5

Adult Culture Wars and Student Achievement

> It's negative, it's wasteful, it's not productive, it's sucking the energy out of the administration, the board, our staff, our teachers.
>
> —Mentor (Ohio) School board member Mary Bryner, on recent curriculum debates (Natanson, 2023*a*)

Darryl Adams, then-superintendent of the Coachella Valley Unified School District in Southern California, received an unexpected letter in November 2013. The correspondence was signed by the legal director of Arab-American Anti-Discrimination Committee (ADC), and the topic was the mascot of the district's lone high school – the "Arab." The image of the mascot featured a cartoon face of a bearded man with a pronounced hook nose wearing a traditional Middle Eastern head-covering. "All of these are examples of gross stereotyping, which must not be tolerated, and must immediately be addressed," the letter demanded. "ADC strongly believes the use of the word and such imagery perpetuates demeaning stereotypes of Arabs and Arab Americans. The 'Arab' mascot image is a harmful form of ethnic stereotyping which should be eliminated."

To be accused of racism came as a surprise to Adams, to say the least. The superintendent was an African-American man who had grown up in the deep South, reported to a majority–minority school board, and oversaw education for an overwhelming Latino student population. The "Arab" mascot was selected in 1921 as an honorific – to pay homage to Algeria and the Middle East as the original source of the date palm, one of the major agricultural crops first grown in the region. For the

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press

(a) Original Mascot (b) Updated Mascot

FIGURE 5.1 Coachella Valley High School "Arab" Mascot

next ninety years, it apparently never occurred to anyone that it might be offensive. The ADC said the group understood this history, but nevertheless expected action: "[I]n the 21st Century, such justifications for these actions are no longer tolerable."

The controversy immediately attracted national and even international attention – with prominent coverage in outlets including *Al Jazeera* and *Buzzfeed*. Graduates of the high school rallied behind their beloved mascot. The issue would come to occupy months of Adams' time, involving shuttle diplomacy to work out a compromise that both sides could ultimately live with. Eventually, they would settle on a new "Mighty Arab" mascot (see Figure 5.1). "It was a good lesson for the kids, for the adults, for the community," Adams recalled nearly a decade later.

But it is interesting to think about what Adams may have spent that time on instead. After all, school district administrators are very busy people and their time is arguably one of the scarcest resources in public education. Every minute spent negotiating a new high school mascot is a minute that could've been invested on improving instruction or revising curriculum with the goal delivering a better education in the classroom. When I asked about this type of zero-sum dynamic, Adams insisted that the controversy had not disrupted the district's academic performance – but did agree that similar distractions could have negative consequences on learning, particularly in places with more combustible racial relations.[1] "It depends on the leadership in the community," he told me. "I

[1] The year before the controversy, Coachella Valley had passed a bond measure that provided funding for a program to provide every student with an iPad. The initiative brought acclaim from the Obama White House, but within a few years, teachers began complaining about a botched implementation and inadequate training. I continue to wonder if, absent the mascot controversy, the iPad program roll out would've gone more smoothly. Adams would ultimately leave for medical reasons, and the board would

can see how, in some situations, it can be tough and can take away from the academic goals."

One such situation occurred in the Dover Area School District, located in a southern Pennsylvania. In 2004, the Dover school board majority adopted a change to the district's high school biology curriculum to include a statement describing "intelligent design" as an alternative theory to evolution and adopted the controversial intelligent design textbook, *Of Pandas and People*, as a reference text. The decision would lead to a lawsuit, and ultimately, a federal court case that would be nicknamed the "Dover Panda Trial" – a spoof of the famous "Scopes Monkey Trial" over evolution that took place in Tennessee nearly a century earlier.

The litigation became a circus – even Bruce Springsteen mentioned the case during a live performance – and turned tiny Dover into a center of media attention, fueling deep political divisions within the community. School board member Bill Buckingham, a pugnacious retired cop who led the charge on behalf of intelligent design, was transparent about the religious motivation for the policy change. "This country wasn't founded on Muslim beliefs or evolution," he told a local reporter. "This country is founded on Christianity, and our students should be taught as such" (Lebo, 2008, pp. 24–25). Yet at trial, he denied religion had anything to do with it. Confronted with his lies, Buckingham admitted he had been deep into an OxyContin addiction at the time of the policy change.

The school district would go on to lose the case and be ordered to pay the plaintiff's legal cost. (They generously asked for only $1 million, half of the actual expense, but a significant sum for a small school district.[2]) Amidst the controversy, voters booted out the anti-evolution incumbents, bringing in a slate of new board members, who immediately fired the superintendent, a man they accused of enabling the previous board majority,[3] and other top administrators, plunging the district deeper into chaos. "It was a huge distraction," Christina Kauffman, who covered the case for the *York Dispatch*, explained. "In terms of the kids, their educational experience was out there for the entire world. It wasn't just the media coverage of it, it was the fact that the adults – the ones who were

fire his top deputies – a lingering consequences of divisive salary negotiations with the district's teachers. "Union issues," as Adams described them.

[2] For context, the district's total spending was less than $40 million at the time.

[3] Like Buckingham, the superintendent had also testified that he never heard the board members discuss a religious motivation behind the policy change, although such discussions were clearly documented by local journalists who had attended the same meetings.

supposed to know what was going on – the adults were fighting. And much like a child growing up in a household where people are fighting, it's not conducive to learning."

I open this chapter with the controversies in Coachella Valley and Dover because they illustrate an important tension at the root of our public education governance institutions. In a well-functioning democracy, the decisions and actions of government officials must reflect the wishes of the governed.[4] Designing political institutions to achieve this goal, however, can require difficult trade-offs. What if increased responsiveness to public opinion results in the politicization of agencies that actually implement policies and impedes their performance (e.g., Gailmard and Patty, 2007; Lewis, 2008)?

Balancing political accountability with effective service delivery poses particularly difficult challenges in the context of public education. As I've argued throughout this book, the core democratic dilemma is that public schools exist to serve the educational needs of children but these students generally cannot vote in the elections through which key policymakers are chosen. Thus, dynamics that we would normally view as desirable in a democracy – elected officials' responsiveness to voter political preferences – could encourage school board members to prioritizing adult interests at the expense of student learning. For example, defending an offensive mascot or pushing through controversial curricular changes. This especially true when adults prioritize symbolic issues – be they cultural, religious, or identity-based – that are largely unrelated to the core academic functions schools are expected to serve.

This chapter examines how high-profile adult political conflicts ultimately affect student learning in the classroom. Traditional theories of democratic accountability do not provide clear expectations in this regard. On one hand, by bringing attention to the inner workings of the school system, high-profile controversies may pique voter interest in local school elections, increasing accountability pressures and diluting the corrosive influence of special interest groups who would otherwise play a dominant role in what are typically low-interest, low-turnout elections (e.g., Anzia, 2011; Berry, 2009; Moe, 2006). On the other, by shifting attention away from student outcomes to issues of interest to adults – including adults without children, the median voter in local

[4] There is considerable debate among philosophers about whether public officials should be most responsive to the (sometimes uninformed) *opinions* or to the *interests* of constituents (Pitkin, 1967).

school board elections – the controversies could distract teachers and district administrators and create unnecessary chaos and staff turnover, disrupting student learning.[5]

The evidence, I show, points firmly in toward the latter dynamics – and the mascot and evolution controversies help illustrate the likely mechanisms. In public schools, as in most government agencies, personnel time and financial resources – in the Dover example, literally $1 million in legal fees – are the ultimate constraints on the quality of public services. On the margin, a district must decide whether to spend a dollar on litigation or an artist to design a new mascot versus spending the same dollar in the classroom. In addition, divisive political conflicts can lead to churn both at the school board level and also among top district administration, and such turnover in itself can be highly disruptive. Certainly, the loss of both the superintendents and their deputies seemed to affect the operations of Coachella Valley and Dover. When superintendents are forced out, other district administrators usually go into a holding pattern, worried whether they'll still have their jobs once the next leader arrives, and avoid making any high-stakes decisions that could stir up controversy. School principals and teachers face the challenge of navigating difficult political terrain and dealing with the potential drama the political conflicts create among the families their schools serve.

The empirical analysis in this chapter combines a decade of data on high-profile "culture war" controversies in local public education that have been tracked as part of the Cato Institute's Public Schooling Battle Map with information on student academic achievement in grades three through eight from the Stanford Education Data Archive (SEDA) (also known as the Stanford Educational Opportunity Project). Unlike much of the analysis earlier in the book, which is descriptive, here I leverage variation in the timing and location of these controversies to estimate the plausibly causal effect of such conflicts. Specifically, the analysis implements what researchers describe as a "difference-in-differences design" – comparing the trajectory of student test scores in the affected districts before versus after the outbreak of each controversy to contemporaneous achievement trends in a "control group" of school districts that do not experience a similar controversy over the same time period.

[5] For example, Berry and Howell (2007) note that media coverage of student achievement plays an important role in making academic outcomes salient to voters and the weight voters put on achievement when deciding whether to reelect incumbents, as I discuss in Chapter 2. By making culture war issues more salient, controversies may cause voters to place less weight on academic achievement.

To preview the findings, I show that local political controversies hurt student performance on state exams, particularly in mathematics. The decline in achievement, on the order of 0.02 to 0.03 standard deviation units, is roughly equivalent to 10 days of learning out of a standard 180-day school year. Moreover, these effects persist for at least four years after the focal event. Examining potential differences both in the nature of the controversies and the students affected, I find that the learning disruption is particularly pronounced in the wake controversies related to racial issues and the teaching of evolution, which both reduce math achievement by between 0.06 and 0.07 standard deviations. However, I do not find evidence that the disruptions impact disadvantaged students more than their peers, leaving the magnitude of the achievement gap between low-income and other students largely unchanged.

MEASURING CULTURE WARS

The data on local education controversies comes from the Cato Institute's Public Schooling Battle Map, which was assembled by Neal McCluskey, the director of Cato's Center for Educational Freedom. McCluskey used a combination of Google News alerts, direct reports from individuals, and the ChoiceMedia.tv "Newswire," an aggregator website for education-related news. As such, the data is largely limited to sufficiently high-profile controversies that attract media attention and likely misses many smaller conflicts in districts with less media coverage. It is thus useful to think of the "treatment" that causes a district to enter the dataset as a bundle that includes both the controversy and the media attention that follows. While the database includes incidents dating back to 2005, it is most complete starting in 2011. Each controversy was classified by McCluskey into one of nine categories: (1) freedom of expression; (2) religion; (3) curriculum; (4) reading material; (5) race/ethnicity; (6) moral values; (7) gender equity; (8) sexuality; and (9) human origins.[6]

It is useful to highlight some of the specific cases, to illustrate the type of school district dramas these categories actually represent. One racial/ethnic controversy, for example, involved an Ohio district where

[6] Events are coded into categories based on the dimension of the controversy that is judged to be most central. For example, several cases in which districts prohibited students from wearing rosary beads are coded as "freedom of expression," although the expression in those cases is religious in nature. Similarly, complaints about books are coded as "reading material" disputes even if the specific books in question are controversial because they deal with racial or sexual topics.

middle school students sang "Cotton Needs Pickin" at the fall concert, to the outrage of many African-American parents in the audience. Another, from the same category, was the tragic case from a Southern California district that disciplined students who skipped school to attend an immigration reform rally. One of these students, apparently traumatized by his punishment, ended up taking his own life, prompting a high-profile lawsuit from his family. Many others dealt with controversial school mascots – in particular, the "Redskins," a popular team name that fell out of favor in recent years.

Freedom of expression controversies often involved school uniform or dress code infractions, including a female student disciplined for wearing a shirt that said "Twin Peaks" on one side and "Save the scenic views" on the other to raise breast cancer awareness. Another example was a school bus driver suspended for refusing to remove a Confederate flag from his truck, which he parked in the school parking lot. Critically, although the term education "culture war" often has an ideological valence in popular use – usually referring to conservative activism – my analysis includes controversies initiated by activists from both sides of the ideological spectrum. For example, many issues targeted by groups on the political left focus on equity and racial justice concerns, similar to considerations that motivated the ADC letter to the Coachella Valley superintendent. (The "Arab" mascot controversy is on one of the observations in the dataset.)

To examine how these controversies affect student academic achievement, I merge the Battle Map dataset with records from the SEDA (Fahle et al., 2021). This collection includes measures of student academic achievement based on test scores in math and English language arts in grade three through eight that states report to the US Department of Education.

Because this vintage of the SEDA data covers the years 2009 through 2018, I impose several data restrictions for technical reasons.[7] First, I exclude from the analytic sample districts that are "always-treated" – dropping all districts with a controversy observed in 2009 or earlier years. Second, I limit the sample to districts that experience only one controversy during this time period, excluding school systems with more than one event in the data.[8] In addition, I exclude state-level controversies

[7] Specifically, the motivation for these restrictions is to avoid bias due to treatment effect dynamics and heterogeneity.

[8] This is done for technical reasons, since we cannot assume that multiple controversies have an additive effect.

that are likely to affect all districts in a given state.[9] In effect, the analysis compares districts that experience one local controversy during the years included in SEDA compared to a control group made up of "never-treated" school districts that do not experience a controversy. The final sample includes approximately 520 such episodes. The total number of events ranges from 30 to 80 per year, and the most common type of controversy focuses on freedom of expression, which accounts for nearly a third of all of the events in the dataset.

EMPIRICAL STRATEGY

An obvious challenge to studying the effects of political controversies is that such events are unlikely to be random. For example, larger districts located in the core of their newspaper circulation area or television media market are probably more likely to see a run-of-the-mill, small-scale conflict escalate after attracting press coverage. In addition to district size, McCluskey (2019) finds that both student racial composition and the partisanship of each district's surrounding county predict the outbreak and frequency of events (see also Table 5.1).

My empirical approach directly accounts for such nonrandom patterns. Casual readers who are not interested in the technical details behind the analysis can feel free to skip the remainder of this section, which describes what is going on under the statistical hood in the analysis.

Specifically, I estimate a difference-in-differences model (Angrist and Pischke, 2009) that controls for time-invariant factors that could affect both the probability of an outbreak of a political controversy and student achievement. In most specifications, I use the following Ordinary Least Squares model:

$$Y_{dgst} = \alpha_d + \beta \text{Controversy}_{dst} + \delta_{st} + \varphi_g + \epsilon_{dgst},$$

where Y_{dgst} represents test scores of students in grade g in district d located in state s during year t. The model includes both district (α_d) and state-by-year (δ_{st}) fixed effects. The latter accounts for both potential changes in the assessments used by individual states over time as well as the impact of state-level policy changes and statewide education-related

[9] As discussed below, my statistical model controls for state-by-year fixed effects that absorb the effects of these statewide controversies.

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press

TABLE 5.1 *Summary statistics, measured in 2010*

		Excluded from analysis		Included in analysis	
	Full SEDA sample	Always treated	Multiple events	Never treated	Treatment districts
Total enrollment (grades 3–8)	1,695.40	4,022.44	17,581.48	1,089.52	3,666.09
Ave. enrollment per grade	287.10	670.88	2,930.25	184.93	612.43
Attending urban schools (%)	0.06	0.15	0.34	0.04	0.14
Attending suburban schools (%)	0.20	0.40	0.40	0.18	0.35
Attending town schools (%)	0.17	0.20	0.06	0.18	0.20
Attending rural schools (%)	0.57	0.26	0.19	0.60	0.31
White (%)	0.75	0.70	0.60	0.75	0.69
Black (%)	0.08	0.12	0.16	0.07	0.11
Hispanic (%)	0.12	0.13	0.17	0.12	0.15
Econ. dis. (ED,%)	0.48	0.43	0.43	0.49	0.44
English learners (%)	0.05	0.06	0.08	0.04	0.06
Special ed (%)	0.14	0.14	0.13	0.14	0.14
Ave. math achievement (grade 5)	−0.00	0.07	0.09	−0.01	0.07
Ave. ELA achievement (grade 5)	−0.01	0.05	0.08	−0.01	0.04
Math Non-ED-ED achievement gap (grade 5)	0.48	0.55	0.64	0.47	0.53
ELA Non-ED-ED achievement gap (grade 5)	0.52	0.58	0.67	0.50	0.56
Districts	13,105	185	367	12,031	521

political controversies. I pool data across all tested grades and some specifications also include grade-specific fixed effects (φ_g), although the results are not sensitive to this choice. $Controversy_{dst}$ is a binary indicator that takes the value of one starting in the year of each controversy. The variable remains "on" in all subsequent years, although I also estimate more flexible event study specifications that examine the dynamics more carefully. All standard errors are clustered by school district to account for serial correlation (Bertrand, Duflo and Mullainathan, 2004).

The primary quantity of interest is the estimate of β. This can be interpreted as the causal effect of controversy on student academic achievement under the assumption that student performance in the affected districts would have followed trends parallel to those in districts not experiencing an event. To probe the plausibility of this assumption, I examine the trajectory of achievement in the affected districts compared to never-treated controls in the years leading up to the controversy using the following event study specification:

$$Y_{dgst} = \alpha_d + \sum_{j \leq -5}^{\geq 5} \pi_j \text{Controversy}_{dst} \cdot 1(\tau_{dst} = j) + \delta_{st} + \varphi_g + \epsilon_{dgst}.$$

Given the staggered nature of the treatment and the relatively short duration of the achievement panel, I bin the event window end points, combining all years at least five years before and five years after each event into two categories.

Table 5.1 provides a summary of the full dataset as well as the analytic sample. While each district is observed for multiple years and test scores are reported across multiple grades, the table summarizes district characteristics in 2010 and student achievement in the fifth grade. Several patterns clearly stand out in the data. First, districts that experience a controversy attracting media attention are considerably larger, more urban, and enroll a smaller share of white students than districts that serve as the control group. On the other hand, treated districts are also somewhat wealthier, as measured by the share of students designated as economically disadvantaged, and have achievement levels that are modestly higher than for the districts in the control group. In addition, the treated districts have somewhat larger achievement gaps between higher- and lower-SES students. Second, districts that experience multiple events during the period of the study and are thus excluded from the analytic

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press

sample are particularly large – enrolling more than 17,000 students on averages in the tested grades – and are especially urban.[10]

CONFLICTS AND TEST SCORES

Table 5.2 reports the standard difference-in-differences estimates. The first two columns show results for math scores while the latter two columns present comparable estimates for English language arts (ELA) achievement. Overall, student math achievement declines by approximately 0.018 standard deviations in the years after a local political controversy. The estimates are the same regardless of whether the model includes grade-level fixed-effects. By contrast, there does not appear to be any impacts on ELA achievement. While the divergence between math and ELA scores may be surprising, it is consistent with other education policy research showing that student performance in math is much more sensitive to policy interventions than ELA achievement (see Fryer, 2014 for overview).[11] Although education scholars have offered a number of plausible explanations for this pattern, the most likely is that students acquire a much larger share of their mathematical knowledge inside the

TABLE 5.2 *Effect of local culture war controversy on student achievement*

	(1) Math	(2) Math	(3) ELA	(4) ELA
Controversy	−0.0183***	−0.0183***	−0.00557	−0.00556
	(0.00668)	(0.00668)	(0.00484)	(0.00484)
Observations	522,359	522,359	548,946	548,946
R-squared	0.035	0.037	0.041	0.046
Number of districts	11,215	11,215	11,252	11,252
Time FE	State-year	State-year	State-year	State-year
Grade FE	No	Yes	No	Yes

Robust standard errors clustered by district in parentheses.
***$p < 0.01$.

[10] Since these large urban districts are excluded from the analysis, caution is warranted in generalizing the findings to these types of districts.

[11] Nearly every analysis of pandemic-related learning losses has found much larger test score declines in math than in ELA, for example.

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press

classroom, in contrast to activities outside of school that complement formal reading instruction.

It is useful to put the effect size into more intuitive units. Hill et al. (2008) provide national benchmark for typical learning gains in both math and ELA during each year of schooling. Focusing on math achievement growth in just grades three through eight – those covered by the SEDA data – student scores increase by about 0.39 standard deviations per year, on average. The estimated effect of 0.018 corresponds to roughly 5 percent of annual gains, which translates to about 10 days of learning assuming a typical 180-day school year.

Figure 5.2 presents estimates from the event study specification. The top panel reports results for math achievement while the bottom panel focuses on ELA scores. Encouragingly, there is no evidence that test scores begin to decline in the years prior to the outbreak of local controversy. However, math scores show a clear decline starting in the year after the event, and they remain significantly lower for at least four years. Although the individual point estimates are quite noisy, there is no evidence of a rebound in the short term – if anything, the impact appears to grow more negative over the first three years. The figure does find some evidence of pretreatment trends, raising some questions about the plausibility of the parallel trends assumption. I will return to this issue, presenting additional analyses and probing the robustness of the results, later in the chapter.

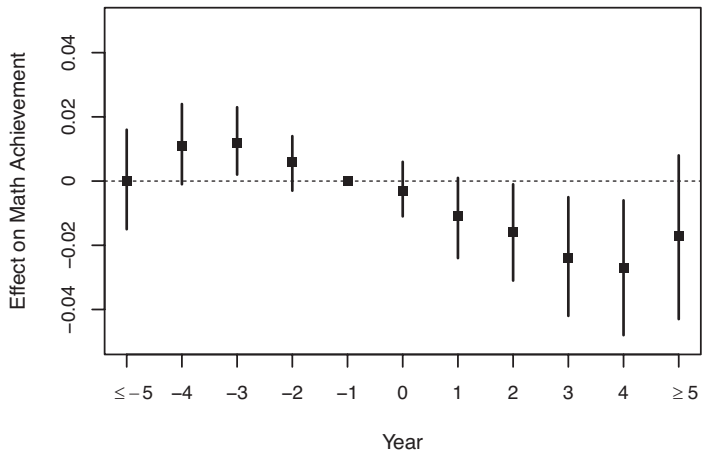
It is important to note that the Cato dataset records only the calendar year of each event, with no additional information about the precise day or month for most of the observations, while the test score data is reported based on academic years.[12] Thus, controversies that occur in the fall of calendar year t can affect achievement no earlier than the following spring, the first testing cycle after the event, which would show up as school year $t + 1$ in the SEDA data. That may explain why effects for year zero are relatively small and not significant and why the impact does not appear until the following spring testing window.

Which Controversies Affect Learning, and for Whom?

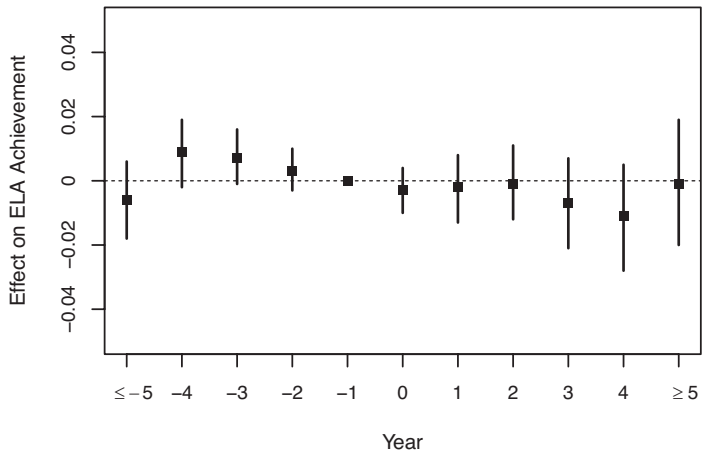
The results presented thus far focus on average achievement. It is important to consider whether the learning disruptions disproportionately

[12] The database began including full dates starting in 2017. For the available years, more than 60 percent of the events occurred in June or later, after the testing cycle had already concluded for that calendar year.

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press



(a) Math



(b) ELA

FIGURE 5.2 Event study plot for the effect of local controversy on student achievement

affect high-risk students. Because the SEDA dataset reports subgroup-specific achievement estimates only for subgroups with at least twenty students, it is not possible to examine race-specific data in most districts. However, many districts to enroll a sufficient number of economically disadvantaged students, proxied in most states by their participation in the federal free- and reduced-price meals program and family participation in cash welfare and Medicaid.

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press

These results shows nearly identical declines in the math scores of both economically disadvantaged students and students not identified as such. As a result, there is no meaningful change in the achievement gap between these student subgroups. Consistent with the above results, there are no effects on ELA achievement for either group. When reestimating the effects separately for each grade, the declines in math achievement appear to be concentrated in elementary school grades, which is consistent with prior research showing that the learning of younger students is more sensitive to educational disruptions compared to the impact on older students.

Local political controversies can affect student learning through a variety of possible channels. First, school boards may replace senior leadership that becomes the focus of community opprobrium, producing turnover in administrative ranks and subsequent disruption in district operations. Voters may also replace current school board incumbents, which can also set in motion a chain reaction of other personnel changes. Second, the controversy can attract a disproportionate share of attention, energy, and resources, distracting school and district leaders from their primary focus on improving instruction and student outcomes. Third, if the controversy surrounds specific curriculum or teaching materials, districts may respond by adopting new curriculum, resulting in significant retraining needs, prep time, and transition costs for educators, which may negatively impact their productivity in the short term. Fourth, the negative publicity may cause higher-achieving students to leave the districts, mechanically reducing the average achievement.[13]

That the decline in test scores is of comparable size for both economically disadvantaged and wealthier students – who are much more likely to have the resources necessary exit the public schools for private options or to move to a different district – provide some evidence against the last explanation. I examined several other outcomes that speak to this question as well. Focusing on total enrollment, I found no evidence that enrollment declined in the districts after a political controversy. In fact, total enrollment actually increased by between 2 and 3 percent after each event.[14] There were also no changes in the demographic and

[13] The version of the SEDA data I use in the analysis includes charter schools located within the geographic boundaries of districts, so students switching from public to nearby charter schools would not affect the results above.

[14] It may be that growing districts, which experience a large influx of new families unfamiliar with local cultural norms, may be particularly likely to have small issues escalate enough to attract media attention.

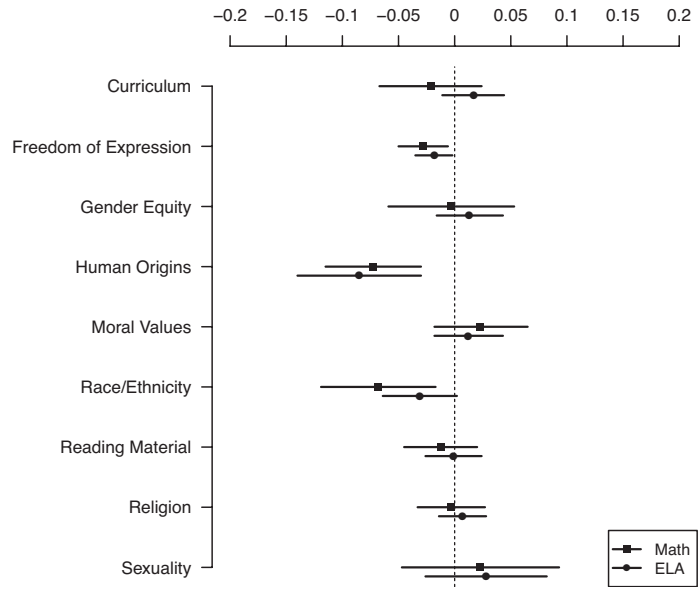


FIGURE 5.3 Student achievement impact by the topic of controversy

socioeconomic composition of students served that would explain the changes in achievement.[15]

Unfortunately, there are no national data on staff, administrator, or school board turnover nor on curricula used by individual school districts, so it is not possible to directly test the remaining three possibilities. However, estimating separate effects for each type of controversy can provide some suggestive evidence about what's actually going on inside school districts.

These results are presented in Figure 5.3. Interestingly, most types of controversies appear to have little if any effect on student learning. Importantly, this is true for controversies surrounding instructional curriculum and those related to reading materials, suggesting that transition costs related to changes in pedagogy or teaching materials are unlikely to be driving the overall findings. Instead, large negative effects appear only for two types of controversies – those dealing with human origins and evolution and disputes surrounding race. As I noted above, that the latter category mostly includes fights unrelated to the actual content of

[15] Although there is a marginally significant decrease in the share of white students, the absolute magnitude of this decline is very small – less than 1 percentage point – and so cannot explain the decrease in test scores.

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press

instruction. In addition, there are smaller but still significant declines following controversies dealing with freedom of expression.[16] Overall, these results suggest that staff turnover and/or the diversion of attention and resources away from learning and day-to-day instruction are likely to represent the most plausible explanations for the observed declines. These mechanisms are consistent with my impression of the aftermath of the events in Coachella Valley and Dover.

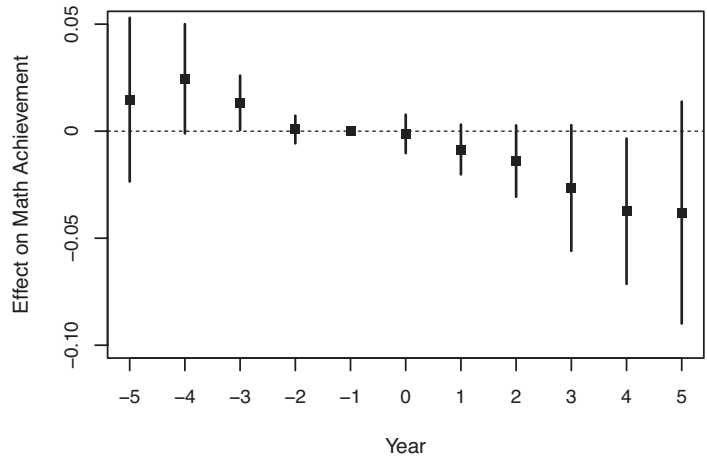
It is also important to note that the point estimates for the declines linked to evolution and racial controversies are considerably larger, on the order of 0.06 to 0.07 standard deviations. These are substantively significant, corresponding to nearly 1.5 months of learning in mathematics in the affected grades. For these types of controversies, there is also evidence of declines in ELA scores, although the point estimates is considerably smaller for ELA achievement for racial controversies.

GETTING MORE IN THE TECHNICAL WEEDS

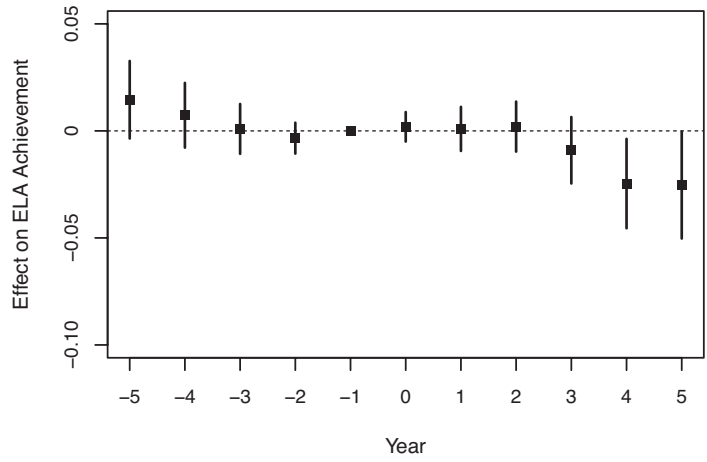
The estimates discussed above come from what is known as a two-way fixed effects (TWFE) regression. A flurry of recent papers in economics have shown that TWFE approaches can breakdown when estimating treatment effects with staggered timing – precisely the application in this chapter. In this section, I address this concern and show that the main results are robust to alternative modeling approaches that address the problems with TWFEs. I also discuss other evidence to support the validity of my findings. As before, readers not interested in the technical details should feel free to skip to the next section.

I begin by replicating the event-study results using an alternative approach developed by de Chaisemartin and D'Haultfœuille (2020). These results are presented in Figure 5.4. Although these estimates are noisier – as we would expect, since the de Chaisemartin and D'Haultfœuille (2020) method uses fewer observations – they are qualitatively quite similar to those produced by TWFE. The figure also shows less noticeable pretreatment trends in the years immediately before each controversy, providing additional support for the parallel trends

[16] Although the decrease is much smaller than for racial and human origins controversies, there are far more events in this category, so the aggregate impact on achievement is nevertheless substantial.



(a) Math



(b) ELA

FIGURE 5.4 Alternative event study estimates for the effect of local controversy on student achievement

assumption necessary to interpret these results as causal effects (rather than mere correlations).

Nevertheless, skeptical readers may still have doubts about whether parallel trends hold. As I noted above, districts hit by a scandal tend to be growing faster than their counterparts in the years immediately prior to the controversy. Maybe there are some unobservable (time-varying)

https://doi.org/10.1017/9781009606349.005 Published online by Cambridge University Press

differences between districts with versus without controversies that confounds the results? One way to address this concern is to examine only districts that experience a controversy and leverage the differences in the *timing*. In other words, we can use "later-treated" school districts as the control group for "earlier-treated" districts, and vice versa (see, e.g., Harvey and Mattia, 2024).[17] Using this alternative approach, I find nearly identical results – a 0.02 standard deviation decrease in math scores in the years after each controversy. In addition, the negative point estimate on ELA scores also becomes statistically significant, corresponding to a 0.015 standard deviation decrease in scores.

IT SHOULD BE ABOUT THE STUDENTS, STUPID

To summarize, this chapter documents that high-profile political controversies dealing with issues only tangentially related to education have a small but persistent negative impact on student learning. For many readers, these results are probably not surprising. After all, a growing literature in other policy domains examining how political controversies affect the quality of public services at the local level tends to find the same thing. Much of this work focuses on policing and has found that public protests and investigations in the wake of high-profile police shootings of civilians can reduce officer effort and proactivity, resulting in higher crime rates (Campbell, 2024; Devi and Fryer, 2020; Rivera and Ba, 2023; Shi, 2009). This finding is known as the Ferguson effect, named after the city that experienced a rise in crime after the death of Michael Brown, an African-American eighteen-year-old, at the hands of a police officer. To my knowledge, this chapter provides the first evidence of similar dynamics in the context of public education. In one media interview about these results, I described them as the "Loudoun County effect" – after a wealthy Virginia community whose school system was thrown into turmoil over sexual assault allegations some parents attempted to link to the district's transgender bathroom policy (Homans, 2023).

These findings have broader implications for debates surrounding educational governance and policy. On the policy front, they should make activists cautious about well-meaning efforts to push local schools to confront important social justice concerns. Although the most salient current

[17] Note that this is the exact opposite of the de Chaisemartin and D'Haultfœuille (2020) approach. That all three methods – standard TWFE, de Chaisemartin and D'Haultfœuille (2020), and earlier- versus later-treated districts – produce substantively identical estimates provides strong evidence for the robustness of the results.

educational policy controversies – including debates about Critical Race Theory and instruction about sexual orientation and gender identity – have a conservative bent, many of the conflicts in the Cato database originated from progressive complaints and umbrage. Such forms of activism may be counterproductive. In recent years, for example, many have raised objections over racial disparities in student discipline (e.g., Liu, Hayes and Gershenson, 2022) and the presence of armed school resource officers in school buildings (e.g., Weisburst, 2019). I have little doubt that advocates working on these issues are sincere and genuinely concerned about student interests. However, it is important to consider the unintended consequences and learning disruptions produced by local controversies that can arise surrounding debates over such issues. Even if the advocacy is ultimately successful, it is not obvious that the net effect for students is positive in the end.

In this chapter, I have offered what I hope is convincing evidence that when adult political conflicts cause school district leaders to take their eye off the ball of educating students, learning suffers. The statistical analysis admittedly offers much less direct evidence of the underlying mechanisms, and so I want to close with two additional case studies that I think help illustrate these mechanisms. As it happens, both caught my attention just as I was finishing this chapter.

In June 2023, PBS "Newshour" ran a segment titled "School boards become battlegrounds for nation's divisions." It featured another rural Pennsylvania school district, Penncrest, and deep conflict within the community over the school board's recent decisions to ban sexually explicit books from the school libraries and also require that transgender athletes to participate in sports teams that match their biological sex. This was explosive stuff for this small, tightly knit community – and it immediately reminded me of the drama in Dover nearly twenty years earlier. Near the end, the segment quoted Penncrest school board member Jeff Brooks, a military veteran and life-long Republican who had opposed both policies. "The biggest effect that [the policy controversy] had is that it has taken time away from our administration being able to focus on the classroom," Brooks said.

Brooks' argument – adult culture wars are taking time and attention from learning – is exactly the point I wanted to make in this chapter, so I reached out to him to learn more. He graciously agreed, and our conversation helped illustrate the kinds of dynamics I worry about most.

When we spoke a few weeks later, I asked Brooks to walk me through specific examples of how the controversies surrounding the new

policies had negatively affected the district's students. He had plenty of receipts. "Every minute that we're spending on something outside of the classroom is time that we could be spending on a curriculum, on teacher development, on classroom observations, on improving literacy skills," he explained. "We don't pay the administration to cater to the board. . . . Back in old times, when I went to college, one of the things that stuck from one of my economics classes was opportunity cost." For example, to implement the district's new book policy, six building principals, three assistant principals, and three librarians had to spend two weeks going through every book in the district's libraries – time they could've directed to something else.[18]

And Brooks told me the latest controversy was just a continuation of other adult drama. Penncrest was once rated in the top 10 percent of the state for achievement but had dropped precipitously as a result of budget cuts, poor administration, and political dysfunction on the board. "Mostly it was just the lack of focus on education, the focus being on every other thing," Brooks said regarding his frustrations with the board. "We've been manufacturing different nonsense drama for a while. . . . With a school board like ours, you can't tell from month to month what the next controversy is going to be." And the dysfunction had other consequences. When the district last tried to fill its superintendent position, only two qualified candidates applied because of the district's reputation. Each of the applicants had their own baggage from prior districts. Uncertain about how long he would have the job, the district's prior superintendent didn't even bother moving in and instead lived out of a hotel.

Brooks had decided not to run for reelection himself. He was tired of being called a "groomer" for speaking out on behalf of LGBT students. As a public official, he did not feel like he could respond with the decorum and restraint an elected office required.

I found the Penncrest example illuminating. It had all of the pieces – misguided priorities, misallocation of scarce staff and financial resources, administrative turnover, negative impact on recruitment, school board member turnout – that I believe helps explain why adult political conflicts like these ultimately filter down to the classroom and negatively impact students.

[18] Another school district, in Virginia, estimated that it took a staff of eleven working forty hours a week all school year to respond to complaints filed by a "serial" book challenger upset about sexually explicit content (Natanson, 2023*b*).

Just a few days before my interview with Brooks, another LGBT-themed school controversy broke out on the other side of country. In Temecula, a wine-growing region in Southern California, a conservative school board majority had just voted against adopting a new social studies textbook because the supplementary material mentioned Harvey Milk, a murdered San Francisco County supervisor who had been the first openly gay elected official in US history. The school board president had called Milk a "pedophile" at the meeting – in apparent reference to Milk dating a sixteen-year-old – and the story blew up. It even attracted the attention of California's publicity hungry governor, Gavin Newsom, who called the board president ignorant on Twitter and announced the state would step in to buy the rejected textbooks and bill the district for the cost. The governor's spokesman noted the move was unusual but pointed to a little-noticed law that allowed the state to intervene when districts faced textbook "shortages." Because the lack of approval of the textbooks meant the district was stuck with old, outdated materials that did not comply with the most recent standards, the administration argued, the law applied.

Despite the attention focused on the social studies controversy, the chaos and dysfunction Temecula was much deeper. A few days after the textbook vote, the conservative board majority fired the district's superintendent. A few months later, the majority disintegrated – one member resigned to move to Texas for work, another lost a recall election by just a few hundred votes. The community seemed evenly divided and the conflict was playing out in the district. However, various political interests were happily to milk the controversy to for political gain.

In an insightful article, veteran California education journalist John Festerwald noted that Gov. Newsom seemed eager to wade into the culture war issue playing out in Temecula. Oddly, however, the governor had not taken similarly proactive action to address outdated reading curriculum – including those using the controversial "three-cueing method" I discuss in Chapter 4 – which arguably had a much bigger impact on the state's students and affected many more school districts. "An unknown number of districts are also using outdated textbooks or early literacy curriculums with disproven reading techniques," Festerwald noted dryly in his piece. But a spokesman for the governor said "the administration doesn't intend to respond similarly to them" (Festerwald, 2023).

The Temecula story seemed to be an uncanny repeat of the Penncrest experience. Adult culture wars were sucking all of the air out of the room, and strategic politicians were taking advantage of the situation to score

political points. All of the media publicity and attention was turning up the volume and the emotions. No one seemed interested in student learning.

But the recent LGBT debates also illustrate why these kinds of battles are unavoidable, given the incentives all of the key actors face. Consider the book *Gender Queer*, which has been the focus of much of the recent controversy over school book bans. The graphic memoir tackles the issue of gender dysphoria and has won wide critical acclaim. Trans rights activists insist it is important to make it available to students struggling with their gender identity, a long-overlooked and disadvantaged population suffering particularly high rates of suicide. However, several of the pages in the book are very graphic, including one depicting a scene in which a character wearing a strap-on gets a blowjob. Conservative activists believe the material is inappropriate for younger (e.g., elementary grade) children.

Certainly there seems to be room for good-faith disagreement – people who want to make the book available to trans kids are not "groomers," and parents concerned about age appropriateness are not "transphobes." Yet a decision has to be made, and there is no room for compromise – either the book is available in libraries serving younger children or it is not. One camp has to win and one has to lose. The more one side advocates for their position, the more the other feels like it must countermobilize. "Nothing enrages parents more than the idea that their children are being turned against them, and few things worry a partisan more than the fear the opposing party is using schools to inculcate its beliefs in the young," journalist Jonathan Chait summarized presciently in a recent article (Chait, 2023).

In some cases, it may be difficult to separate the (adult) political dimension of conflicts from reasonable disagreement about (student) academic goals. For example, disputes over Critical Race Theory and history instruction (discussed in Chapter 4) and evolution are fundamentally tied to content standards. How these disputes are ultimately resolved may impact student performance on tests aligned with those standards. Deciding whether or not to teach about intelligent design ultimately impacts how students perform on science exams (see Arold, 2024). In the real world, it may not always be possible to separate purely political arguments from legitimate academic or pedagogical considerations. But even in such instances, it is worth taking trade-offs into account. The benefits of adopting better curriculum, to take one example, must be

weighed against the learning disruptions and operational dysfunction the associated advocacy campaigns on behalf of such adoptions may cause.

Unfortunately, asking one side in a culture war to show some forbearance while other pushes full steam ahead can feel a lot like unilateral disarmament. And the Cold War analogy is actually quite appropriate. During this fraught period, both the US government and the Soviet Union had to choose between investing scarce resources in guns to retain parity in the arms race or in "butter" – other public services essential to social welfare. Both obviously preferred to spend less on guns and more on butter – but only if they believed the other side would do the same. As long as one side kept buying more guns, the other side had to do the same. Some observers have argued that the Soviet Union ultimately collapsed because it simply ran out of money – it could not spend as much as the arms race demanded while maintaining sufficient investment on other government services. Similar dynamics plague today's education culture wars.