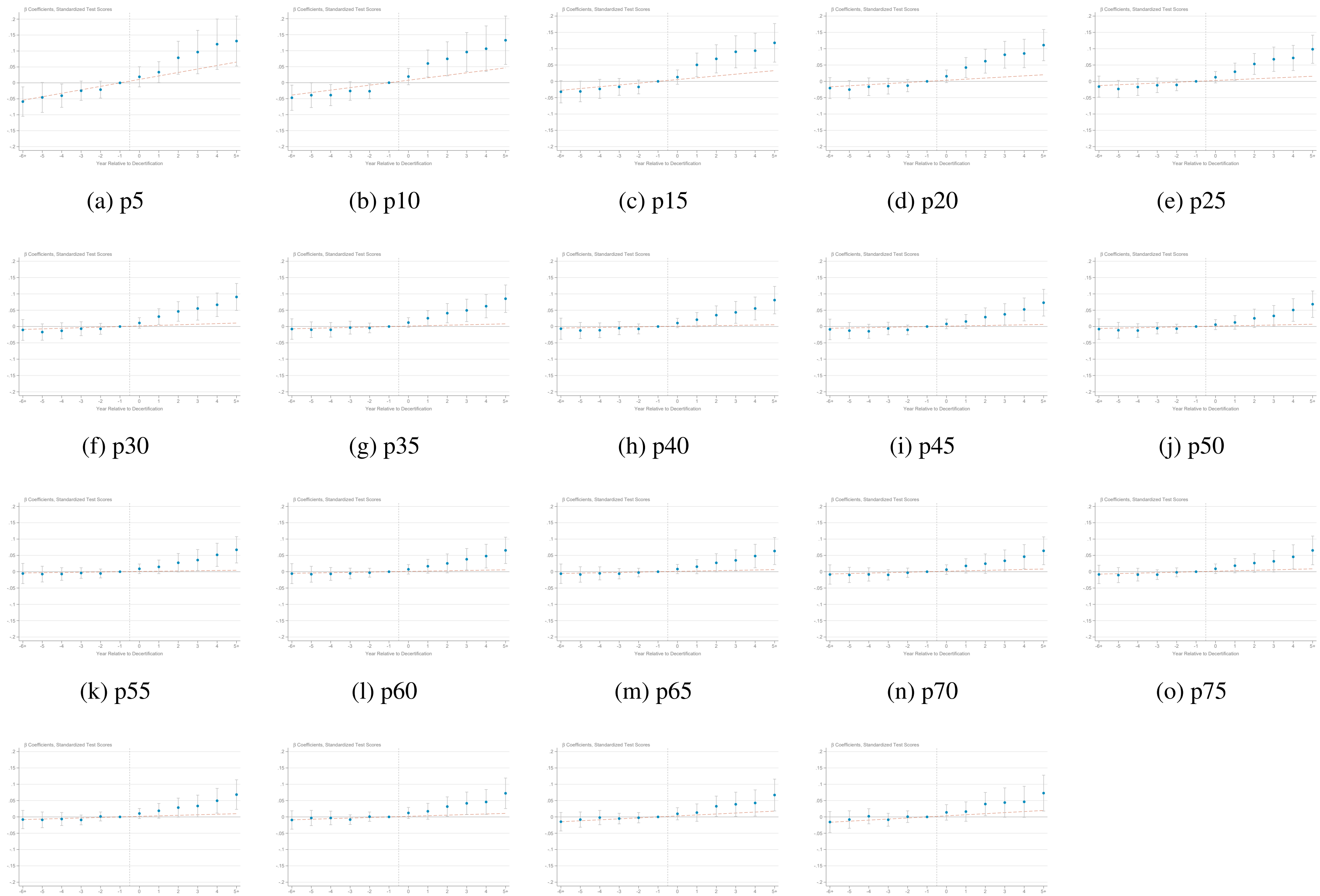
(a) Test Scores



(b) Attendance Rate

Note: The dependent variable is the variance in student test scores in panel A and the variance in student attendance rate in panel B. The coefficients are normalized relative to the school year before the union decertification event. The blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. The figure uses data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A16: Test Score Event Studies by Quantile



(a) p5 (b) p10 (c) p15 (d) p20 (e) p25

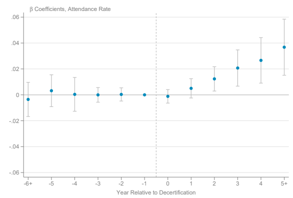
(f) p30 (g) p35 (h) p40 (i) p45 (j) p50

(k) p55 (l) p60 (m) p65 (n) p70 (o) p75

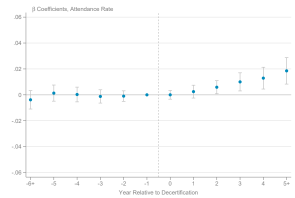
(p) p80 (q) p85 (r) p90 (s) p95

Note: Each panel reports the event study coefficients from an unconditional quantile "RIF" regression as in Firpo, Fortin, and Lemieux (2009) at different achievement percentiles. The dependent variable is average standardized test scores. The coefficients are normalized relative to the school year before the union decertification event. Event time is binned at six years before the event and five years after the event. Each figure uses data from 2006–2019. I superimpose a linear pre-trend (in red) to clearly see the deviation from the trend for the lowest percentile plots. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

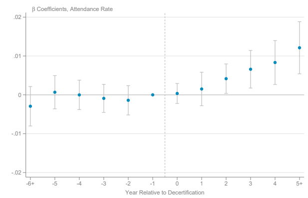
Figure A17: Attendance Event Studies by Quantile



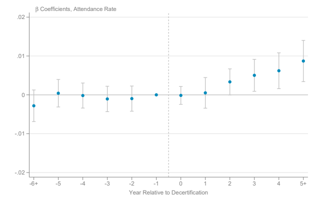
(a) p5



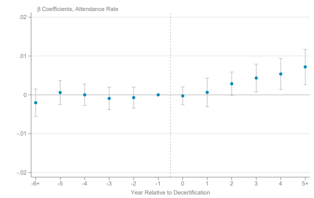
(b) p10



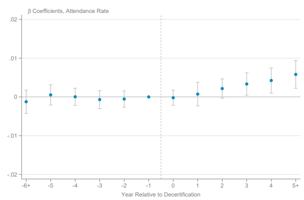
(c) p15



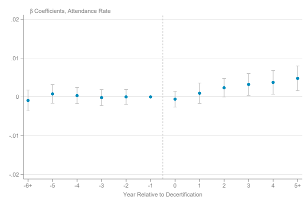
(d) p20



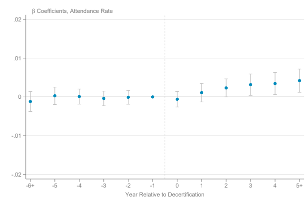
(e) p25



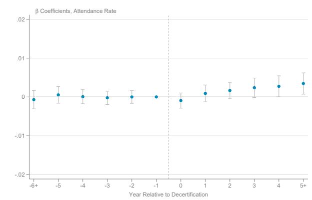
(f) p30



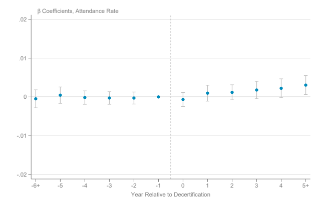
(g) p35



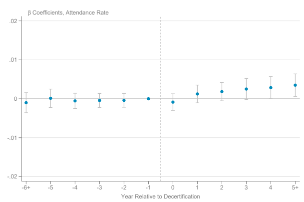
(h) p40



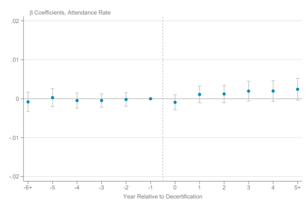
(i) p45



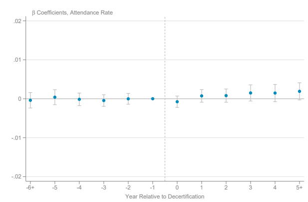
(j) p50



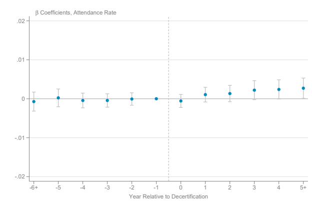
(k) p55



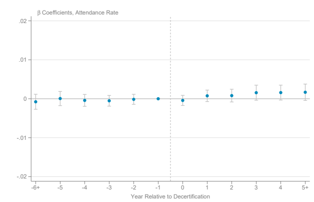
(l) p60



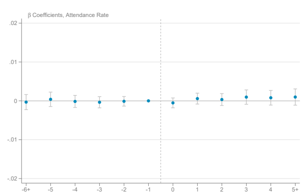
(m) p65



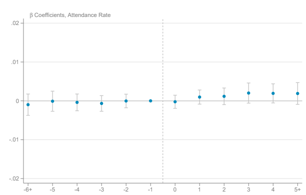
(n) p70



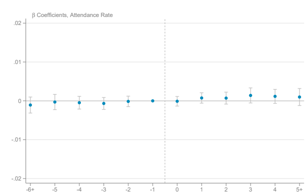
(o) p75



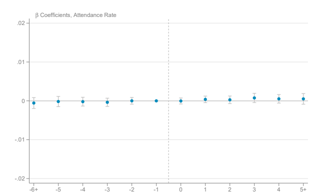
(p) p80



(q) p85



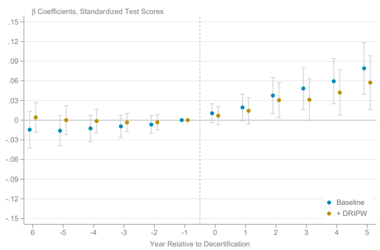
(r) p90



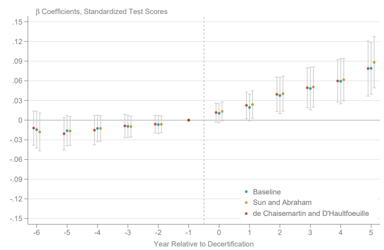
(s) p95

Note: Each panel reports the event study coefficients from an unconditional quantile "RIF" regression as in Firpo, Fortin, and Lemieux (2009) at different achievement percentiles. The dependent variable is student attendance rate. All panels use the same y-axis except for percentiles 5 and 10 as the treatment effects are too large for a consistent scale. The coefficients are normalized relative to the school year before the union decertification event. Event time is binned at six years before the event and five years after the event. Each figure uses data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A18: Standardized Test Scores, Robustness to Alternative Specifications

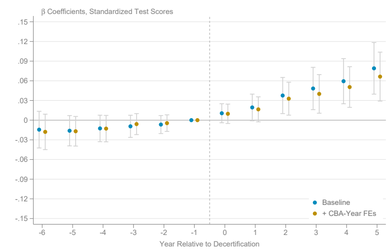


(a) DRIPW weights

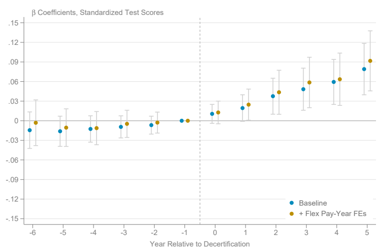


(b) TWFE Estimators

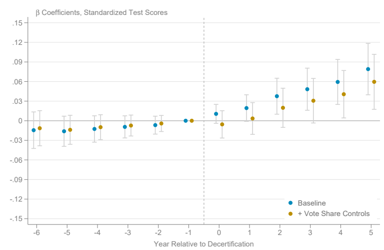
(c) + Missing Districts



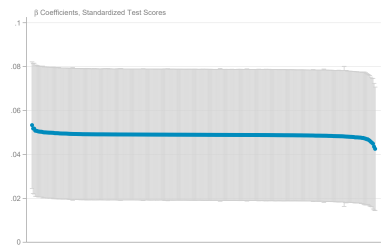
(d) + CBA FEs



(e) + Flex Pay FEs



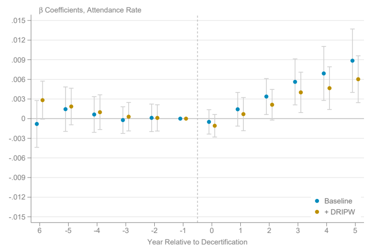
(f) + Election Vote Shares



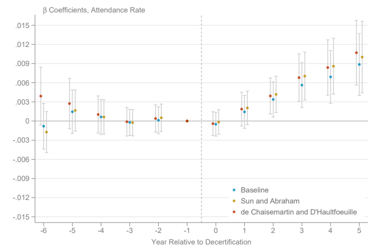
(g) Leave-One-Out Estimates

Note: The dependent variable in all panels is student average standardized test scores across all subjects. Panel A uses the Callaway and Sant'Anna (2021) estimator with doubly-robust inverse propensity weights, adjusting for student demographic characteristics including race (Black, Hispanic, Asian, other), FRPL status, limited English status, and gender. Panel B adds the estimators from de Chaisemartin and d'Haultfoeuille (2020) and Sun and Abraham (2021) to the baseline model (in blue). Panel C adds in the missing districts, assigning the event year as 2012 (see Section VI). Panel D uses CBA-year fixed effects as categorized in Biasi and Sarsons (2022). Panel E adds flexible pay-year fixed effects as categorized in Biasi (2021). Panel F controls for a cubic polynomial in the recertification election vote share. The event study coefficients are normalized relative to the school year before the union decertification event. Event time is binned at six years before the event and five years after the event. Panel G plots 322 different regression estimates from equation 1, successively dropping each district at a time. All panels use data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

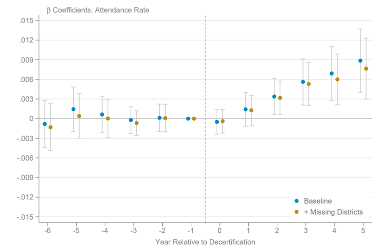
Figure A19: Attendance Rate, Robustness to Alternative Specifications



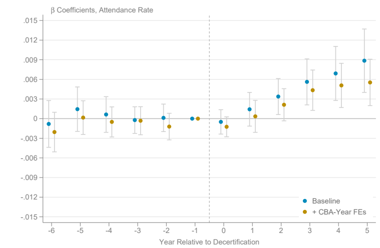
(a) DRIPW weights



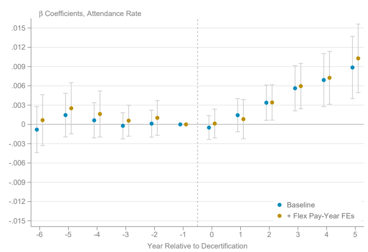
(b) TWFE Estimators



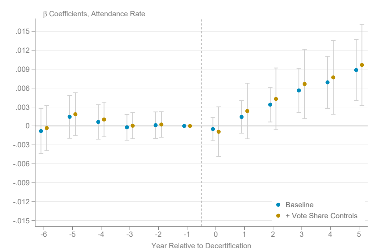
(c) + Missing Districts



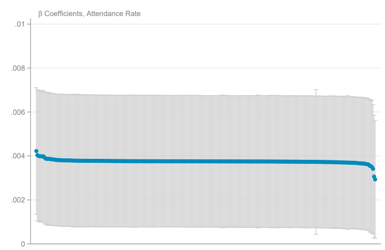
(d) + CBA FEs



(e) + Flex Pay FEs



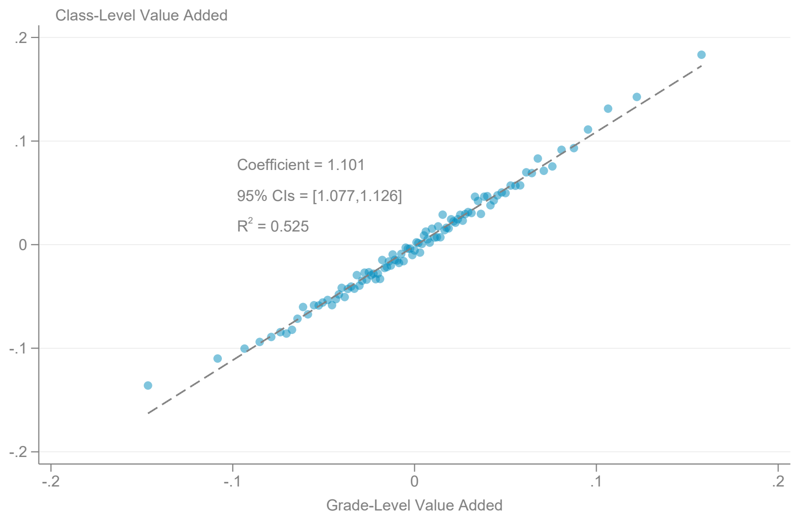
(f) + Election Vote Shares



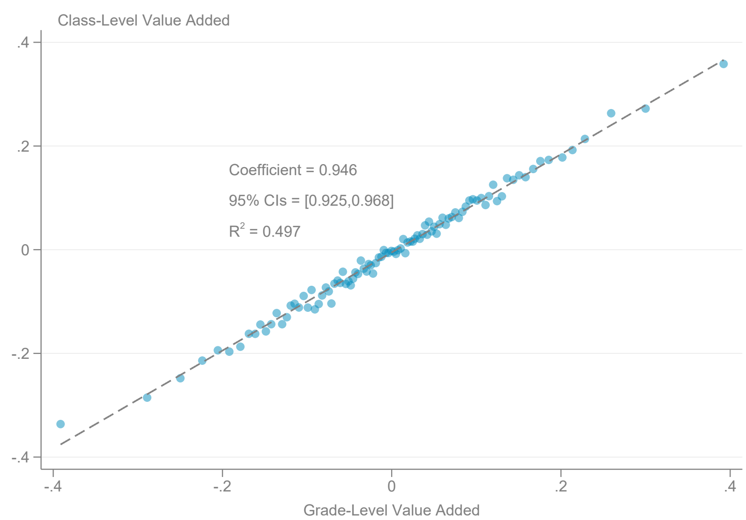
(g) Leave-One-Out Estimates

Note: The dependent variable in all panels is student attendance rates. Panel A uses the Callaway and Sant'Anna (2021) estimator with doubly-robust inverse propensity weights, adjusting for student demographic characteristics including race (Black, Hispanic, Asian, other), FRPL status, limited English status, and gender. Panel B adds the estimators from de Chaisemartin and d'Haultfoeuille (2020) and Sun and Abraham (2021) to the baseline model (in blue). Panel C adds in the missing districts, assigning the event year as 2012 (see Section VI). Panel D uses CBA-year fixed effects as categorized in Biasi and Sarsons (2022). Panel E adds flexible pay-year fixed effects as categorized in Biasi (2021). Panel F controls for a cubic polynomial in the recertification election vote share. The event study coefficients are normalized relative to the school year before the union decertification event. Event time is binned at six years before the event and five years after the event. Panel G plots 322 different regression estimates from equation 1, successively dropping each district at a time. All panels use data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A20: Forecast Bias between Class- and Grade-Level Value Added



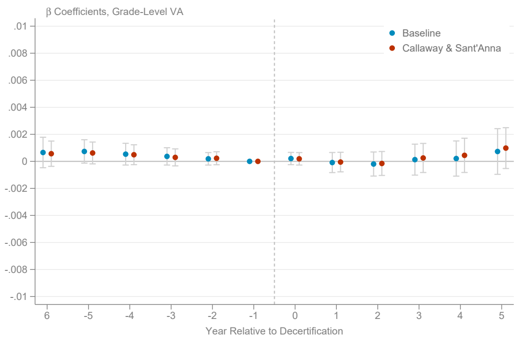
(a) Shrunken Estimates



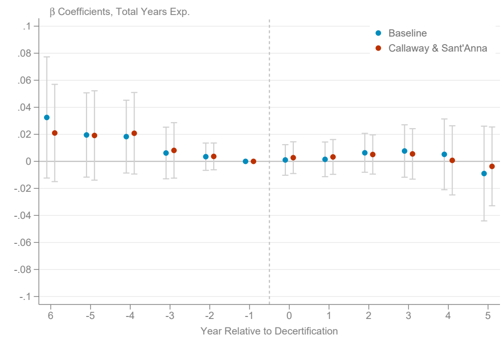
(b) Unshrunk Estimates

Note: Each figure shows a binscatter of class-level value added (y-axis) by grade-level value added (x-axis) over the 2017–19 time period. Panel A uses the shrunken estimates; panel B uses the unshrunken estimates. The dashed line plots the slope from the OLS regression of class-level value added on grade-level value added.

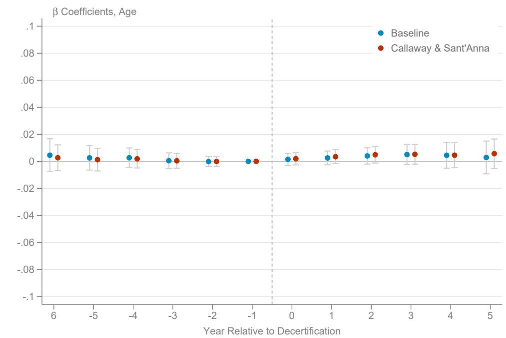
Figure A21: Teacher Compositional Changes



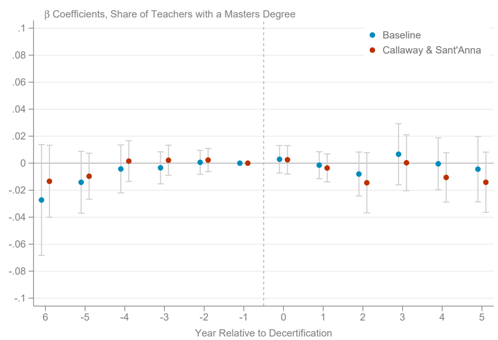
(a) Grade-Level VA



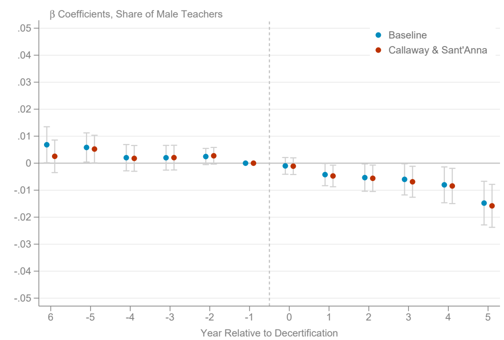
(b) Log Years of Experience



(c) Log Age



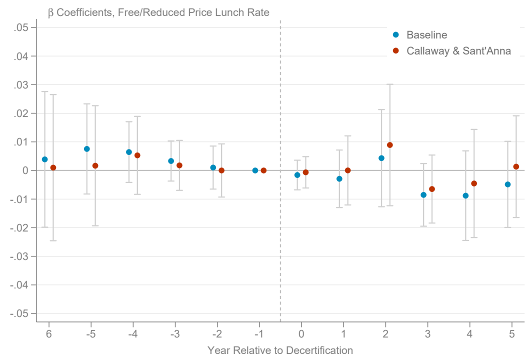
(d) Advanced Degree



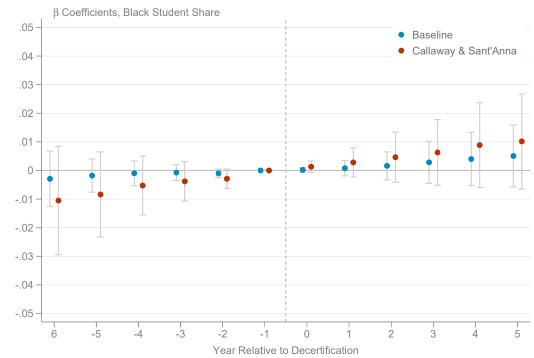
(e) Male

Note: The dependent variable in column 1 is the 2006–2011 grade-level value added score. The dependent variables in the remaining columns are years of experience on a logarithmic scale (column 2), age on a logarithmic scale (column 3), an indicator variable for having an advanced degree (column 4), and an indicator variable for being male (column 5). The coefficients are normalized relative to the school year before the union decertification event. The blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. Each panel uses data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

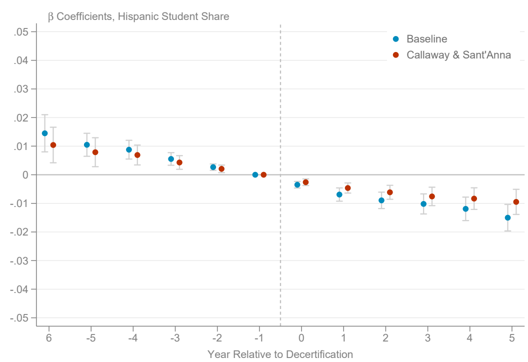
Figure A22: Student Compositional Changes



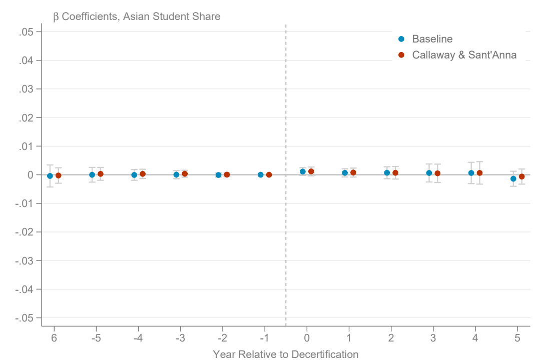
(a) Free/Reduced Price Lunch



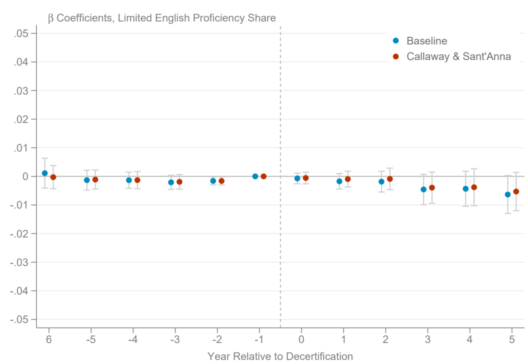
(b) Black



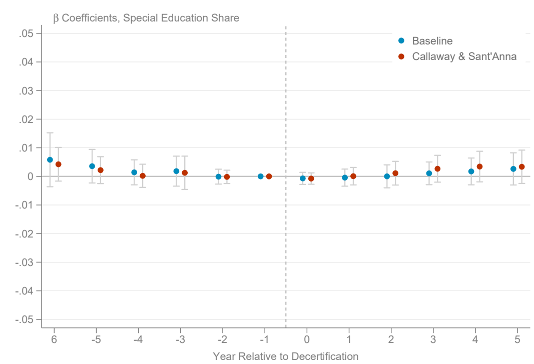
(c) Hispanic



(d) Asian



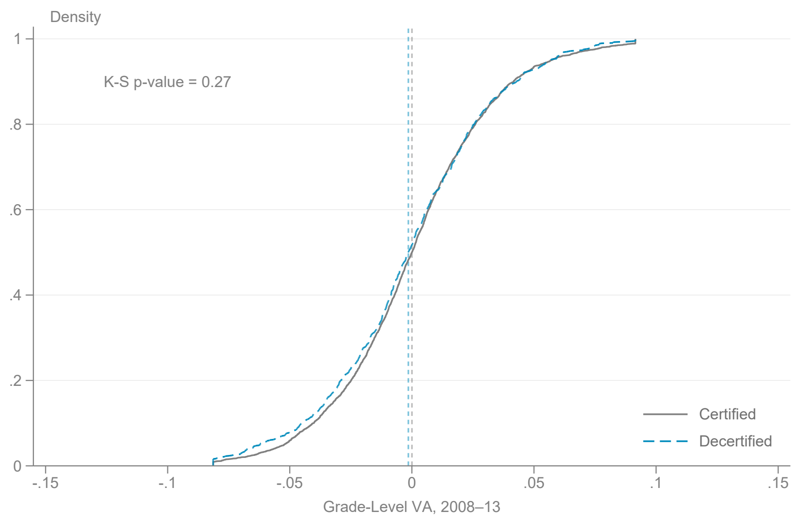
(e) Limited English



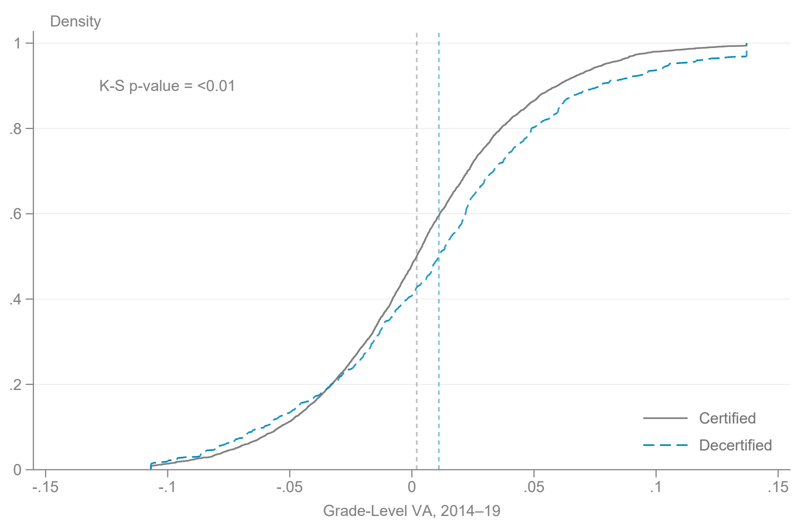
(f) Special Education

Notes: The dependent variables in panels A–F are indicator variables for FRPL, Black, Hispanic, Asian, limited English proficiency, and special education status, respectively. The coefficients are normalized relative to the school year before the union decertification event. The blue dots represent the baseline two-way fixed effects specification (binned at event time -6 and +5) and the red dots use the Callaway and Sant'Anna (2021) estimator. Each panel uses data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A23: Grade-Level VA Distributions, Before and After Decertification



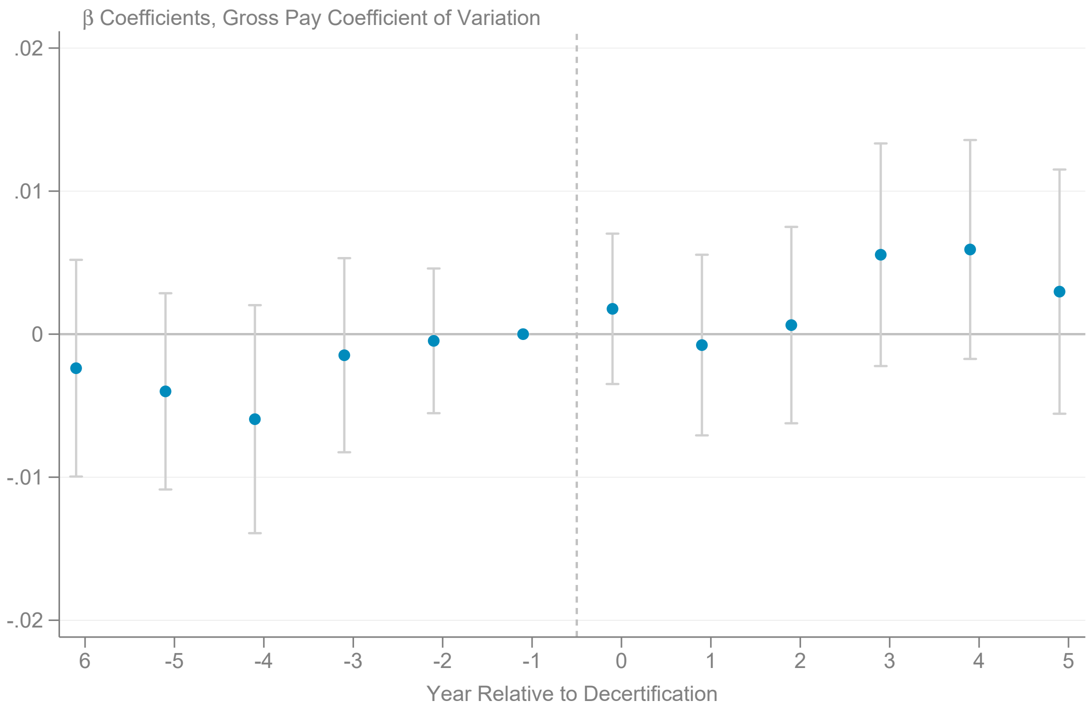
(a) Grade-Level VA, 2008–13



(b) Grade-Level VA, 2014–19

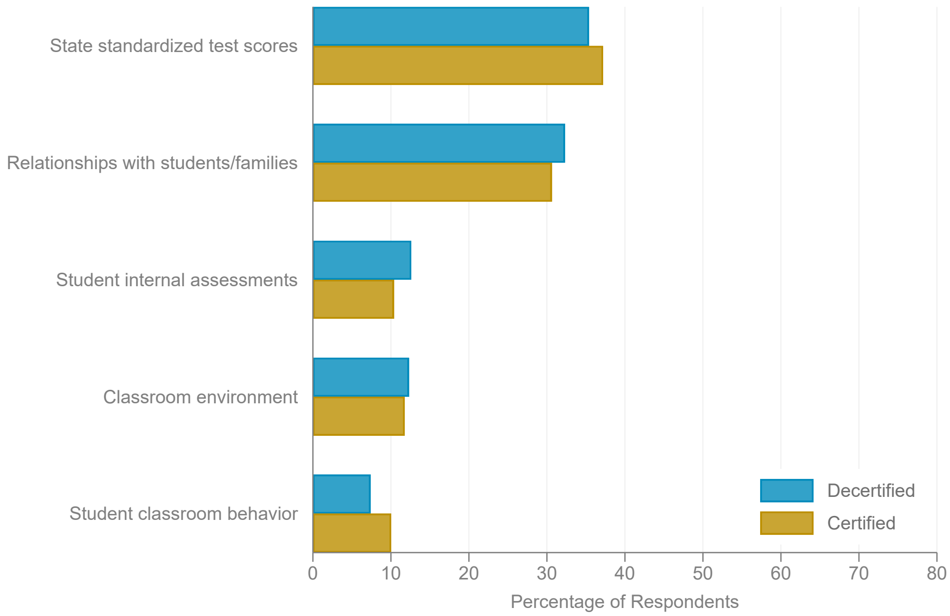
Note: Each figure plots the distribution of grade-level value-added scores by certified (gray) or decertified (blue) status. The sample includes districts that decertified in 2014 and districts that always remained certified. Only teachers who taught in both the pre- and post-2014 period are included. The grade-level measure in panel A is constructed using the 2008–13 school years (before any decertifications occurred). The measure in panel B uses all years between 2014–2019. In the top left corner, I report the p-value from a Kolmogorov–Smirnov test of equality between the two distributions.

Figure A24: Coefficient of Variation around Decertification

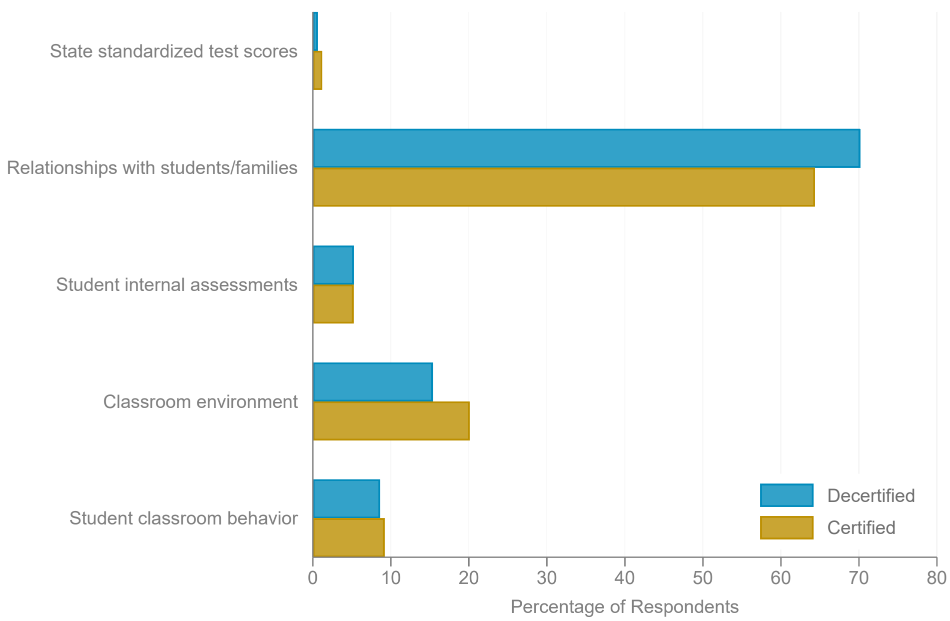


Note: This figure plots the coefficient of variation for log salaries conditional on fixed effects for experience. The coefficients are normalized relative to the school year before the union decertification event. The figure uses data from 2006–2019. The gray bars represent 95 percent confidence intervals based on standard errors that are clustered by school district.

Figure A25: Survey Evidence: What Teachers Feel is Most Important



(a) What Administrators Value



(b) What Teachers Value

Note: This figure uses survey data for teachers who reported teaching grades 4–8 and in one of the four core subjects. Respondents were asked to rank the five items from most important to least important. Both panels display the share of teachers who ranked the relevant item first. Panel A displays what teachers felt administrators valued the most; panel B displays what teachers themselves valued the most. See Appendix Section E for details.

Table A1: Correlates of Union Membership and Recertification Election Voting, Yes Votes > 99%

	(1) Member	(2) Voter
Years of Experience	0.013***	0.005***
	(0.001)	(0.0004)
Advanced Degree	0.042***	0.039***
	(0.012)	(0.008)
Value Added	-0.154	0.047
	(0.111)	(0.049)
Mean	0.55	0.77
Number of Teachers	7,552	7,552
Number of Districts	188	188
Observations	27,555	27,555

Note: This is a reproduction of Table 1 columns 5–6, but only using election years where greater than 99 percent of votes were in favor of recertification. One standard deviation of the distribution of value added scores is 0.074. The dependent variable in columns 1–2 are union membership and recertification election voting, respectively. All regressions include school and district-by-year fixed effects. Standard errors are two-way clustered by school district and individual. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A2: Correlates of Union Membership and Recertification Election Voting, Unshrunk Value Added

	(1) Member	(2) Member	(3) Member, Certified	(4) Member, Decertified	(5) Member	(6) Voter
Years of Experience	0.013***	0.013***	0.013***	0.012***	0.013***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0004)
Advanced Degree	0.027***	0.027***	0.029**	0.023	0.028**	0.038***
	(0.010)	(0.010)	(0.012)	(0.017)	(0.012)	(0.007)
Value Added		-0.067**	-0.088***	-0.015	-0.086***	0.013
		(0.027)	(0.031)	(0.041)	(0.031)	(0.017)
Mean	0.44	0.44	0.52	0.27	0.52	0.77
Number of Teachers	11,941	11,941	9,142	4,764	9,196	9,196
Number of Districts	385	385	217	234	219	219
Observations	70,229	70,229	48,578	21,651	49,479	49,479

Note: This is a reproduction of Table 1, but using the unshrunken value-added estimator. One standard deviation of the distribution of value added scores is 0.074. The dependent variable in columns 1–5 is union membership. Columns 1–2 use a sample of all teachers with non-missing years of experience, education, and value added. Columns 3–4 use the sample of districts that are certified or decertified, respectively. Columns 5–6 restrict the sample to district-years when a recertification election was held. The dependent variable in columns 5–6 are union membership and recertification election voting, respectively. All columns use data from 2016–2022. All regressions include school and district-by-year fixed effects. Standard errors are two-way clustered by school district and individual. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A3: Correlates of Union Membership and Recertification Election Voting, 1-Step Value Added

	(1) Member	(2) Member	(3) Member, Certified	(4) Member, Decertified	(5) Member	(6) Voter
Years of Experience	0.013***	0.013***	0.013***	0.012***	0.013***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0004)
Advanced Degree	0.027***	0.027***	0.029**	0.023	0.028**	0.038***
	(0.010)	(0.010)	(0.012)	(0.017)	(0.012)	(0.007)
Value Added		-0.132***	-0.147***	-0.099	-0.143***	-0.019
		(0.036)	(0.043)	(0.062)	(0.043)	(0.026)
Mean	0.44	0.44	0.52	0.27	0.52	0.77
Number of Teachers	11,941	11,941	9,142	4,764	9,196	9,196
Number of Districts	385	385	217	234	219	219
Observations	70,229	70,229	48,578	21,651	49,479	49,479

Note: This is a reproduction of Table 1, but using the one-step value-added estimator as described in Appendix Section C. One standard deviation of the distribution of one-step value added is 0.111. The dependent variable in columns 1–5 is union membership. Columns 1–2 use a sample of all teachers with non-missing years of experience, education, and value added. Columns 3–4 use the sample of districts that are certified and decertified, respectively. Columns 5–6 restrict the sample to district-years when a recertification election was held. The dependent variable in columns 5–6 are union membership and recertification election voting, respectively. All columns use data from 2016–2022. All regressions include school and district-by-year fixed effects. Standard errors are two-way clustered by school district and individual. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A4: Correlates of Union Membership and Recertification Election Voting, Errors-in-Variables Adjustment

	(1) Member	(2) Member, Certified	(3) Member, Decertified	(4) Member	(5) Voter
Years of Experience	0.013***	0.013***	0.012***	0.013***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.0005)
Advanced Degree	0.028***	0.030**	0.023	0.029**	0.037***
	(0.010)	(0.011)	(0.018)	(0.012)	(0.007)
Value Added	-0.338**	-0.466***	-0.045	-0.453***	0.076
	(0.140)	(0.170)	(0.205)	(0.177)	(0.099)
Mean	0.44	0.52	0.27	0.52	0.77
Number of Teachers	11,941	9,142	4,764	9,196	9,196
Number of Districts	385	217	234	219	219
Observations	70,229	48,578	21,651	49,479	49,479

Note: This is a reproduction of Table 1 columns 2–6, but using errors-in-variables regression to account for the measurement error of the value added regressor. One standard deviation of the distribution of value added scores is 0.074. The dependent variable in columns 1–4 is union membership. Column 1 use a sample of all teachers with non-missing years of experience, education, and value added. Columns 2–3 use the sample of districts that are certified and decertified, respectively. Columns 4–5 restrict the sample to district-years when a recertification election was held. The dependent variable in columns 4–5 are union membership and recertification election voting, respectively. All columns use data from 2016–2022. All regressions include school and district-by-year fixed effects. Standard errors are bootstrapped with 1,000 replications and two-way clustered by school district and individual. I use a reliability value of 0.44, which is determined according to the Kane and Staiger (2008) estimator (see Appendix Section C). The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A5: Additional Behavioral Outcomes

	(1) Any Suspension	(2) Any Expulsion	(3) HS Dropout
Post-Decertification	0.014	0.0001	0.001
	(0.012)	(0.0002)	(0.002)
Mean	0.05	0.001	0.03
Number of Districts	322	322	289
Observations	9,753,947	9,753,947	831,798

Note: The dependent variables in column 1–2 are any out-of-school suspensions or expulsions in a given year, respectively. The dependent variable in column 3 is whether a student dropped out of high school. Columns 1–2 use data from 2007–2019 as the suspension/expulsion data started to be tracked in 2007. Column 3 uses data from 2006–2019 for all students in grade 12. There are fewer districts in column 3 as some school districts do not have a high school. All specifications include district and year fixed effects. Standard errors are clustered by school district.

Table A6: Test Score Effects, Interaction with Various District Characteristics

	(1)	(2)	(3)	(4)
		Outcome = Student test scores		
X =	Mean Experience	Mean Turnover	Mean FRPL	Mean Enrollment
Post-Decertification	0.052***	0.052***	0.058***	0.059***
	(0.019)	(0.017)	(0.019)	(0.016)
Post-Decertification $\times$ $\mathbb{1}(X >$ Median)	-0.009	-0.007	-0.024	-0.013
	(0.021)	(0.023)	(0.021)	(0.019)
Number of Districts	322	322	322	322
Observations	5,004,965	5,004,965	5,004,965	5,004,965

Note: The dependent variable is average standardized test scores. In each specification, I add an interaction between the post-decertification coefficient and an indicator variable for whether the district-level average characteristic is above the median level of all districts. For instance, column 1 includes an interaction between the post-decertification indicator and an indicator for above-median district-level average teacher experience (in 2011). The interaction term in column 2 is the district-level average teacher exits in 2011. The interaction term in column 3 is the district-level average student FRPL rate in 2011. The interaction term in column 4 is the district-level average student enrollment in 2011. All specifications include district and year fixed effects. Standard errors are clustered by school district.

Table A7: Teacher Exit/Entry around Decertification

	(1) Exit District	(2) Exit District	(3) Enter District	(4) Enter District
Post-Decertification	-0.001	-0.006	0.004	0.002
	(0.003)	(0.005)	(0.003)	(0.004)
Post-Decertification × Above-Median VA		-0.002		0.001
		(0.005)		(0.004)
Above-Median VA		0.001		-0.005**
		(0.002)		(0.002)
Mean	0.10	0.08	0.10	0.05
Number of Districts	322	313	322	313
Observations	811,102	121,998	752,011	112,363

Note: This table summarizes how decertification affected measures of either exit from or entry to the school district. Columns 1 and 3 use the full sample of teachers; columns 2 and 4 use teachers with grade-level value added scores. The grade-level value added is measured pre-Act 10 (2006–11) as described in Appendix Section C. Above-median VA is an indicator variable for a teacher having a value-added score above the median for all teachers. The dependent variable in columns 1–2 is whether a teacher left the district in year t. The dependent variable in columns 3–4 is whether a teacher entered the district in year t. Columns 1–2 use data from 2006–2019. Columns 3–4 use data from 2007–2019 (as it is unclear if the person was new to the district/profession in 2006). There are fewer districts in columns 2 and 4 as some school districts do not have a high school. All specifications include district and year fixed effects. Standard errors are clustered by school district. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A8: Teacher Characteristics

	(1) Grade-Level VA	(2) Log Years Exp.	(3) Log Age	(4) Advanced Degree	(5) Male
Panel A: All Teachers					
Post-Decertification	-0.00003	-0.020	0.001	0.011	-0.010***
	(0.001)	(0.021)	(0.005)	(0.014)	(0.003)
Mean	0.00	2.37	3.73	0.53	0.24
Number of Districts	313	322	322	322	322
Observations	121,972	809,675	809,675	809,675	809,675
Panel B: Exiters					
Post-Decertification	-0.001	-0.018	0.0002	0.015	-0.010
	(0.002)	(0.027)	(0.007)	(0.012)	(0.007)
Mean	-0.00	2.17	3.72	0.47	0.23
Number of Districts	309	322	322	322	322
Observations	9,180	82,677	82,677	82,677	82,677
Panel C: Entrants					
Post-Decertification	0.003	0.034	0.010	-0.003	-0.009
	(0.002)	(0.036)	(0.007)	(0.011)	(0.007)
Mean	0.00	1.12	3.49	0.29	0.21
Number of Districts	287	322	322	322	322
Observations	5,340	76,572	76,572	76,572	76,572

Note: This table summarizes whether characteristics of the teachers changed post-decertification for all teaching staff (panel A), just those who left the district (panel B), and those who entered the district (panel C). The dependent variables are the 2006–2011 grade-level value added score (column 1), years of experience on a logarithmic scale (column 2), age on a logarithmic scale (column 3), an indicator variable for having an advanced degree (column 4), and an indicator variable for being male (column 5). There are fewer districts in Column 1 because value added cannot be estimated for the nine districts that are comprised of a single high school. All specifications include district and year fixed effects. Standard errors are clustered by school district. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A9: Student Characteristics

	(1) FRPL	(2) Black	(3) Hispanic	(4) Asian	(5) Limited English	(6) Special Ed.
Post-Decertification	-0.007	0.004	-0.017***	0.001	-0.002	-0.002
	(0.009)	(0.005)	(0.002)	(0.002)	(0.003)	(0.003)
Mean	0.38	0.11	0.10	0.04	0.06	0.14
Number of Districts	322	322	322	322	322	322
Observations	10,529,416	10,529,416	10,529,416	10,529,416	10,529,416	10,529,416

Note: This table regresses various student characteristics on the post-decertification indicator. The dependent variables in columns 1–6 are indicator variables for FRPL, Black, Hispanic, Asian, limited English proficiency, and special education status, respectively. All specifications include district and year fixed effects. Standard errors are clustered by school district. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A10: Changes to Teacher Compensation and District Spending

	(1) Log Salaries	(2) Log Per-Pupil Spending
Post-Decertification	0.006	0.002
	(0.007)	(0.008)
Observations	810,079	4,498
Number of Districts	322	322

Note: This table summarizes how decertification affected teacher compensation and district spending. The dependent variable in column 1 is teacher salaries divided by a worker's full-time employment level (where full-time = 100). The dependent variable in column 2 is per-pupil spending at the district-level. All outcomes are represented on a log scale. All specifications include district and year fixed effects and use data from 2006–2019. Standard errors are clustered by school district. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A11: Changes in Teacher Salaries by Value Added and Experience

	(1)	(2)
	Outcome = Log Salaries	
Post-Decertification	0.004	0.045***
	(0.006)	(0.011)
Post-Decertification × Value Added	-0.016	
	(0.032)	
Post-Decertification × Experience		-0.003***
		(0.0005)
Value Added	0.045**	0.046**
	(0.022)	(0.019)
Experience	0.022***	0.022***
	(0.0003)	(0.0003)
Number of Districts	313	313
Observations	95,933	95,933

Note: This table indicates whether teachers were paid more following decertification by either value added (column 1) or experience (column 2). The dependent variable is log teacher salary. The sample uses teachers for whom value added can be estimated. Column 1 interacts the post-decertification term with a teacher's value added score (estimated over years 2017–19). Column 2 interacts the post-decertification term with a teacher's years of total experience. All specifications use data from 2006–2019. Standard errors are clustered by school district. There are fewer school districts than in the other main tables because value added cannot be estimated for the nine districts that are comprised of a single high school. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A12: Selection into Taking the Survey

	(1) Experience	(2) ELA	(3) Math	(4) Male	(5) Non-White	(6) Advanced Degree
Took Survey	0.642***	-0.018*	-0.038***	0.039***	0.0002	0.016*
	(0.197)	(0.011)	(0.012)	(0.009)	(0.006)	(0.010)
Mean	14.60	0.34	0.33	0.23	0.07	0.50
Observations	61,573	61,573	61,573	61,573	61,573	61,573

Note: The dependent variables are total experience (column 1), an indicator variable for whether the individual teaches ELA or math (columns 2–3), and an indicator variable for whether the teacher is male, non-white, or has an advanced degree (columns 4–6). These characteristics are regressed on "Took Survey" an indicator variable for whether the teacher filled out at least part of the online survey. Each observation represents one teacher. See Appendix Section E for survey details. Standard errors are clustered by school district. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels, respectively.

Table A13: Survey Evidence: Teacher Perceptions on Job Attributes

	(1)	(2)	(3)	(4)	(5)	(6)
		Satisfied with		Secure from	Discretion over	
	Hours Worked	Work Env.	Salary	Lay Off	Curriculum	Work Duties
Decertified	0.922**	0.042	-0.015	0.016	0.084***	-0.013
	(0.477)	(0.028)	(0.028)	(0.025)	(0.025)	(0.024)
Certified mean	47.59	0.55	0.29	0.79	0.57	0.37
FDR q-values	0.048	0.100	0.222	0.222	0.002	0.222
Number of Districts	265	265	265	265	265	265
Observations	2,283	2,283	2,283	2,283	2,283	2,283

Note: The column 1 dependent variable is the number of reported hours worked per week. The dependent variables in columns 2–6 are whether the respondent selected "strongly agree" or "agree" to each of the survey questions. The specific outcomes are satisfaction with the work environment (col. 2), satisfaction with one's salary (col. 3), feeling secure from being laid off (col. 4), having discretion over one's curriculum (col. 5), and having agency over one's work duties (col. 6). The means are weighted to match the distribution of covariates observed in the full teacher sample. See Appendix Section E for survey details. Standard errors are clustered by school district. False-discovery rate q-values (Anderson 2008) are reported. The symbols *, **, *** denote statistical significance at the 10, 5, and 1 percent levels according to the adjusted q-values, respectively.

B. Data Appendix

A. Data Sources

Wisconsin Campaign Contributions: The Wisconsin Campaign Finance Information System provides publicly available data on all campaign contributions/spending.[38] I download all data where the receiving registrant was the WEAC PAC or one of the 13 associated regional chapters. This data includes the amount of money contributed for each reporting period (often at a quarterly level), the contributor's name, and the contributor's address. Since 2016, the WEAC PAC and the PACs of its 13 regional chapters have provided quarterly reports listing each contributor. To merge these listings with the DPI administrative data, I sum each person's contributions to an annual level.

DPI Staff Data: The DPI website has a public administrative dataset of all public school workers.[39] This dataset contains information on demographic characteristics (gender, race, birth year), salary and benefits information, occupational information, and data on one's place of work. I append the 2006 to 2022 data together into a single dataset. Starting in 2015, the data have individual identifiers. I back-fill the pre-2015 identifiers by someone's name and year of birth. Likewise, I create new identifiers for people who do not appear in the post-2015 data using their name and date of birth. There are some rare cases where someone has the same name and year of birth and are not identified by the post-2015 identifier codes. I therefore additionally use the geographic work information to separately create identifiers for these people. Individuals can appear in the dataset more than once in a year if they worked more than one job assignment. In cases where this job assignment spans multiple schools, I assign the person to the school where they work for the majority of their time (each position has an "FTE" indicating how much of their time was spent in each position). I also drop workers who appear in the data but who have no salary information, have a salary of $0, or are missing full-time employment information.[40]

DPI Student Data: The DPI collects student-level data for all public-school students across the state. I use individual-level data under a data use agreement with the DPI. From 2006–2022, I have data on student demographics (school attended, grade, race, gender, limited English proficiency status, grade repetition status, special education status, and free or reduced price lunch status) and standardized test score information. I also have data on student attendance (number of days attended and scheduled to attend) and whether the student dropped out of the education system. From 2007–2022, I have disciplinary information including suspensions and expulsions. In 2017, the state started collecting student-staff links. This dataset includes each course that a student took, the staff member(s) who taught the particular course, and the staff member's role in a particular course.

From 2006 to 2014, Wisconsin's universal standardized test was called the Wisconsin Knowledge and Concepts Examination (WKCE). Students in grades 3–8 and 10th grade took exams in reading and math. Students in grades 4, 8, and 10 were additionally tested in science, social studies, and English language art/writing. In 2015, Wisconsin replaced the WKCE test with the

[38] See https://cfis.wi.gov/Public/Registration.aspx?page=ReceiptList.

[39] See https://publicstaffreports.dpi.wi.gov/PubStaffReport/Public/PublicReport/AllStaffReport.

[40] These observations are typically support staff and substitutes. The DPI does not require that salary information be reported for non-professional staff.

Badger exam for students in grades 3–8. One minor change was the relabeling of the reading exam as English and language arts (ELA); I treat these as the same across years. Additionally, for 10th graders, the WKCE exam was replaced with the ASPIRE exam, a pre-ACT assessment. Similar to the ACT, the ASPIRE exam tested students on ELA, science, and math.[41] From 2016–2022, Wisconsin changed exams again to the Forward exam.

WERC Data: WERC publicly posts the results of all union recertification elections online.[42] I download all election dates that involve teachers' unions (typically these take place in the fall). This information includes the eligible unit population (i.e., the number of teachers in a school district), the total number of votes cast, the number of votes in favor of representation, and the number of votes against representation. Additionally, I received data via FOIA request which lists the names of all voters for each district-year recertification election.

B. Merging Administrative Datasets

DPI to WCFIS: First, I merge the administrative salary data from DPI to the WCFIS campaign contributions data. The unique information common to both datasets is a person's full name and the regional WEAC chapter they work for in a given year. The regional chapter in the WCFIS data comes from the fact that each regional chapter has a separate WEAC PAC. For the DPI data, this information is linked via district identifiers from here: https://weac.org/region-finder-test/.

In the DPI data, 99 percent of people are uniquely identified by name, year, and WEAC region. Of the 1 percent who are not uniquely identified, 90 percent cannot be identified from one other person, while 10 percent cannot be identified from two or more people. For these duplicates, there are three possible scenarios: (1) The number of duplicates matches the number of unique people in the WCFIS data. (2) No one appears in the WCFIS data. (3) There is one or more people in the WCFIS data, but it is not immediately clear which person matches which teacher. In the first case, all of the duplicates are considered members, while in the second, no one is a member. In the third case, I denote membership status as missing.

I first match people exactly by name (and regional chapter), but supplement this using the Stata `reclink` package. The `reclink` algorithm applies a score between 0 and 1 denoting the likelihood of a match. By hand, I then go through all matches above 0.75 to check which matches are correct.[43] Common reasons for why the fuzzy match catches true matches are: first name nicknames (e.g., Andy for Andrew), hyphenated last names that are missing one of the parts of a name, small misspellings in long or unique names, or cases where a person moved districts but the WCFIS information still has them listed at their previous regional chapter.

Appendix Figure A5 shows that the merged dataset appears to track both the counts and shares of union members from surveys and administrative statistics. One concern, however, is that panel A suggests that the 2016 merge undercounts the number of union members since the itemization requirement started mid-school year (in January 2016). For most people, this is not a problem because they appear in the spring 2016 data (as the PAC has to make quarterly campaign finance

[41] Unlike the actual ACT exam, the ASPIRE exam was on a three-digit scale. Students received a composite ELA score, but also separate scores for reading, writing, and English. 10th-graders also continued to take social studies separately under the Badger/Forward exam.

[42] See http://werc.wi.gov/representation-election-updates/.

[43] The matches below 0.75 are very poor quality.

reports). For other people though, they are only listed in the fall of each year giving the annual \$19.99. To alleviate this issue, I denote the membership variable as missing in 2016 for anyone who appears in the 2017 data as giving \$19.99 in the fall, but is not observed as being a member in 2016. Alternatively, all results are robust to just dropping the 2016 data.

DPI to WERC: Next, I merge the DPI administrative data to the list of WERC recertification election voters. This is relatively easier to assign as both datasets include the school district name. In the DPI data, 99.9 percent of teachers are uniquely identified by their name and school district, while 0.1 percent cannot be identified from one or more people. I follow the same procedure as above, assigning voter status to all teachers if all of the duplicates appear on the voter list or to none of them if none appear. In the remaining few cases where there is a mismatch, I denote the voter information as "missing" because there is no way to say which teacher voted. Finally, I once again use the fuzzy match as described above to catch other true matches.

C. Data Variables

Any expulsion: An indicator variable equal to 1 if a student was expelled in a given year.

Any suspension: An indicator variable equal to 1 if a student received an out-of-school suspension in a given year.

Attendance rate: The number of days a student attended divided by the number of days a student was scheduled to be in attendance in a school year. If the student had a half day absence, the value equals 0.5 and if a full day absence the value is 0. If the absence reason is authorized the day is not counted as absent.

Average standardized test score: The average state test score across all subjects. For grades 3, 5, 6, and 7 this is the average across math and ELA. For grades 4, 8, and 10 this is the average across math, ELA, science, and social studies.

FRPL status: An indicator variable equal to 1 if a student qualified for free- or reduced-price lunch in a given year.

High school dropout status: An indicator variable equal to 1 a student discontinued school enrollment without obtaining a high school completion credential. I limit the analysis of this variable to current 12th graders.

Limited English proficiency status: An indicator variable equal to 1 if a student was considered limited English proficient in a given year.

Special education status: An indicator variable equal to 1 if a student was in special education in a given year.

Student race indicators: Indicator variables for each race category that appears in the data: Alaskan Native, Asian, Black - African American, Native Hawaiian - Pacific Islander, White, Hispanic, Two or More Races. In 2011, the DPI added "Two or More Races" as a category. To lessen the impact of this change, I recode student race as the modal choice within students.

C. Value Added Measures

A. Kane and Staiger (2008) Two-Step Approach

In the main analysis, I follow the Kane and Staiger (2008) two-step procedure using Bayes shrinkage (hereafter, I refer to this as "KS"). In what follows, I detail the procedure and also discuss alternative estimators.

Recall that value added estimation is interested in predicting a teacher-specific effect on student test scores. To start, consider the following model of student test scores:

$$A_{ijt} = \mathbf{X_{ijt}}\beta + v_{ijt},$$

where A_{ijt} is a student's test score in a given year, $\mathbf{X_{ijt}}$ is a vector of observable covariates, and v_{ijt} is the unobserved residual. The residual can be expanded as

$$A_{ijt} = \mathbf{X_{ijt}}\beta + \mu_j + \theta_{jt} + \epsilon_{ijt},$$

where μ_j is the teacher value-added effect, θ_{jt} is an unobserved classroom effect unrelated to teacher j, and ϵ_{ijt} is an unobserved student-level effect.

To estimate the μ_j parameters, the two-step estimator regresses student-level residuals on teacher-by-year fixed effects. Specifically in the first step, I regress

$$A_{ijt} = \alpha + \mathbf{X_{ijt}}\beta + v_{ijt}, \tag{5}$$

where A_{ijt} is a student's math or ELA test score in a given year and X_{ijt} is a vector of controls. The controls include a cubic polynomial of lagged test scores in both ELA and math interacted with a student's grade as well as classroom-level averages of these lagged test score polynomials. I also control for lagged disciplinary measures (grade repetition, log days disciplined, and log days absent) in the same fashion at the student and class-level, as well as grade-by-year fixed effects, classroom size, and student and classroom average demographics (special education, limited English proficiency, and FRPL status; gender and race indicators). Observations are dropped when missing either the test score of interest or the lagged score. If a different variable is missing, the observation is included with an indicator variable denoting missing status.

In the second step, the residuals from the above equation are regressed on teacher-by-year indicators:

$$v_{ijt} = \mathbf{T_{jt}}\delta + \epsilon_{ijt}. \tag{6}$$

To come up with a time invariant measure of value-added, the KS estimator weights the annual average residuals by their reliability, where school years with greater precision are given more weight. Specifically, for each teacher, a time-invariant measure v_j is constructed as follows:

$$v_j = \sum_t w_{jt}\bar{v}_{jt},$$

where the yearly weights w_{jt} are

$$w_{jt} = \frac{h_{jt}}{\sum_t h_{jt}}$$

and

$$h_{jt} = \frac{1}{Var(\bar{v}_{jt}|\mu_j)} = \frac{1}{\hat{\sigma}_\theta^2 + (\frac{\hat{\sigma}_\epsilon^2}{n_{jt}})} = \frac{n_{jt}}{n_{jt}\hat{\sigma}_\theta^2 + \hat{\sigma}_\epsilon}.$$

In the above equation n_{jt} is the number of students the teacher j taught in year t, $\hat{\sigma}_\theta^2$ is the variance in the unobserved classroom component, and $\hat{\sigma}_\epsilon^2$ is the variance in the student component. To estimate the variance in the student component, KS use the within classroom variance in v_{ijt} defined as $\hat{\sigma}_\epsilon^2 = Var(v_{ijt} - \bar{v}_{jt})$. While unobserved, the variance of the classroom is estimated as the within-classroom variance not explained by the student or teacher components: $\hat{\sigma}_\theta^2 = Var(v_{ijt}) - \hat{\sigma}_\mu^2 - \hat{\sigma}_\epsilon^2$. To estimate the teacher component, KS use the covariance between a teacher's average residuals in two consecutive years weighted by the number of students: $\hat{\sigma}_\mu^2 = Cov(\bar{v}_{jt}, \bar{v}_{jt-1})$.

In practice, the above weights put more emphasis on years where the teacher taught more students.

The $\bar{v}_j$ term above is the unshrunken value-added estimate. To apply the shrinkage term, consider the following general empirical Bayes set-up:

$$VA_j^S = w_j\bar{v}_j + (1 - w_j)\bar{v},$$

where $\bar{v}_j$ is the unshrunken estimate and $\bar{v}$ is the average value added across teachers, which is 0 by construction. The weights w_j are typically the signal variance over the true variance. KS estimate the shrunken value added with the following weighting scheme:

$$VA_j^S = \left(\frac{\hat{\sigma}_\mu^2}{var(\bar{v}_j)}\right)\bar{v}_j,$$

where

$$var(\bar{v}_j) = \hat{\sigma}_\mu^2 + \left(\sum_t h_{jt}\right)^{-1}.$$

In practice, teachers with a high $\sum_t h_{jt}^{-1}$ (lower estimation error) have shrunken value-added estimates that are similar to the unshrunk version. This occurs when they have many students over the years of estimation.

I follow KS and Chetty, Friedman, and Rockoff (2014) and drop classrooms with fewer than five students or greater than 50 as these may be mismeasured. Additionally, I drop classes where more than 25 percent of students are in special education.

B. One-Step Approach

In the one-step approach, teacher fixed effects are included directly in the step 1 regression:

$$A_{ijt} = \alpha + \mathbf{X_{ijt}}\beta + \mathbf{T_{jt}}\delta + v_{ijt}.$$

I then follow the same Bayes shrinkage procedure as described in the two-step process. Bacher-Hicks and Koedel (2023) describe the tradeoffs between the one- and two-step procedures. In general, the estimators give similar results, but are slightly different because the control variables are identified by either within teacher variation (one-step model) or within- and across-teacher variation (two-step model). Therefore, the downside of the two-step procedure is that any correlations between teacher quality and the control variables will be attributed to the control variables. Conversely, the one-step model may increase attenuation bias (Parsons, Koedel, and Tan 2019).

C. Grade-Level Value Added

In the section on mechanisms, I use a value-added measure at the grade level to circumvent the problem that classroom level links are only available from 2016–17 to 2017–19. This allows for the construction of a value-added metric over any time period in the sample; the downside is that there is estimation error given that I cannot observe exactly which teachers taught which students. In what follows, I describe the procedure that essentially assigns grade-level average residuals to each teacher in the grade. For example, if there are three 5th-grade teachers, then each teacher is assigned the grade-level residual for that year. This is very similar to Biasi (2021), except that I do not use team fixed effects.

The grade-level measure follows KS as closely as possible but with the following changes. First, the control variables in the step 1 regression use grade-level averages rather than class-level averages and grade size rather than class size. In the second-step, instead of regressing the step-one residuals on teacher-year indicators, I create average school-by-grade-by-year residuals. The time-invariant unshrunken VA is then

$$v_j = \sum_t w_{jgt} \bar{v}_{jgt},$$

where $\bar{v}_{jgt}$ is the average residuals in grade g and year t that teacher j taught in. In other words, if the grade-level math team has three teachers, then they each get the same $\bar{v}_{jgt}$ for that year. However, because teachers move across grades and schools, they will likely have different v_j terms in the aggregate. The weights w_{jgt} are as before except for the following changes:

First, the variance of the student component $\hat{\sigma}^2_\epsilon$ is over the grade-level rather than the teacher-level:

$$\hat{\sigma}^2_\epsilon = Var(v_{igt} - \bar{v}_{gt}).$$

Second, in place of the teacher component $\hat{\sigma}^2_\mu$, I use the covariance between a teacher's grade-level residuals over time:

$$\hat{\sigma}^2_\mu = Cov(\bar{v}_{jgt}, \bar{v}_{jgt-1}).$$

The covariance is weighted by the number of students in the grade divided by the number of teachers in the grade.[44] Finally, the classroom component $\hat{\sigma}^2_\theta$ is replaced with the grade-level equivalent:

$$\hat{\sigma}^2_\theta = Var(v_{igt}) - \hat{\sigma}^2_\mu - \hat{\sigma}^2_\epsilon.$$

The only other difference between the KS classroom value-added and the grade-level measure

[44] In contrast, weighting by the total number of students in a grade results in putting the most weight on grades where each individual teacher is the least likely to see any one student.

is that the weights use the number of students in the grade divided by the number of teachers in the grade. In practice, the grade-level measure is likely to mirror a teacher's "true" effect when more years are included and when a teacher moves grades and/or schools more frequently. However, if two teachers always work the same assignment together, then they will be given the average of their two true effects. Figure A20 examines how well the grade-level measure predicts the classroom-level measure. Using the shrunken estimates, there is forecast bias of roughly 0.1 (panel A). For the unshrunken estimates, there is forecast bias of about 0.05 (panel B).

D. Conceptual Framework

I propose a very simple framework to understand why an individual would choose to be a dues-paying union member in the presence of a clear free-rider incentive. As predicted by Olson (1965), membership will decline to zero without compulsory mechanisms because a worker can receive all of the collective benefits without bearing any of the individual costs. However, other research suggests that there are in fact individual excludable benefits to membership such as social reasons (Booth 1985, Naylor and Cripps 1993) and job protection or representation rights (Blanchflower et al. 1990, Murphy 2020).

Following these prior studies, I model the decision to join the union as a trade-off between the private benefits and the private costs, as any collective benefits (e.g., wage bargaining) can be attained by free-riding. What makes this framework unique is the consideration of whether there is positive or negative selection into union membership based on a teacher's underlying performance. The implications will depend on whether workers join for social reasons and/or representation benefits and whether either of these forms of benefits interact with a worker's type.[45] This simplified framework shares similarities with the model in Dal Bó, Finan, and Rossi (2013), which examines selection into the public-sector workforce.

Consider a teacher of type θ who earns wages w and has the choice of whether to join the union or not. Initially, I assume that w does not depend on type, but rather is solely a function of experience x and education e as in a "step-and-lane" compensation system. A teacher pays a cost c of being in the union. This cost can be thought of as the annual membership dues, which are typically around \$1,000 per year regardless of income. The benefits b of being a union member can be broken up into two parts: an individual benefit b_i and a collective benefit b_C. Similar to Booth (1985), the collective benefit simply depends on the share of local workers M who are members of the union $b_C(M)$, where $0 \leq M \leq 1$ and $\frac{\partial b_C}{\partial M} > 0$. This captures the idea that the collective benefits strengthen as more people are union members. The individual benefits can be further decomposed as follows:

$$b_i(\theta) = r(\theta) + a(\theta) + \epsilon_i,$$

where r is the representation benefit as in Murphy (2020) and a is the social benefit as in Booth (1985). I assume that $r > 0$ and $a > 0$, indicating that all workers derive some positive value to each of these components. Both terms are indexed by θ, suggesting that each of these components may be more valuable for workers of a certain type. However, I remain agnostic about the sign of this relationship. Finally, ϵ represents an idiosyncratic taste for unions unrelated to θ, which can be positive or negative.

[45] Unions may also offer other excludable benefits, such as access to professional development programs or discounts on goods and services. I abstract away from these other individual benefits for simplicity.

When not joining the union, each teacher receives wages and the collective benefit, which they cannot be excluded from. Each teacher's utility when not joining the union is simply

$$U^n = w(x, e) + b_C(M).$$

In comparison, a union member's utility is

$$\begin{aligned} U^u &= w(x, e) + b_C(M') + b_i(\theta) - c \\ &= w(x, e) + b_C(M') + r(\theta) + a(\theta) + \epsilon_i - c, \end{aligned}$$

where M' represents the fact that the numerator of M increases by one. This implies that a teacher joins the union if

$$\begin{aligned} r(\theta) + a(\theta) + \epsilon_i + \underbrace{\big(b_C(M') - b_C(M)\big)}_{\Delta b_C \approx 0} \geq c \implies \\ r(\theta) + a(\theta) + \epsilon_i \geq c. \end{aligned}$$

In line with the free-rider incentive, $b_C(M)$ falls out of the expression because an individual receives the collective benefits regardless of whether they are a member. Therefore, a teacher joins the union if their value of representation and social benefits exceeds the cost of union dues (and net of any personal taste for unions ϵ).

What does this imply for whether there exists positive or negative selection into membership? The prediction is theoretically ambiguous because it depends on the direction of both $\frac{\partial r}{\partial \theta}$ and $\frac{\partial a}{\partial \theta}$. However, if we assume that $\frac{\partial r}{\partial \theta} < 0$, meaning higher performers value representation less than lower performers, then the selection can be summarized by the direction of $\frac{\partial a}{\partial \theta}$:

(1) $\frac{\partial a}{\partial \theta} > 0$: If social benefits a and worker quality θ are positively related (e.g, as in Dal Bó, Finan, and Rossi 2013), then the prediction is ambiguous. The terms could cancel each other out, resulting in no selection, or the sign will depend on which of the two effects dominates. For example, if $\frac{\partial a}{\partial \theta} > \frac{\partial r}{\partial \theta}$ in absolute value, there exists positive selection into membership.

(2) $\frac{\partial a}{\partial \theta} <= 0$: If there is no relationship or a negative interaction between a and θ, then there will exist negative selection into membership.

While I am not able to measure r or a directly, the framework provides general insights regarding the sign of selection into membership to consider when examining the descriptive evidence. For instance, if there exists positive selection in the data, then there likely is a correlation between social factors and performance that exceeds that of the representation benefits. However, if selection is negative, then the framework suggests that representation benefits are relatively more important.

Next, I consider two extensions: (1) applying the framework to the question of why people vote in the recertification elections and (2) considering whether workers can be paid by θ.

Support For Union Status: The model can also be applied to the case of who supports the union existing as a legal entity in the recertification elections. In contrast to membership, it costs an individual nothing but the time it takes to submit a ballot online. Therefore, the model predicts that all people should vote in favor of union status, except those who have extreme distaste for unions above their representation and social benefits:

$$r'(\theta) + a'(\theta) + \epsilon_i \geq 0.$$

Here, I refer to the representation and social benefits as r' and a' to distinguish from the membership case because the benefits are likely different when being a member. That is, one only gets the representation benefits by paying dues, but may still vote in favor of recertification for that reason (i.e., $r \geq r' > 0$). Likewise, social reasons may be different between membership and casting a vote. For example, a teacher may feel more peer pressure to pay dues relative to casting a vote in the recertification elections. Nevertheless, people will still vote in favor of recertification for these reasons, consistent with the idea of expressive voting (e.g., see Hamlin and Jennings 2019). Importantly, it need not be the case that the sign of selection is the same between membership and recertification election voting. For example, suppose case (1) in the main text holds where $\frac{\partial r}{\partial \theta} < 0$ and $\frac{\partial a}{\partial \theta} > 0$, but where the r effect dominates. Then, the sign of the election terms can be similar $\frac{\partial r'}{\partial \theta} < 0$ and $\frac{\partial a'}{\partial \theta} > 0$, but now where $\frac{\partial a'}{\partial \theta} > \frac{\partial r'}{\partial \theta}$ in absolute terms. In other words, people support recertification for both reasons but the social benefit is more important than the representation benefit. This could be possible because an individual does not actually receive the representation benefit without paying dues.

Pay for Performance: What if teachers can be paid by θ? If teachers believe that the union upholds a seniority-driven pay schedule, then higher value-added teachers should be less likely to view the union favorably, all else equal. This should sharpen the prediction of negative association between membership and value added following a simple Roy (1951) model of selection.

E. Survey Details

In April 2024, I sent an online Qualtrics survey to the majority of the state teacher workforce. I created a mailing list of teachers' emails using public school districts' online staff directories. In total, I sent the email to about 49,000 teachers from 85 percent of school districts in the state. The missing districts were those who did not post their staff emails online.

Teachers were invited to participate in an online research study regarding their jobs and teachers' unions. For participating, they were automatically entered into a raffle drawing where 60 winners received a $50 Amazon gift card. Email recipients were told that the survey would take about 5–10 minutes to complete; in practice, the median completion time was about 5.7 minutes.

I received full or partial responses from 2,942 individuals, a response rate of about 6 percent. Roughly 93 percent of respondents completed the entire survey. Survey respondents were slightly selected on demographic characteristics relative to the full teacher sample (Appendix Table A12). To account for this, I use entropy balancing weights using the Stata `ebalance` package (Hainmueller and Xu 2013). This reweights the survey sample to match the observable characteristics from the target population, though results are very similar when unweighted as well.

Below, I display screenshots of the relevant survey portions that are used in the paper. First, respondents answered questions regarding their work hours and what grades/subjects they taught:

Berkeley
UNIVERSITY OF CALIFORNIA

First, I would like to ask about your job.

About how many hours per week do you work in your primary job?

(Please include hours both in and out of school: e.g., class prep, grading papers, etc.)

[ ] **hours**

Do you teach students in grades 3–8?

○ Yes

○ No

Do you teach math, reading, science, or social studies?

○ Yes

○ No

The first question regarding hours worked appears in Table A13 column 1. If a respondent answered "Yes" to the next two questions, they were then asked questions about what they value as a teacher and what they feel administrators value (next page). This was only shown to 4–8 grade teachers in the four core subjects because these are the grades where state standardized testing may be a relevant component in their evaluations.

When considering your success as a teacher, which factors do you think **school administrators** feel are most important?

Please rank from most important (#1) to least important (#5).

- Student internal assessments (for example: unit tests, classroom assessments)
- Classroom environment
- Your relationships with students/families
- Student test scores on state Forward exam
- Student classroom behavior

When considering your success as a teacher, which factors do **you** feel are most important?

Please rank from most important (#1) to least important (#5).

- Student internal assessments (for example: unit tests, classroom assessments)
- Classroom environment
- Your relationships with students/families
- Student test scores on state Forward exam
- Student classroom behavior

I summarize responses to these two questions in Appendix Figure A25, where each bar in the figure is the share of respondents who selected each choice as their number one option. The order of options was randomized, but each respondent saw the same order for the two different questions.

Next, respondents were asked questions regarding their workplace:

Berkeley
UNIVERSITY OF CALIFORNIA

How would you agree/disagree with the following statements:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I am satisfied with the work environment in my school.	○	○	○	○	○
I am satisfied with my salary.	○	○	○	○	○
I have agency over which work duties I'm given.	○	○	○	○	○
I feel secure in my current job position from being laid off.	○	○	○	○	○
I have discretion in setting my classroom's curriculum.	○	○	○	○	○

Table A13 columns 2–6 report the results by certification status. The share in the table corresponds to those who selected "Strongly agree" or "Agree."

Next, respondents were asked about unions:

Berkeley
UNIVERSITY OF CALIFORNIA

Next, I would like to ask you a few questions about labor unions.

Should workers have the right to form unions in the workplace?

- ○ Yes
- ○ No
- ○ I don't know

Is the teachers' union in your school district legally certified as the exclusive bargaining representative?

- ○ Yes, the union is certified
- ○ No, the union is not certified
- ○ I don't know

Have you been a member of the teachers' union in the past five years?

- ○ Yes
- ○ No

The share answering "Yes" to the first question about the right to form unions appears in Table 4 column 1. If a person answered yes to the last question, they saw the following:

Berkeley
UNIVERSITY OF CALIFORNIA

Why are/were you a member of the teachers' union? (Check all that apply)

- ☐ Better pay and benefits
- ☐ Representation in case of conflict
- ☐ A voice in the decision-making process
- ☐ I don't have a choice
- ☐ Co-workers are members
- ☐ Job security
- ☐ I felt it was the right thing to do
- ☐ write-in

Respondents could check any number of options or write a comment in the text box. The results from this question appear in Figure 9. Finally, respondents were asked whether they had union representation if investigated for disciplinary action:

Do you currently have the right to union representation if being investigated for disciplinary action?

- ○ Yes
- ○ No
- ○ I don't know

The share who responded yes to this question appears in Table 4 columns 2–4, where the last two columns are broken out by whether the person said yes or no to being a union member in the past five years.