ESSAYS IN APPLIED MICROECONOMICS

by

KYLE RAZE

A DISSERTATION

Presented to the Department of Economics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2022

DISSERTATION APPROVAL PAGE

Student: Kyle Raze

Title: Essays in Applied Microeconomics

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

Glen Waddell	Chair
Michael Kuhn	Core Member
Jonathan Davis	Core Member
David Liebowitz	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2022

DISSERTATION ABSTRACT

Kyle Raze

Doctor of Philosophy

Department of Economics

June 2022

Title: Essays in Applied Microeconomics

In this dissertation, I consider two potential sources of racial disparities: racially biased discipline at school and changes to voting rights protections.

In the first chapter, I examine patterns of racial disparities in school discipline. Racial gaps in the adjudication of student misconduct are well documented—for similar behaviors, students of color are more likely to be disciplined and discipline tends to be harsher. While students of color do receive harsher punishments, on average, I show that this differential depends on the racial composition of incidents. Consistent with administrators moving toward equal treatment when variation in race is more salient, multi-race incidents evidence no differentials. In fact, when a white student is implicated in the same incident as a student of color, punishments imposed on students of color are indistinguishable from those imposed on white students in all-white incidents.

In the second chapter, I turn to the effects of changes to voting rights protections on racial disparities in voter turnout. Existing research shows that the Voting Rights Act of 1965 increased turnout among Black voters, which then generated economic benefits for Black communities. In *Shelby County v. Holder* (2013), the Supreme Court invalidated the enforcement mechanism responsible for these improvements, prompting concerns that states with histories of discriminatory

election practices would respond by suppressing Black turnout. I estimate the effect of the *Shelby* decision on the racial composition of the electorate using triple-difference comparisons of validated turnout data from the Cooperative Congressional Election Study. The data suggest that the *Shelby* decision did not widen the Black-white turnout gap in states subject to the ruling.

This dissertation includes both unpublished co-authored material and previously published solo-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Kyle Raze

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene

DEGREES AWARDED:

Doctor of Philosophy, Economics, 2022, University of Oregon
Master of Science, Economics, 2017, University of Oregon
Bachelor of Arts, Political Science and Economics, 2015, University of Oregon

AREAS OF INTEREST:

Applied microeconometrics, labor economics, public economics

PROFESSIONAL EXPERIENCE:

Instructor, Department of Economics, University of Oregon, 2018–2022
Teaching assistant, Department of Economics, University of Oregon, 2016–2022

GRANTS, AWARDS AND HONORS:

Kleinsorge Summer Research Award, University of Oregon, 2020
Graduate Teaching Award, University of Oregon, 2020

PUBLICATIONS:

Raze, K. (2022). Voting rights and the resilience of Black turnout. *Economic Inquiry*, (60)3, 1127–1141.

ACKNOWLEDGEMENTS

To my grandfathers, Dean and Jim, for their wisdom and commitment to public service.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

DOES THE SALIENCE OF RACE MITIGATE GAPS IN DISCIPLINARY OUTCOMES?

EVIDENCE FROM SCHOOL FIGHTS

This dissertation includes both unpublished co-authored material and previously published solo-authored material. This chapter (Chapter 1) includes material that was co-authored with Glen Waddell: I performed the statistical analysis and we both contributed to the writing of the manuscript. Chapter 2 includes solo-authored material that is forthcoming publication in the July 2022 issue of *Economic Inquiry*.

1.1 Introduction

With an extensive literature identifying racial disparities in many outcomes, any degree of hysteresis in the production of racial disparities suggests that there are gains to correcting early differences in the experiences of those of different racial backgrounds. In this way, school environments are an important setting to consider—it is in these formative years that students are making human capital investment decisions and forming expectations of their own comparative advantages and relative strengths.[1] In this chapter, we examine patterns of racial disparities in the adjudication of student misconduct and provide suggestive evidence of a mechanism that can explain their origins.

While school discipline can mitigate externalities associated with disruptive behavior (Carrell & Hoekstra, 2010; Kinsler, 2013; Pope & Zuo, 2020), disciplinary interventions impose significant costs on disciplined students. Namely, the disciplinary actions commonly available to school administrators are inseparable from interruptions to the direct inputs into the production of human capital; suspensions, expulsions,

---

[1] For example, individual expectations about comparative advantage have been shown to influence consequential decisions about major choice in university settings (Arcidiacono, Aucejo, & Spenner, 2012; Card & Payne, 2021).

and other forms of exclusionary discipline decrease instructional time and disrupt the continuity of instruction. As a result, exclusionary discipline can hinder academic performance (Anderson, Ritter, & Zamarro, 2019; Bacher-Hicks, Billings, & Deming, 2019; Craig & Martin, 2019; Sorensen, Bushway, & Gifford, 2022; Steinberg & Lacoe, 2018). Exposure to harsh exclusionary discipline regimes has also been shown to decrease educational attainment and increase the likelihood of arrest and incarceration (Bacher-Hicks et al., 2019). Equity concerns notwithstanding, racially biased discipline could therefore lead to significant and long-lasting economic inefficiencies in the production of human capital.

The existence of race-based disparities in disciplinary outcomes is well documented (Anderson & Ritter, 2017; Barrett, McEachin, Mills, & Valant, 2021; Gopalan & Nelson, 2019; Kinsler, 2011; Liu, Hayes, & Gershenson, 2022; Ritter & Anderson, 2018; Shi & Zhu, 2022; Welsh & Little, 2018). Existing research suggests that a significant portion of the average discipline gap between white students and students of color arises across schools, as Black students are more likely to live in school districts with higher rates of exclusionary discipline (Anderson & Ritter, 2017; Barrett et al., 2021; Gopalan & Nelson, 2019; Kinsler, 2011; Ritter & Anderson, 2018). Yet gaps typically remain after conditioning on student characteristics and school fixed effects (Anderson & Ritter, 2020; Barrett et al., 2021; Beck & Muschkin, 2012; Gopalan & Nelson, 2019; Liu et al., 2022; Shi & Zhu, 2022), which leaves open the possibility that school officials treat students of color less favorably than white students who engage in similar behaviors. Even so, within-school comparisons may nevertheless fail to isolate the average response of school officials to the race of their students. Within-school disparities are consistent with differential treatment on the basis of race, but they are also consistent with systematic but unobserved differences in student behavior.

To adjust for unobserved differences in student behavior, several recent studies compare the punishments of students implicated together in the same incident. Barrett et al. (2021) use administrative data from Louisiana to compare the suspension lengths associated with incidents in which exactly two students were suspended for fighting on the same day in the same school. Shi and Zhu (2022) leverage the availability of incident identifiers in North Carolina to make within-incident comparisons among all incident types. As will be the case in our setting, the North Carolina data include cases that did not end in punishment, and having an incident identifier circumvents the need for a same-day, same-school, same-incident-type matching rule. Likewise, Liu et al. (2022) leverage incident identifiers in administrative data on student referrals, but from a large California school district. All three studies detect statistically significant within-incident differences in the number of days suspended (i.e., among fights that end in suspension in the case of Barrett et al., 2021, and among all incidents in Liu et al., 2022, and Shi & Zhu, 2022). Together, these studies suggest that school administrators exhibit biases that disfavor students of color when adjudicating cases of student misconduct. In terms of magnitude, they each find similarly-sized racial disparities in suspension lengths— Black students are suspended for roughly one twentieth of a day longer than white students implicated in the same incident.[2] Within-school disparities are typically larger in magnitude (e.g., Gopalan & Nelson, 2019; Anderson & Ritter, 2020; Barrett et al., 2021; Shi & Zhu, 2022).

That being said, within-incident race differentials are identified from a very specific subset of incidents—those in which there was variation in race. Thus, any identification strategy that relies on incident fixed effects necessarily decouples those

_____

[2] Shi and Zhu (2022) and Liu et al. (2022) also document statistically significant differences for Black students on the extensive margin of suspension, with smaller effects in North Carolina (Shi & Zhu, 2022) than in California (Liu et al., 2022).

incidents that have variation in race from those that do not. Moreover, when an administrator adjudicates students of color and white students "side-by-side," one would expect racial differences to be more salient, which may induce outcomes that are not representative of the adjudication of students of color in other contexts. For example, variation in the racial composition of incidents may lead to "signal jamming" behavior (Fudenberg & Tirole, 1986; Holmström, 1999), by which administrators anticipate that the most reliable signal of their treatment of race is likely to be found in their adjudication of multi-race incidents. If there are professional repercussions to exhibiting explicit biases, it would be in administrators' interest to pay closer attention to their treatment of students of color, which would move them toward equal treatment in multi-race incidents. Or perhaps administrators more easily suppress implicit biases when differences in race are more evident, which again implies that racial disparities in adjudication outcomes would vary across the racial composition of incidents. The psychology literature on preference reversals in joint evaluation (Bazerman, Moore, Tenbrunsel, Wade-Benzoni, & Blount, 1999; Hsee, Loewenstein, Blount, & Bazerman, 1999) also suggests that racial composition may directly matter to outcomes, insofar as the adjudication of students of color apart from white students leads decision makers to put less weight on equal treatment (i.e., a "difficult-to-evaluate attribute" in the spirit of Hsee et al., 1999) than they would in the joint evaluation of students of color and white students side-by-side.

From a variety of perspectives, there is good reason to anticipate direct effects of racial composition on outcomes. Regardless of the particular mechanism, however, one of the takeaways from our analysis will be that existing gaps in adjudication outcomes are not seeming to arise from differential treatment within multi-race incidents. By relying on within-incident variation to identify race-based differentials in the

16

adjudication of misconduct, we fear that within-incident comparisons identify the effect of race in an environment that is neither representative of the choices or incentives of administrators nor representative of the experiences of students of color.

We therefore seek to contextualize the adjudication of multi-race incidents by comparing adjudication outcomes across joint incidents that are similar but differ in their racial composition. In multi-race fights, we find no race-based differences in outcomes. However, the magnitude of the difference in punishments across same-race fights within schools is often large—fights that involve only students of color elicit harsher punishments than those involving only white students. In high schools, for example, students of color in fights involving *only* students of color receive suspensions that are two thirds of a day longer than those assigned to white students in all-white fights. Yet, in the same environment, being implicated with a white student renders the punishment of students of color indistinguishable from the punishment of white students in all-white fights. In other words, purging all within-incident disparities in punishment would do little to close the gap in disciplinary outcomes between students of color and their white peers.

A causal interpretation of these findings suggests that the salience of differences in race within an incident moves administrators toward the equal treatment of students on the basis of race.[3] By the same token, the large disparities across same-race fights suggest that within-incident comparisons can severely understate the extent of differential treatment.

---

[3] We note that the results of the supplemental student fixed effects analysis in Shi and Zhu (2022) are not necessarily at odds with our findings. In that analysis, the authors consider how *relative* differences in adjudication outcomes vary across the racial composition of incidents that a student participates in. They find that Black students receive harsher punishments *relative to the other student in an incident* when the other student is of a different racial background. By measuring outcomes relative to others in the same incident, rather than levels, this analysis leaves open the possibility that the absolute severity of punishments imposed on students of color decreases in multi-race incidents.

We proceed in Section 1.2 by describing the empirical environment and data we use to test for race-based differentials in punishment. We then present estimates of racial disparities in expulsions and suspensions in Section 1.3. Therein, we progress from raw race-based gaps in adjudication outcomes to within-incident differences. In Section 1.4 we incorporate same-race incidents to better understand the origins of the racial disparities we identify in Section 1.3. Finally, in Section 1.5, we discuss implications of our findings.

1.2 Data

To document racial differences in the severity of sanctions for alleged misconduct, we consider fighting infractions reported by public schools in Washington to the Office of the Superintendent of Public Instruction between 2014–15 and 2017–18.[4] Similar to Shi and Zhu (2022) and Liu et al. (2022), our administrative data include incident identifiers and infractions that did not result in students receiving exclusionary discipline. We complement Barrett et al. (2021), Shi and Zhu (2022), and Liu et al. (2022) by identifying disparities in a setting in which racial resentment is persistently lower (Smith, Kreitzer, & Suo, 2020).[5] While there are still significant race-based gaps in the adjudication of outcomes in Washington, both across and within schools, one might reasonably expect school administrators in Washington to respond differently to race than administrators in other states.

---

[4] For our purposes, public schools include traditional public schools as well as public charters and alternative schools—we exclude infractions from special education schools and juvenile correctional institutions.

[5] This is also consistent with the data collected by Project Implicit, which suggests that implicit racial attitudes in Washington are the second lowest among US states. For more information, see Chris Mooney, "Across America, Whites Are Biased and They Don't Even Know It," *Washington Post*, 8 December 2014, https://www.washingtonpost.com/news/wonk/wp/2014/12/08/across-america-whites-are-biased-and -they-dont-even-know-it/ (accessed 1 June 2022), and Jordan Axt, "Mapping Geographical Variation in Implicit Racial Attitudes," *Project Implicit*, https://implicit.harvard.edu/implicit/user/jaxt/blogposts/ piblogpost005.html (accessed 1 June 2022).

While our data will facilitate the ability to identify fights, on other margins we will be limited. There is a degree of difficulty in capturing race categorically, generally, and coarse racial categories prevent us from distinguishing between students who report more than one race. For example, while it is easy to imagine that students who identify as both Black and white experience different disciplinary outcomes than students who identify as both Asian and white, the data record both types as "two or more races." Similarly, the data do not allow us to distinguish between race and ethnicity, as students who report Hispanic ancestry are coded as "Hispanic," regardless of their race. As a result, the available racial categories can complicate the interpretation of specific racial gaps, as students perceived by administrators as one race (e.g., Black) may be coded in the data as another (e.g., Hispanic, or as two or more races). Moreover, the considerable racial diversity in the sample can limit our ability to precisely estimate specific gaps, such as those between monoracial Black and white students. Thus, to economize on statistical power, we conduct the analysis around incidents that involve only white students, incidents that involve only students of color, and those that involve both white students and students of color, defining "students of color" as those who do not identify as white non-Hispanic.[6] That said, the qualitative conclusions from a more granular analysis of specific racial disparities (i.e., Black-white, Hispanic-white, and gaps between the remaining students of color and white students) are unchanged, and are similar to those documented in Section 1.3 and Section 1.4, but with less precision.

1.2.1 Sample selection

We restrict our attention to incidents that (i) are well defined, (ii) are more likely to have well-defined sets of participants, (iii) are narrow enough in scope that we can

---

[6] Specifically, we define students of color as those who identify as (i) solely Black, (ii) Hispanic (of any race), (iii) solely Asian, (iv) solely Pacific Islander, (v) solely Native American, or (vi) two or more races.

argue that any remaining racial disparities are not likely to be explained by incident heterogeneity, and (iv) are not so rare that they lack economic significance. A set of incidents that satisfies these criteria provides as close to as-good-as-random variation as possible while still allowing us to contextualize multi-race incidents with a set of similar, but same-race incidents.

To satisfy these criteria, we consider infractions for "fighting without major injury" among boys.[7] In addition to being included in mandatory federal reporting, the fights in our sample are well-defined by the state. State guidance defines "fighting without major injury" as "mutual participation in an incident involving physical violence" and specifically conditions on incidents in which no "persons on school grounds require professional medical attention" (Reykdal, Weaver-Randall, & Ireland, 2018). The state also provides examples of disqualifying injuries; fights that result in "stab or bullet wounds, concussions, fractured or broken bones, or cuts requiring stitches" would be adjudicated in a separate category of offense. Thus, if fights between students of color tend to be worse in some unobservable way that rationalizes harsher penalties, "worse" must not be so much worse as to imply "cuts requiring stitches." In that way, "worse" has an upper bound of "not so much worse that there are stitches." Moreover, the state directs school officials to exclude "verbal confrontations, tussles, or other minor confrontations." Collectively, these describe a fairly narrow band of student activity over which we can examine differences in adjudication outcomes between students of color and white students.

Relative to other forms of joint misconduct, it can also be argued that fights are the least likely to originate from race-based selection into the sample. Consider "disruptive conduct," for example, which the state defines as any behavior "that

---

[7] Relative to the boys in our data, girls are rarely implicated for fighting. Boys' infractions for fighting outnumber those of girls by a ratio of four to one.

materially and substantially interferes with the educational process" (Reykdal et al., 2018). The relative subjectivity permitted in determining what constitutes disruptive conduct would leave much more room for race-based selection into infractions. In contrast, well-defined conditions and mandatory reporting supports that selection into fights is less likely to depend on the subjective judgments of teachers, so our focus on fights tips toward limiting potential measurement error in the classification of incidents. To the extent that there are concerns about selection into fights, those concerns should be heightened considerably in the analysis of other types of incidents. As for framing the external validity of comparisons across fights, we note that fights are the most common type of multi-student incident in our data, and while there will surely be some students who escape the eyes of teachers, the "jointness" of fights leaves us more confident that we have captured the set of relevant actors.[8]

There are a total of 66,355 fighting infractions among boys in our data. While schools are required to use the same incident identifier for incidents that involve multiple students, 33 percent of fighting infractions occur in schools that never report the same incident identifier for multiple students. Thus, our analysis will speak only to schools that follow the reporting guidelines.[9] In schools that do report matching incident identifiers, not all fighting infractions have an incident identifier that matches that of another student in the infraction data. As a worst case, one might imagine that white students systematically avoid fighting infractions, leaving an "excess" of

---

[8] Further, note that no state reports data on victims, to our knowledge, and to the extent the victim is observable to those adjudicating student conduct (but not to the econometrician), there may also be missing race components to the adjudication of other categories of misconduct. Considering fights between students—fights being well-defined and subject to mandatory reporting—mitigates such concerns.

[9] Schools that follow the reporting guidelines tend to be less white, more urban, and more economically disadvantaged (as measured by the fraction of students who qualify for free or reduced-price meals) than schools that do not.

students of color among the reported fights. If it is the less severe white infractions that select out of reporting, then the measurable within-incident race differentials would understate the extent of discriminatory adjudication in our identifying sample. While this possibility is not unique to our setting, the safest inference going forward might be to interpret our estimated differentials as lower bounds of the effect of race on outcomes. In total, we observe 16,279 infractions from 7,641 multi-student incidents that implicate at least two boys for fighting. To further ensure the comparability of the fights in our sample, we discard 576 infractions from multi-student incidents that include girls or that implicate other students for non-fighting behaviors, though our findings are not sensitive to these restrictions.

1.2.2 Outcomes

We consider three margins of formal exclusionary discipline as outcomes for each infraction: whether the student is expelled, whether the student receives any suspension or expulsion, and the length of suspension conditional on all students being suspended within an incident.[10] In Table 1 we provide average disciplinary outcomes by school level (i.e., elementary, middle, high) for (i) all fighting infractions in Washington, (ii) all fighting infractions at schools that use the same incident identifier (across students) when multiple students are involved in individual fights, and (iii) all multi-student fights at these schools. Fewer than one percent of fighting infractions in high school result in expulsion (the most severe punishment we observe in the data) and expulsions for fighting are rarer still in elementary and middle schools—too rare to build reasonable inference from. On other margins, punishments vary significantly across school levels—the rate of formal exclusionary discipline (suspension or expulsion)

---

[10] We exclude suspension lengths longer than 20 school days (approximately one calendar month) to limit the influence of rare long-term suspensions (over 99 percent of suspensions are shorter than 20 days). However, the inferences we make are not sensitive to perturbations of this 20-day cutoff.

Table 1. Summary statistics

	Grades PK–5		Grades 6–8		Grades 9–12	
	μ	σ	μ	σ	μ	σ
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All fighting infractions						
Expelled? (= 1 if yes, = 0 if no)	0.00	0.01	0.00	0.04	0.01	0.08
Observations	29,523		24,999		11,832	
Incidents	27,025		20,979		9,799	
Suspended or expelled? (= 1 if yes, = 0 if no)	0.40	0.49	0.83	0.37	0.91	0.28
Observations	29,523		24,999		11,832	
Incidents	27,025		20,979		9,799	
Suspension length (days)	1.45	1.27	2.26	1.72	3.48	2.39
Observations	11,424		20,264		10,325	
Incidents	10,430		16,839		8,562	
Panel B: All fighting infractions from schools that report at least one multi-student fight						
Expelled? (= 1 if yes, = 0 if no)	0.00	0.00	0.00	0.04	0.01	0.09
Observations	12,018		14,599		7,026	
Incidents	9,533		10,582		5,002	
Suspended or expelled? (= 1 if yes, = 0 if no)	0.40	0.49	0.86	0.35	0.92	0.28
Observations	12,018		14,599		7,026	
Incidents	9,533		10,582		5,002	
Suspension length (days)	1.54	1.39	2.30	1.74	3.54	2.41
Observations	4,572		12,132		6,125	
Incidents	3,588		8,709		4,367	
Panel C: Multi-student fights						
Expelled? (= 1 if yes, = 0 if no)	0.00	0.00	0.00	0.04	0.01	0.09
Observations	4,556		7,587		3,560	
Incidents	2,141		3,703		1,666	
Suspended or expelled? (= 1 if yes, = 0 if no)	0.44	0.50	0.89	0.31	0.94	0.24
Observations	4,556		7,587		3,560	
Incidents	2,141		3,703		1,666	
Suspension length (days)	1.35	0.83	2.16	1.51	3.47	2.22
Observations	1,868		6,495		3,159	
Incidents	905		3,181		1,501	

Notes: Sample means (μ) and standard deviations (σ) of punishment outcomes considered in Section 1.3 and Section 1.4. The alternative to an expulsion or a suspension is either "no intervention" or "other intervention." Suspension lengths are conditional on all students being suspended within each incident. The sample in Panel A consists of boys' infractions for "fighting without major injury" between 2014–15 and 2018–18. The sample in Panel B consists of boys' fighting infractions from schools that report fights with matching incident identifiers. The sample in Panel C consists of fighting infractions from multi-student fights that implicate at least two boys for fighting, but do not include girls or implicate other students for non-fighting behaviors. A fight is classified as "multi-student" if two or more students have a matching incident identifier.

doubles between elementary and middle school, and the average suspension is over one day longer in high schools than in middle schools.[11] Within each grade span, average disciplinary outcomes are similar across samples, though exclusionary discipline rates are somewhat higher in the multi-student sample than in the larger samples, and suspensions are somewhat shorter.

1.2.3 Student characteristics

We observe a total of 41,520 students in the full sample and 12,855 students in the multi-student sample. Roughly 41 percent of students in the multi-student sample are white (non-Hispanic), 27 percent are Latino (Hispanic origin of any race), 16 percent are Black, 9 percent report more than one race, 3 percent are Asian, 2 percent are Pacific Islander, and 2 percent are Native American.

We derive controls for socioeconomic status, disability, and past achievement from an extended panel that dates back to 2009–10. We measure socioeconomic status using persistent eligibility for free or reduced-price meals. While there are significant racial disparities in socioeconomic status within each grade level, the vast majority of infractions in our sample implicate students from low-income households—this is true for white students and students of color alike. Using up to nine years of data, we determine whether a student is (i) always eligible, (ii) never eligible, or (iii) sometimes eligible for free or reduced-price meals. In doing so, we follow others who have argued that persistent eligibility provides a better proxy for current household income than current eligibility (Michelmore & Dynarski, 2017). To measure special-education status, which is an important predictor of punishment, we derive two proxies from state testing data. The first indicates whether a student has previously taken a state test that is

---

[11] While schools are not required to report infractions that do not result in suspension or expulsion, 95 percent of fighting infractions are from schools that report infractions (for fighting or other behaviors) that result in "no intervention" or "other intervention."

intended to be taken by students with disabilities and the second indicates whether a student has previously taken an alternative state test that is intended to be taken by students with an individualized education program. We control for observed English Language Arts (ELA) and math achievement levels from the most-recent grade tested. While there are significant racial disparities in achievement within each grade level, the plurality of infractions in our sample are from low-achieving students—this is true for white students and students of color. As a general rule, our objective in modeling punishment outcomes is not to control for ability, but rather to control for what an administrator observes (and may consider) when adjudicating misconduct.[12]

Using an additional year of infraction data, we also control for each student's infraction history, measured as the number of infractions from the previous school year. Students who select into fights typically have an infraction from the previous school year, and students of color tend to have more past infractions than their white peers.[13]

1.3 Within-incident comparisons

Given the potential for differential selection into incidents (by students, for example) and the potential for differential adjudication of incidents (by different vice principals), the difference in the average punishment received by white students and by students of color is not likely capturing the causal relationship of interest (i.e., the change in punishment induced by an all-else-equal change in the perception of student race by school officials). For example, if baseline differences in misconduct or punishment vary across schools and there are more students of color in schools

---

[12] For this reason, we include students with test scores that are unobservable to both the econometrician and school administrators in our analysis (e.g., elementary students who have not yet been tested, as tests are not available until the third grade). We allow for any level differences for those without test scores with an indicator variable, though results are robust to their exclusion.

[13] The results we report in Section 1.3 and Section 1.4 are not sensitive to controlling specifically for the number of fighting infractions from the previous school year.

with higher baseline levels of misconduct or higher average punishments, then it may well look like students of color are treated more harshly without there ever being any individual actor (e.g., a vice principal) treating students of color differently. Such differences in outcomes are important, but the policy implications can be quite different if no individual actors are implicated as part of the mechanism that produces differential outcomes.

Below, we consider expulsion and the extensive and intensive margins of suspension, and provide estimates of the gap in outcomes for students of color across several specifications. In the end, we will approach a within-incident comparison where one may be more inclined to interpret estimates as causal. We will then re-direct our efforts toward identifying a mechanism that can explain the advent of race-based differentials in punishment.

1.3.1 Expulsions

In Figure 1 we begin by reporting unconditioned differences in the adjudication of student misconduct, and then progressively restrict the identifying variation that contributes to the estimated difference. The left-most estimate in Panel A is the raw difference in expulsion rates between white students and students of color in high school. Conditional on receiving an infraction, expulsion rates for students of color are 0.66 percentage points, or 213.2 percent, higher ($p < 0.001$) than those for white students. Relative to the sample standard deviation (σ) of expulsion in the estimation sample, this difference corresponds to an effect size on the order of 0.08σ.

In Column (2) we control for student attributes (e.g., grade, eligibility for free or reduced-price school meals, past achievement, and proxies for the receipt of special education services), past infractions, and school-by-year fixed effects, absorbing any variation in punishment across schools into the error term for the sample of all fighting

26

Figure 1. Punishment disparities in school fights

Coefficient: ○ Student of color

Grade span: ● PK-5　● 6-8　● 9-12

Panel A: Expelled? (= 1 if yes, = 0 if no)

Panel B: Suspended or expelled? (= 1 if yes, = 0 if no)
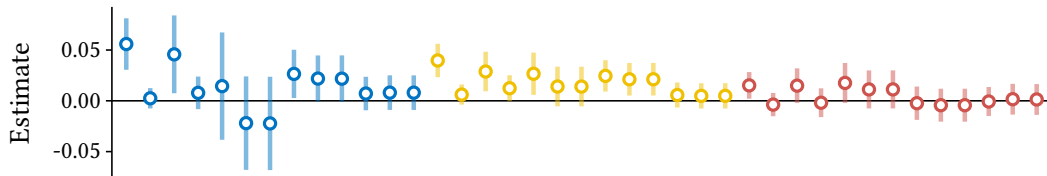
Panel C: Suspension length (days, conditional on all being suspended)
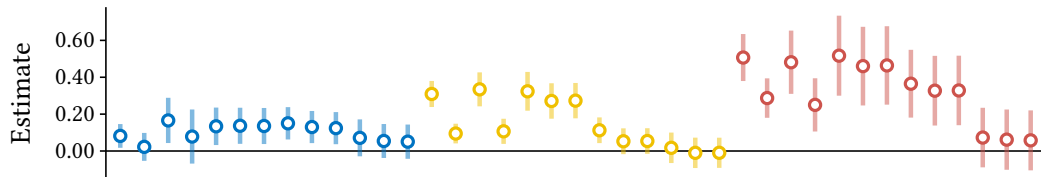
Notes: Open circles show OLS estimates of racial punishment gaps. Each estimate is from a different regression. The leftmost estimate in each grade span describes a raw punishment gap, and the rightmost estimate describes a within-incident punishment gap from the fully specified model (e.g., see Equation 1.1). The unit of observation is an infraction for "fighting without major injury." The reference category consists of white students' infractions. Solid circles below each set of estimates describe the attributes of each regression: an opaque circle indicates the presence of an attribute and a translucent circle indicates the absence of an attribute. Vertical lines outline 95% confidence intervals adjusted for clustering at the school-by-year level.

[a] All fighting infractions from schools that report at least one multi-student fight.

infractions. In contrast to Kinsler (2011), who estimates a similar specification for suspensions using data from North Carolina, this fails to decrease the variation in expulsion that is attributable to race, and within-school variation in expulsions are still suggestive of significant gaps in the adjudication of infractions for students of color compared to white students. The introduction of school-by-year fixed effects does attenuate race differentials for suspension outcomes, considered further below, but significant gaps remain nonetheless.

In Column (3) we consider the unconditioned race differential for fighting infractions from schools that report fighting infractions with matching identifiers—it is within these schools that we will have the ability to restrict identifying variation to within-incident variation. As in the full sample, the unconditioned gap in these schools implies that students of color are significantly more likely than white students to be expelled for fighting. This difference is larger in magnitude than the point estimate in Column (1), though the confidence intervals do overlap. Likewise, the addition of controls and school-by-year fixed effects in Column (4) has little impact on the magnitude of the estimated race differential.

In columns (5) through (7) we restrict the sample to fighting incidents that explicitly implicate more than one student. For completeness, we again produce estimates of the unconditioned differences and thereafter collapse toward our preferred specification. In columns (6) and (7), for example, we control first for student attributes and then also for past infractions.

We begin to approach something that may defensibly justify a causal interpretation in the Column (8) of Figure 1, where we control for school heterogeneity with the inclusion of school-by-year fixed effects. However, it is in columns (11) through (13) that we absorb any unobserved heterogeneity that is specific to incidents—

this is where we are most confident in having retrieved estimates that warrant a causal interpretation. We execute these within-incident comparisons by estimating models of the form

$$\mathbb{1}(\text{Expelled} = 1)_{ikst} = \beta \, \text{SoC}_i + X'_{ij}\Theta + \lambda_k + v_{ikst} \, , \qquad (1.1)$$

where $\mathbb{1}(\text{Expelled} = 1)_{ikst}$ captures the expulsion of student i associated with their involvement in incident k in school s during year t.[14] Incident fixed effects (λ_k) capture unobserved heterogeneity in incidents (nested within schools). Student controls (X'_i) adjust for level differences that arise from within-incident variation in student attributes (i.e., grade, eligibility for free or reduced-price school meals, math and reading achievement levels from the previous school year, and proxies for the receipt of special education services) and past infractions (from the previous school year). Our parameter of interest (β) absorbs the average difference in expulsion rates for students of color ($\text{SoC}_i = 1$) relative to white students ($\text{SoC}_i = 0$). The error term (v_{ikst}) captures any remaining variation, and we allow for clustering at the school-by-year level.

If students of color are systematically more culpable (e.g., more contributory, or associated systematically with actions that are deemed more severe, or more worthy of punishment), then it would not be surprising to observe punishment differentials that disfavor students of color. This constitutes the assumption that implies a causal interpretation of $\hat{\beta}$—we assume that students of color are not differentially culpable, conditional on the full set of controls and incident fixed effects. If selection into misconduct has school officials being less lenient toward students of color, then estimates of racial gaps in punishment could understate the extent of discriminatory adjudication. That said, across the fighting infractions we consider, we are less concerned that differential selection into incidents explains our results—the severity

---

[14] No student (i) has multiple infractions (j) within the same incident (k).

of these behaviors presents teachers with few opportunities to exercise discretion in deciding whether to refer students to the principal's office for discipline.

We find no statistically significant difference in the probability of expulsion for students of color relative to white students in the same incident. In the preferred specification, the probability of expulsion is 0.48 percentage points higher (227%, 0.08σ), on average, for students of color, but the difference is indistinguishable from zero at conventional significance levels ($p = 0.339$). Though the difference in probability is statistically insignificant, note that we cannot rule out meaningful effect sizes at the upper bound of the 95-percent confidence interval.

1.3.2   Suspensions

In panels B and C of Figure 1 we repeat a similar exercise for suspensions— in Panel B we model the extensive margin, and in Panel C we estimate models of suspension length conditional on suspension.[15] As expected, suspensions vary systematically with race—unconditioned, students of color experience rates of formal exclusionary discipline that are 5.59 percentage points higher (15.1%, 0.11σ, $p < 0.001$) in elementary schools, 3.97 percentage points higher (4.9%, 0.11σ, $p < 0.001$) in middle schools, and 1.52 percentage points higher (1.7%, 0.05σ, $p = 0.023$) in high schools. Racial disparities are also large on the intensive margin of suspension—conditional on being suspended for fighting, students of color receive suspensions that are 0.08 days longer (5.8%, 0.06σ, $p = 0.013$) in elementary schools, 0.31 days longer (14.8%, 0.18σ, $p < 0.001$) in middle schools, and 0.51 days longer (15.9%, 0.21σ, $p < 0.001$) in high schools. However, when we restrict the identifying variation to that existing within

---

[15] For the analysis in Panel B we model "suspended or expelled" together as there are two margins around which we anticipate student selection. Namely, there are students who are at the margin of being suspended, and (likely different) students who are at the margin of being expelled—defined this way, exit is necessarily toward lesser consequences. For the analysis in Panel C we restrict the sample to incidents that result in suspensions for all students.

incidents we find that students of color are neither suspended at higher rates than white students implicated in the same incident, nor suspended for any longer, conditional on being suspended. Within-incident variation in student race does not support the claim that there are significant differences in suspensions experienced by students of color, *on average.*

In fact, after accounting for unobserved incident-specific heterogeneity, no margin of punishment supports that there are statistically significant differences in the disciplinary actions imposed on students of color—this is true across elementary, middle, and high schools. Although some estimates have relatively wide confidence intervals—namely those of expulsion and suspension length gaps in high school—other are precise zeros, giving us an additional degree of confidence that the adjudication of the infractions of students of color is not systematically different from that of white students implicated in the same fight. If white students and students of color are equally culpable, on average, for their involvement in a fight, then differential treatment within incidents explains very little of the aggregate racial disparities.

1.4   Where do gaps arise, then?

One explanation for the absence of significant gaps in punishment within multi-race fights is that within-incident variation in race offers a degree of salience that enables the equal treatment of students of color. For example, it could be easier for administrators to suppress implicit biases within incidents. Alternatively, it could be that explicit biases are more costly to act on within incidents, where one cannot appeal to incident heterogeneity (e.g., "It was a really bad fight") as a justification for harsher punishment.

That punishment gaps attenuate when we identify off of within-incident variation is also consistent with race-based differences in parents' inclinations to

advocate for their children, or for their advocacy to exert varying degrees of influence on punishments. For example, if advocacy varies more across race than within, we should expect advocacy-driven variation in outcomes to be partially absorbed by the incident fixed effect. However, for a differential-advocacy story to explain the variation we see in the data, it would need to be the case that administrators respond to the advocacy given to a white student *and* extend it to others involved in the same fight, regardless of race. In that way, administrators still appear better able to maintain equality norms within fights than they do across fights.

By absorbing the unobserved heterogeneity associated with specific incidents, the foregoing analysis identifies only those factors that vary within incidents—we necessarily lose the context that would come from the comparison of multi-race fights alongside same-race fights, where some of the mechanisms that induce equal treatment are absent. In Figure 2 we therefore consider the punishments of students of color across multi-student fights—dropping the incident fixed effects from the earlier analysis allows for the comparison of multi-race and same-race fights.[16] Specifically, we estimate models of the form

$$\text{Punishment}_{ikst} = \beta \, \text{SoC}_i + \tau \, \text{Multiracial}_k + \phi \, \text{SoC}_i \times \text{Multiracial}_k + X_i' \Theta + \lambda_{st} + v_{ikst} \, ,$$

$$(1.2)$$

where Punishment_{ikst} is the disciplinary intervention assigned to student i associated with their involvement in incident k in school s during year t. As before, we control for student attributes and past infractions, and identify racial gaps across same-race fights (β) with the conditional variation that exists within the same school during the same school year. As selection into multi-race fights may differ, we absorb any level effect

---

[16] Multi-race fights make up 34.8 percent of multi-student fights, while 42.5 percent implicate only students of color—the remaining 22.7 percent implicate only white students.

associated with multi-race fights in τ.[17] However, our interest is in how that treatment changes for students of color, across the changing racial composition of fights (ϕ). In a way, we are asking whether being implicated with a white student induces changes in the punishments assigned to students of color.

A causal interpretation of the within-incident differences in Figure 1 required the assumption that students of color are not differentially culpable, conditional on controls and incident fixed effects. In Figure 2, however, to interpret $\hat{\phi}$ as causal we must also be willing to assume that there is no differential selection into multi-race fights—it cannot be that it is the less-culpable students of color who are selecting into fights with white students. In other words, to explain away the variation we observe in the data (conditioning on school-by-year fixed effects, student characteristics, and past infractions) one must simultaneously believe that (i) students of color who select into fights with only other students of color are somehow more deserving of punishment than white students who select into fights with only other white students and (ii) students of color who select into fights with white students are somehow less deserving of punishment than students of color who select into fights with only other students of color.

1.4.1 Results

In Figure 2 we plot two coefficient estimates from each model. The first is the estimated difference in outcomes for students of color, identified off of same-race fights (i.e., the average difference in outcomes across fights that implicated only students of color and fights that implicated only white students). The second is the estimated difference in outcomes for students of color in multi-race fights (i.e., the average

---

[17] Point estimates of τ are generally positive, capturing that multi-race fights tend to be punished more heavily than same-race fights. In high school expulsions and suspension length, and in elementary school suspension length, $\hat{\tau}$ is small but statistically significant, though in all models $\hat{\tau}$ is smaller in magnitude than both $\hat{\beta}$ and $\hat{\phi}$.

difference in outcomes experienced by students of color in fights that implicated a white student).

In Panel A we consider expulsions among high school students. Our preferred specifications identify that (i) students of color *in fights that only implicate students of color* experience significantly higher rates of expulsion (2.24 percentage points, 0.27σ, $p = 0.006$) and (ii) the increase in expulsion rates for students of color is offset when there is a white student implicated in the same fight (-1.75 percentage points, -0.21σ, $p = 0.022$). The sum of those coefficients—that is, the marginal effect of being a student of color in a multi-race fight—is indistinguishable from zero (0.49 percentage points, 0.06σ, $p = 0.324$), which is consistent with the presence of a white student *fully* offsetting the average difference in expulsion rates. The expulsion gap persists when we restrict the sample to fights that involve exactly two students (1.26 percentage points, 0.2σ, $p = 0.041$), as does the offsetting difference for being implicated with a white student (-1.38 percentage points, -0.22σ, $p = 0.047$). As in the multi-student sample, the sum of the coefficients indicates that students of color are no more likely to be expelled when they are implicated with white students (-0.12 percentage points, -0.02σ, $p = 0.798$). In either sample, "impact" estimates—the percentage changes over the mean of the reference group—are undefined, as no white student in an all-white incident is expelled for fighting.

In Panel B of Figure 2 we consider suspension disparities for each grade level. Again, we find that students of color are only more likely to experience exclusionary punishment when they are implicated with only other students of color. This is most evident in middle schools, where students of color are 4.6 percentage points more likely to receive exclusionary punishment when they are implicated with only students of color (5.3%, 0.2σ, $p = 0.003$), but no more likely when they are implicated with white

Figure 2. Punishment disparities across school fights by racial composition

Coefficient: ○ Student of color ▫ Student of color × multiracial fight

Grade span: ● PK-5   ● 6-8   ● 9-12

Panel A: Expelled? (= 1 if yes, = 0 if no)

Panel B: Suspended or expelled? (= 1 if yes, = 0 if no)
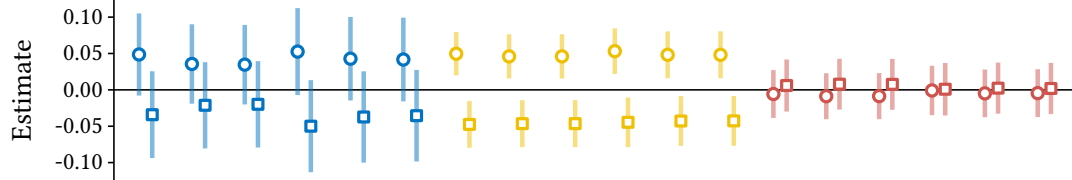
Panel C: Suspension length (days, conditional on all being suspended)
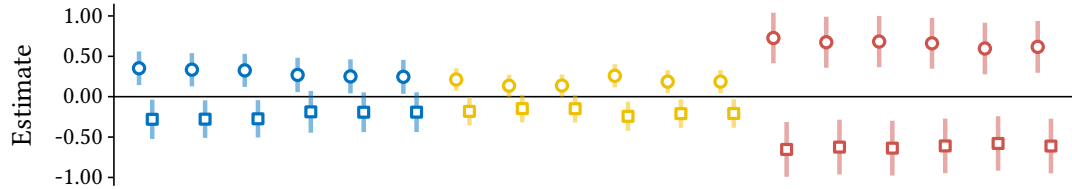
Notes: Open circles and squares show OLS estimates of coefficients from Equation 1.2. Each set of two estimates is from a different regression. The unit of observation is an infraction, and the sample consists of infractions from multi-student fights in which all students receive infractions for "fighting without major injury." The reference category consists of white students' infractions from all-white fights. Solid circles below each set of estimates describe the attributes of each regression: an opaque circle indicates the presence of an attribute and a translucent circle indicates the absence of an attribute. Vertical lines outline 95% confidence intervals adjusted for potential clustering at the school-by-year level.

students (-0.04 percentage points, 0%, 0σ, $p = 0.96$). The same pattern is also evident in elementary schools, where students of color experience a larger gap when they are implicated with only students of color (3.47 percentage points, 7.7%, 0.12σ, $p = 0.214$) than when they are implicated with white students (1.48 percentage points, 3.3%, 0.05σ, $p = 0.128$), though precision falls short of conventional significance levels. In high schools, however, we do not observe significant differences for students of color when they are implicated with only students of color (-0.86 percentage points, -0.9%, -0.05σ, $p = 0.593$) or when they are implicated with white students (-0.1 percentage points, -0.1%, -0.01σ, $p = 0.894$). Across all grade levels, inferences are robust to restricting the sample to fights that involve exactly two students.

In Panel C we consider the intensive margin of suspension, where we find similar patterns, and with enough precision to suggest that the patterns we identify are a significant part of the data-generating process. Relative to white students in same-race fights, suspensions for students of color in fights that implicate only other students of color are, on average, 0.33 days longer in elementary school (26.9%, 0.57σ, $p = 0.002$), 0.14 days longer in middle school (7.3%, 0.11σ, $p = 0.048$), and 0.68 days longer in high school (22.6%, 0.41σ, $p < 0.001$). However, when implicated with white students, students of color receive suspensions that are no longer than those for white students in all-white fights—this is true in elementary school (0.06 days, 4.3%, 0.09σ, $p = 0.281$), in middle school (-0.01 days, -0.6%, -0.01σ, $p = 0.802$), and in high school (0.04 days, 1.6%, 0.03σ, $p = 0.607$). As in Panel B, this pattern is robust to restricting the sample to fights that involve exactly two students.

The offsetting differences in Figure 2 suggest that the within-school disparities in Figure 1 are driven by differences in punishment *across* same-race fights. When implicated in fights with at least one white student, students of color are punished

no differently, on average, than white students implicated for fighting in the same school during the same school year. When implicated in fights without white students, however, students of color receive systematically harsher punishments than those imposed on white students.[18]

Depending on the grade level and margin of punishment, the magnitude of the difference in punishment across same-race fights is often large—in several cases the point estimate is nearly identical to the unconditioned race gap. Indeed, some of the race differentials that we estimate across same-race fights—especially those on the intensive margin of suspension—are larger in magnitude than within-school gaps documented elsewhere in the literature (e.g., Anderson & Ritter, 2020; Barrett et al., 2021; Kinsler, 2011; Shi & Zhu, 2022).

If students of color and white students in same-race fights are equally culpable (after conditioning on school-by-year fixed effects and the full set of controls), then the empirical regularities we document are consistent with disparate treatment of all-white fights and fights involving only students of color. The full characterization of the data-generating process—with within-incident variation coming from multi-race fights—strongly suggests that the presence of a white student moves administrators toward equal treatment, consistent with administrators correcting biases when racial differences are more salient.

1.5   Conclusion

Racial disparities in the incidence of exclusionary discipline have increased since race-based gaps in suspensions were first documented (Children's Defense Fund, 1975;

---

[18] Recall that the vast majority of infractions in our sample are from low-income students—this is true for both students of color and white students. The relative dearth of students from higher-socioeconomic-status backgrounds increases our confidence that the patterns we document in Figure 2 reflect the salience of race rather than the salience of socioeconomic status. Estimates of β and ϕ are qualitatively similar when we stratify each sample by socioeconomic status and run them separately, though we lose precision in the sub-samples with higher socioeconomic status.

Losen, Hodson, Keith II, Morrison, & Belway, 2015). In an effort to reduce discipline gaps, policymakers have begun to roll back strict "zero tolerance" discipline policies that have been shown to have a disparate impact on students of color (Curran, 2016). For example, some school districts have implemented policies that mandate the elimination of exclusionary interventions for low-level offenses (Craig & Martin, 2019; Lacoe & Steinberg, 2018; Pope & Zuo, 2020; Steinberg & Lacoe, 2018), while others have experimented with less punitive disciplinary interventions, such as restorative justice (Glenn, Barrett, & Lightfoot, 2020). However, the ultimate success of any disciplinary reform depends, in part, on the ability of school officials to enforce policies without partiality. With evidence that educators hold biases that disfavor students of color (Chin, Quinn, Dhaliwal, & Lovison, 2020), the extent to which those biases manifest in disparate treatment can have important implications for the effectiveness of education reforms, including those concerning the use of discipline.

In our analysis, we inquire into the source of racial disparities in the adjudication of student misconduct. In incidents that implicate at least two boys for fighting, we find that students of color receive harsher punishments, on average. Across same-race fights in the same school during the same school year, students of color are more likely to be suspended or expelled than white students, and tend to receive longer suspensions conditional on being suspended. However, within multi-race fights, the punishments imposed on students of color are statistically indistinguishable from those imposed on white students. Moreover, we document a pattern, evident across grade levels and robust to a variety of alternative specifications, in which the presence of a white student *fully* offsets within-school punishment differentials for students of color.

We find encouragement insofar as the data-generating process supports that biases are correctable where race and equality norms are more salient. That being said,

our results imply that purging all within-incident disparities in punishment would do little to close the gap in disciplinary outcomes between students of color and their white peers. Our results also raise questions about the prospects of school accountability measures that leverage incident identifiers to detect discrimination—relying on within-incident comparisons to monitor unequal treatment would falsely signal an equality in outcomes, understating the extent of differential treatment.

CHAPTER 2

VOTING RIGHTS AND THE RESILIENCE OF BLACK TURNOUT

From Raze, K. (2022). Voting rights and the resilience of Black turnout. *Economic Inquiry*, (60)3, 1127–1141.

2.1 Introduction

For its profound impact on the political participation and representation of Black Americans, the Voting Rights Act of 1965 is often described as "the most successful piece of civil rights legislation ever adopted" by Congress.[1] By eliminating poll taxes, literacy tests, and other discriminatory election policies, the Voting Rights Act produced lasting increases in Black turnout (Ang, 2019; Filer, Kenny, & Morton, 1991; Fresh, 2018), which then increased the number of Black local elected officials and elevated support for subsequent civil rights legislation (Bernini, Facchini, & Testa, 2018; Schuit & Rogowski, 2017). Ultimately, the expansion of voting rights generated significant economic benefits for Black communities (Avenancio-León & Aneja, 2019; Cascio & Washington, 2014; Facchini, Knight, & Testa, 2020).

Most of the political and economic consequences of the Voting Rights Act are attributable to an enforcement mechanism known as preclearance. Preclearance required state and local jurisdictions identified by a coverage formula to secure federal approval for any election-related policy change.[2] The formula identified "covered" jurisdictions as those with histories of literacy tests or significant racial or linguistic disparities in voting. To secure federal approval, a covered jurisdiction had to demonstrate that the policy change would not restrict voting rights based on race or membership in a language minority group. If federal authorities withheld approval,

---

[1] US Department of Justice. "Introduction to Federal Voting Rights Laws." 6 August 2015. https://www.justice.gov/crt/introduction-federal-voting-rights-laws-1 (accessed 29 December 2020).

[2] *Allen v. State Board of Elections*, 393 U.S. 544 (1969).

then the jurisdiction could not legally implement the proposed change. Between 1965 and 2013, federal authorities received 556,268 proposals (over 11,000 per year) for election-related policy changes in covered jurisdictions.[3] At issue, then, is that the coverage formula was invalidated by the Supreme Court of the United States, in *Shelby County v. Holder* (2013). In this way, what is known as "the *Shelby* decision" ended 48 years of direct preclearance oversight. Previously covered jurisdictions can now adopt new election policies *without* federal approval.

The removal of preclearance could enable officials in previously covered jurisdictions to suppress voting for political advantage. While white voters in covered jurisdictions lean Republican, Black voters exhibit strong and enduring support for Democrats (Kuriwaki, 2020). Any successful attempt at suppressing the Black vote would likely benefit Republican candidates. Political conditions at the time of the ruling were favorable to selectively manipulating the costs of voting—Republicans held the governorship and both houses of the legislature in 12 out of 15 states with covered jurisdictions.[4]

The *Shelby* decision appears to have prompted previously covered jurisdictions to adopt policies that increase the cost of voting, perhaps differentially for prospective Black voters. Hours after the ruling, the Attorney General of Texas announced that "the State's voter ID law will take effect immediately".[5] State election officials in Alabama and Mississippi implemented strict voter identification laws that had either failed to

---

[3] US Department of Justice. "Section 5 Changes by Type and Year." 6 August 2015. https://www.justice.gov/crt/section-5-changes-type-and-year-2 (accessed 29 December 2020).

[4] Ballotpedia. "State government trifectas, pre-2014." https://ballotpedia.org/State_government_trifectas\#State_government_trifectas.2C_pre-2014 (accessed 25 November 2020).

[5] Brennan Center for Justice. "The Effects of *Shelby County v. Holder*." 6 August 2018. https://www.brennancenter.org/our-work/policy-solutions/effects-shelby-county-v-holder (accessed 4 February 2021).

secure federal approval before *Shelby* or were under review at the time of the ruling.[6] Other covered states, such as North Carolina, also enacted voter identification laws with additional restrictions on early voting and registration (US Commission on Civil Rights, 2018). A federal court later overtuned North Carolina's election policy changes on the grounds that they "target African Americans with almost surgical precision."[7]

Not all election policy changes in covered jurisdictions were blocked by the courts after *Shelby*. Covered states were more likely than uncovered states to purge voters of color from voter rolls (Brater, Morris, Pérez, & Deluzio, 2018; Feder & Miller, 2020) and covered counties were more likely than uncovered counties to close polling places, though not in North Carolina (Shepherd, Fresh, Eubank, & Clinton, 2021).[8] The proliferation of election-related policy changes in covered jurisdictions suggests that preclearance oversight was binding.

Given the response of covered jurisdictions after the ruling, fears that *Shelby* could attenuate or even reverse the effects of the Voting Rights Act are well-founded. However, the worst-case scenario presupposes a change in the composition of the electorate, for which there is little existing evidence. I bring new evidence to bear on the impact of the *Shelby* decision by considering whether the removal of preclearance decreased turnout among eligible Black voters relative to white voters in previously covered states.

The impact of preclearance removal may depend on the relative importance of (i) increases in the cost of voting and (ii) counter-mobilization against new voting

---

[6] Brennan Center for Justice. "New Voting Restrictions in America." 19 November 2019. https://www.brennancenter.org/new-voting-restrictions-america (accessed 25 November 2020).

[7] *North Carolina State Conference of the NAACP v. McCrory*, 831 F.3d 204 (2016).

[8] Rob Arthur and Allison McCann. "How the Gutting of the Voting Rights Act Led to Hundreds of Closed Polls." *Vice News*, 16 October 2018. https://news.vice.com/en_us/article/kz58qx/how-the-gutting-of-the-voting-rights-act-led-to-closed-polls.

restrictions. Although increases in the cost of voting exert downward pressure on turnout, counter-mobilization can stimulate turnout (Biggers & Smith, 2020; Valentino & Neuner, 2017), making the net effect of policy changes made possible by the *Shelby* decision theoretically ambiguous. Understanding the impact of the *Shelby* decision therefore requires empirical evaluation.

I conduct such an evaluation using validated turnout and registration data from six biennial waves of the Cooperative Congressional Election Study (2008-2018) and state-level variation in preclearance coverage. To identify the impact of preclearance removal, I employ a triple-difference design that uses states that were unaffected by *Shelby* as a control group for covered states. Building upon difference-in-differences variation in preclearance oversight, I incorporate a third difference by comparing the turnout of Black and white voters within each state. This allows me to flexibly control for differential trends in aggregate turnout and registration rates that could otherwise confound estimates of the effect of *Shelby* on the relative turnout of Black voters. The resulting triple-difference comparison considers, in effect, how changes in Black-white voter participation gaps differ across covered and uncovered states after *Shelby*.

I find that the removal of preclearance requirements did not significantly reduce the relative turnout or registration of eligible Black voters. If anything, Black turnout may have *increased* relative to white turnout in covered jurisdictions after *Shelby*. I document that *Shelby* increased relative turnout in covered states during the 2016 election, with the largest effects concentrated among covered states where all levels of government were previously subject to preclearance. Midterm elections exhibit smaller effects that are statistically indistinguishable from zero, but point estimates remain positive. Across several increasingly flexible specifications, I find little evidence to substantiate the claim that *Shelby* differentially restricted Black voter participation

43

in covered states. The non-negative association of *Shelby* with relative turnout and registration is consistent with counter-mobilization efforts outweighing increases in the relative cost of voting, on average.

I proceed in Section 2.2 by reviewing the extant literature on the Voting Rights Act, the *Shelby* decision, and the mechanisms through which *Shelby* could facilitate changes in relative turnout. I then turn to the data and research design in Section 2.3 before discussing the results in Section 2.4. Finally, in Section 2.5, I conclude with a brief discussion of the implications of my findings.

2.2 Background

2.2.1 The Voting Rights Act

Before the Voting Rights Act was passed in 1965, many states required prospective Black voters to pass literacy tests and pay poll taxes before voting.[9] States with these barriers exhibited vast racial disparities in voter registration and turnout (US Commission on Civil Rights, 2018). The Voting Rights Act and its subsequent revisions ameliorated racial disparities in voter participation by banning the use of literacy tests, providing the legal framework to challenge poll taxes and other discriminatory policies through litigation, and requiring federal approval of new state and local election policies in jurisdictions with histories of discrimination.

Before *Shelby*, litigation and preclearance were the primary enforcement mechanisms of the Voting Rights Act. Litigation targeted previously enacted election policies, and overturning a policy through this channel required plaintiffs to demonstrate that the policy was intended to discriminate or that it exhibited

---

[9] Literacy tests and poll taxes disenfranchised a broad class of otherwise eligible voters, but grandfather clauses re-enfranchised those whose ancestors were eligible to vote before the passage of the 15th Amendment, which guaranteed universal male suffrage. Few, if any, citizens of color had voting rights before the 15th Amendment was ratified, so grandfather clauses almost exclusively exempted white citizens from literacy tests and poll taxes.

discriminatory effects (Ho, 2017). In contrast, preclearance targeted policies before implementation, requiring covered jurisdictions to demonstrate to federal authorities (namely, the US Attorney General or the US District Court for the District of Columbia) that each proposed policy change was free of both discriminatory intent and discriminatory effects. A crucial distinction for the research design I outline in Section 2.3 is that preclearance enforcement was limited to covered states before *Shelby*, whereas the right to litigate applied, and continues to apply, nationally.

The extant literature on the Voting Rights Act leverages the coverage formula to identify causal effects of preclearance oversight on voter registration, turnout, and political representation. Using difference-in-differences variation in preclearance coverage within North Carolina, Fresh (2018) shows that preclearance generated large increases in voter registration and turnout rates in covered counties after the Voting Rights Act was enacted. Similarly, Ang (2019) leverages the expansion of preclearance coverage in the 1975 renewal of the Voting Rights Act and finds that preclearance produced lasting increases in the turnout rates of newly covered states and counties. Supplemental results from survey data suggest that the increases in overall turnout were driven by differential increases in voting among citizens of color. To document the effects of the Voting Rights Act on representation in local government, Bernini et al. (2018) exploit triple-difference variation in preclearance exposure within states of the former Confederacy. They show that preclearance increased the number of Black local elected officials in relatively Black counties. Using data on congressional roll-call votes and difference-in-differences variation in preclearance coverage, Schuit and Rogowski (2017) demonstrate that members of Congress who represented covered jurisdictions were more likely to support civil rights legislation after 1965 than those who represented uncovered jurisdictions. Taken together, the evidence suggests

that the Voting Rights Acts—and preclearance in particular—increased the political participation and representation of Black Americans.

Models of distributive politics predict that expansions of specific constituencies, such as those associated with the Voting Rights Act, can trigger changes in the distribution of public resources (Cox & Mccubbins, 1986; Lindbeck & Weibull, 1987). Specifically, if politicians elicit support by redistributing public resources, then an increase in the share of voters from a group with distinct preferences should induce an increase in public resources directed toward that group. For example, the re-enfranchisement of relatively poor, Black voters in the South appears to have increased government spending on redistributive social assistance programs (Husted & Kenny, 1997).

The quasi-experimental literature on the downstream consequences of the Voting Rights Act reinforces the importance of distributive politics in shaping economic outcomes. Using a triple-difference design that leverages within-state variation in racial composition and the use of literacy tests, Cascio and Washington (2014) show that the abolition of literacy tests in the South increased state transfers to predominately Black school districts. Using a similar research design, Facchini et al. (2020) find that preclearance decreased the rate at which Black individuals were arrested for non-felony offenses in covered counties with elected sheriffs and a high concentration of Black voters. To identify the effects of enfranchisement on labor market disparities, Avenancio-León and Aneja (2019) employ a triple-difference design that compares racial differences in labor market outcomes within covered counties to those within neighboring uncovered counties. They find that preclearance increased the relative wages of Black workers through additional government employment opportunities and enhanced anti-discrimination protections. Through re-enfranchisement and increased

representation, the Voting Right Act generated substantial economic benefits for Black Americans.

Other episodes of mass enfranchisement also demonstrate how reshaping the electorate induces policy changes that affect economic outcomes. For example, the enfranchisement of women in the United States during the early twentieth century prompted significant increases in public health spending which then decreased child mortality (Miller, 2008). Similarly, the *de facto* enfranchisement of less-educated Brazilians expanded access to public health care services which then increased utilization among uneducated mothers and improved infant health (Fujiwara, 2015). Mass disenfranchisement exhibits comparable effects—most pertinently, the imposition of poll taxes and literacy tests on Black voters in the Southern United States after Reconstruction reduced the allocation of public goods in Black communities (Naidu, 2012). A worry, then, is that the policy changes made possible by the *Shelby* decision could disenfranchise Black voters and ultimately reverse the economic gains brought about by the Voting Rights Act.

2.2.2 *Shelby County v. Holder*

On June 25, 2013, the Supreme Court ruled in a 5-4 decision that the coverage formula governing preclearance is unconstitutional. In an oft-cited passage of his majority opinion, Chief Justice Roberts argued that the coverage formula is "based on 40-year-old facts having no logical relation to the present day," and thus violates equal sovereignty of the states. Moreover, he argued that preclearance was "intended to be temporary," but Congress repeatedly renewed these sections without major revisions, most recently in 2006.[10] While the court did not rule on the constitutionality of preclearance itself, the invalidation of the coverage formula rendered preclearance

---

[10] *Shelby County v. Holder*, 570 U.S. 529 (2013).

unenforceable. Congress has since failed to enact new coverage formula that would comply with the *Shelby* decision and restore preclearance oversight.

After the ruling, previously covered states enacted a variety of election reforms.[11] Beyond new voter identification laws and other legislation, the *Shelby* decision also triggered an increase in voter list maintenance activity—which can remove (or "purge") otherwise eligible voters from registration lists—and may have prompted the closure of polling places. Using a difference-in-differences design, Feder and Miller (2020) estimate that voter purge rates increased in covered counties after *Shelby*, corroborating the findings of a Brennan Center report that documented a differential increase in purge rates in covered states (Brater et al., 2018).[12] An analysis of polling place data by *Vice News* documents that polling place closures were more common in covered counties than in uncovered counties after *Shelby*.[13] In contrast, Shepherd et al. (2021) find no systematic change in polling place locations within North Carolina, a state with covered and uncovered counties. Although the evidence on difficult-to-observe local policy responses remains sparse, the evidence on voter purges demonstrates that election administration in covered states responded systematically to the *Shelby* decision.

The existing literature has yet to establish whether the *Shelby* decision suppressed turnout among citizens of color. Ang (2019) provides a preliminary difference-in-differences analysis of the impact of *Shelby* on turnout. An event study

---

[11] Jaime Fuller. "How has voting changed since *Shelby County v. Holder*?" *Washington Post*, 7 July 2014. https://www.washingtonpost.com/news/the-fix/wp/2014/07/07/how-has-voting-changed-since-shelby -county-v-holder/?utm\%7B_\%7Dterm=.0e6ef3dabd32.

[12] Neither Feder and Miller (2020) nor Brater et al. (2018) use data that facilitate tests for race-specific heterogeneity, so it remains unclear whether Black voters were more likely to purged from the rolls.

[13] Rob Arthur and Allison McCann. "How the Gutting of the Voting Rights Act Led to Hundreds of Closed Polls." *Vice News*, 16 October 2018. https://news.vice.com/en_us/article/kz58qx/how-the-gutting-of -the-voting-rights-act-led-to-closed-polls.

of county election returns suggests that the difference in turnout between covered and uncovered counties decreased in 2016 relative to the difference in 2012, and event studies of self-reported turnout from the November Current Population Survey (CPS) voter supplement suggest that this decrease was concentrated among citizens of color. However, the same event studies also show that turnout was differentially increasing in covered jurisdictions before 2012, illustrating the difficulty of isolating the causal impact of *Shelby*. Using regression discontinuity and difference-in-differences comparisons of counties within North Carolina, Gibson (2020) finds no evidence that *Shelby* reduced turnout rates in covered counties, either overall or among specific racial subgroups. The North Carolina state government was also subject to preclearance, though, so the data foreclose on the ability to measure the consequences of statewide policy changes that would affect uncovered counties.

Whereas Ang (2019) and Gibson (2020) measure the impact of *Shelby* on absolute turnout (i.e., the Black turnout rate or the white turnout rate), I measure the impact of *Shelby* on relative turnout (i.e., the Black-white turnout gap). Changes in absolute turnout can provide important signals about the health of democratic institutions (e.g., a reduction in the turnout of citizens of any race could indicate an erosion of voting rights), but changes in relative turnout are important, too, in that they can bring about changes in the distribution of public resources. A focus on relative turnout also facilitates a research design that is robust to differential trends in absolute turnout rates between covered and uncovered states before *Shelby*.

2.2.3 Conceptual framework

Many of the policy changes made possible by the *Shelby* decision may have made voting more costly. For example, voter purges may have forced some prospective voters to re-register, a process that imposes nontrivial administrative burdens

(Braconnier, Dormagen, & Pons, 2017). Similarly, the closure or relocation of polling places may have imposed significant information and travel costs on voters (Cantoni, 2020; Clinton, Eubank, Fresh, & Shepherd, 2021). Absent some other mechanism, any increase in the cost of voting would decrease the likelihood of voting for those at the margin, putting downward pressure on aggregate turnout (Downs, 1957; Riker & Ordeshook, 1968). Thus, if the *Shelby* decision ultimately led to differential increases in the cost of voting for Black citizens, then one would expect Black turnout to decrease relative to white turnout.

Such a narrow focus on the costs of voting neglects the potential for counter-mobilization efforts to offset or even outweigh any new inconveniences of voting. Experimental evidence suggests that exposure to news portraying voter identification laws as discriminatory and disenfranchising generates anger that materializes in stronger voting intentions among Democrats, but not independents or Republicans (Valentino & Neuner, 2017). Likewise, quasi-experimental evidence suggests that challenging an individual's right to vote increases the probability that they will vote in subsequent elections, consistent with the predictions of psychological reactance theory (Biggers & Smith, 2020). Other quasi-experimental evidence suggests that political campaigns respond to strict voter identification laws by intensifying outreach toward voters of color (Cantoni & Pons, 2021).[14] Whether instigated by campaigns or voters themselves, the perception of selective disenfranchisement can trigger a backlash that mobilizes voters. If *Shelby* induced policy changes that were differentially salient for Black voters, then one would expect counter-mobilization to exert upward pressure on Black turnout relative to white turnout.

---

[14] Increased campaign activity might explain, in part, why strict voter identification laws tend to have minimal effects on turnout (Cantoni & Pons, 2021; Neiheisel & Horner, 2019).

Given the potential for the *Shelby* decision to put downward pressure on turnout through increasing the cost of voting and upward pressure on turnout through counter-mobilization, the net effect of *Shelby* is theoretically ambiguous.[15] The sign of the estimated effect of *Shelby* on relative turnout can therefore elucidate the relative importance of increases in the cost of voting and counter-mobilization. A decrease in the relative turnout of Black voters, for example, would suggest that the effects of differential increases in the cost of voting outweigh the effects of race-specific counter-mobilization, whereas an increase in relative turnout would suggest that the effects of race-specific counter-mobilization outweigh the effects of differential increases in the cost of voting. The results in Section 2.4 support the latter case in which the counter-mobilization response dominates.

2.3 Data and research design

2.3.1 Preclearance coverage

Preclearance coverage provides a source of identifying variation for estimating the impact of *Shelby* on the composition of the electorate. Figure 3 illustrates preclearance coverage when the coverage formula was invalidated in 2013.[16]

Fully covered states were explicitly identified by the coverage formula. Before 2013, all state and local authorities responsible for running or overseeing elections in fully covered states were required to secure preclearance for any policy change related

---

[15] Viewed this way, it is possible for *Shelby* to increase Black turnout while decreasing white turnout—even if *Shelby* induced policy changes that differentially increased the cost of voting for Black citizens. For example, it remains illegal for election officials to discriminate explicitly on the basis of race, so policy changes that target Black voters would likely affect some white voters, too. In this case, a sufficiently strong Black counter-mobilization response paired with a sufficiently weak white one would have the seemingly paradoxical consequence of simultaneously increasing Black turnout while decreasing white turnout.

[16] US Department of Justice. "Jurisdictions Previously Covered by Section 5." 11 September 2022. https://www.justice.gov/crt/jurisdictions-previously-covered-section-5 (accessed 29 December 2020).

to elections.[17] With the exception of Alaska and Arizona, most fully covered states were located in the South.

Figure 3. Preclearance coverage
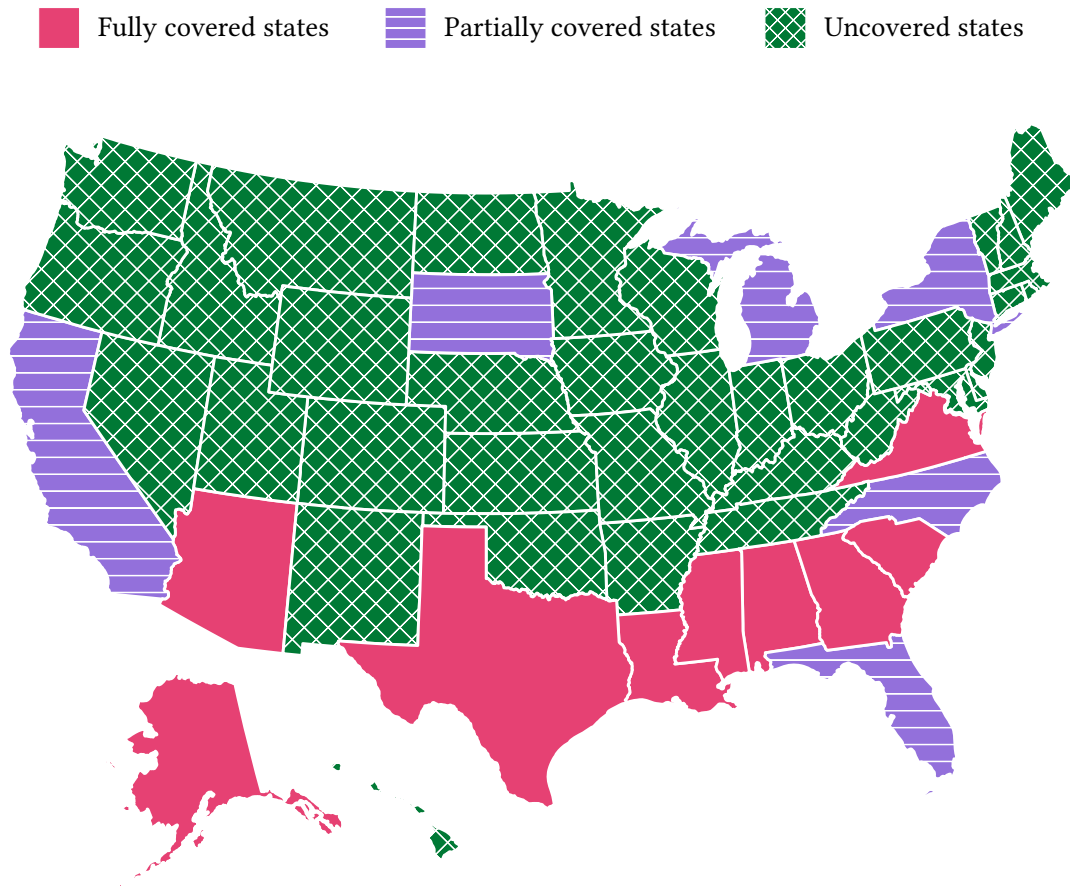
Notes: Preclearance coverage at the time of the *Shelby* decision. Before *Shelby*, all state and local authorities within fully covered states were required to obtain federal approval before changing any election policy. All state authorities within partially covered states were required to obtain federal approval, but most local authorities were not. State and local authorities within uncovered states were not required to obtain federal approval. The need for federal approval in covered states ended after the Supreme Court invalidated the coverage formula in 2013.

---

[17] Some counties in Virginia—a fully covered state—had "bailed out" of coverage before *Shelby*, which exempted them from preclearance requirements, but most counties were subject to preclearance.

Partially covered states were not explicitly identified by the coverage formula, but each had at least one local jurisdiction subject to preclearance. The presence of a covered jurisdiction in an otherwise uncovered state effectively extended coverage to the state government, as a statewide policy change would likely affect voting in the covered jurisdiction. Before 2013, state officials overseeing elections in partially covered states were required to obtain preclearance for any election policy change, but most local officials—namely those in uncovered jurisdictions—were not. For example, Michigan had two covered townships. A statewide voter identification law would require voters in those townships to provide identification as a precondition to voting, which would necessitate federal approval for the policy. In contrast, local officials in Wayne County, Michigan—an otherwise uncovered jurisdiction—could move polling places or change voting hours without federal approval.

Uncovered states were not identified by the coverage formula and had no covered local jurisdictions. Neither state nor local officials overseeing elections in these states were required to obtain preclearance for election policy changes before the *Shelby* decision. While the West, Midwest, and Northeast census regions contained most uncovered states, there were seven uncovered states in the South. In the analyses that follow, I use voters in uncovered states as a control group to estimate the effect of *Shelby* on the composition of voters in fully covered and partially covered states. Given the differences in the intensity of preclearance oversight, I estimate separate effects for fully covered and partially covered states.[18]

---

[18] One might anticipate additional margins of heterogeneity among covered states. For example, the Voting Rights Act of 1965 brought full preclearance oversight to states with the worst records of racial discrimination in election administration (i.e., Alabama, Georgia, Louisiana, Mississippi, South Carolina, and Virginia) whereas the 1975 reauthorization extended coverage to states that lacked voting disparities severe enough for coverage in the original 1965 legislation (i.e., Alaska, Arizona, and Texas). The results in Section 2.4 do not change, however, when I drop respondents from the expansion states, and point estimates are nearly identical when I estimate separate effects for expansion states and the initial group of fully covered states.

2.3.2 Turnout and registration data

Measuring changes in relative turnout and registration among Black and white voters requires data on turnout and registration by race, but official election returns do not typically publish race-specific turnout rates and most states do not release race-specific registration rates. For this reason, I use repeated cross sections from the Cooperative Congressional Election Study (CCES; Kuriwaki, 2020), a large-sample survey of the voting-age population administered after presidential and midterm elections. Each weighted cross section provides a representative sample of eligible voters within each state.

Unlike the November CPS and other national election surveys, the CCES validates self-reported turnout and registration against state voter files to correct for the tendency of respondents to over-report their political participation; only those with a verified record of voting are counted as having voted and only those with a verified registration record are counted as being registered. As a result, implied turnout rates from validated CCES turnout data are closer to actual state turnout rates derived from official election returns. Relying on validated voter participation data also sidesteps the potential for differential overreporting that coincides with exposure to the *Shelby* decision. The extent of overreporting in the CPS, for example, varies over time and across states (McDonald, 2021), and may also vary by race (Ansolabehere, Fraga, & Schaffner, 2021).[19]

CCES data are available for each presidential and midterm election since 2006, but I discard observations from the 2006 midterm election because the validation of

---

[19] To the extent that overreporting varies by race within states over time, triple-difference comparisons of self-reported voter participation could confound the effects of the *Shelby* decision with coincident changes in overreporting. In spite of the potential for differential overreporting, I provide estimates using CPS self-reports in the Appendix. As with data from the CCES, data from the CPS do not support the claim that *Shelby* caused significant reductions in relative turnout or registration among Black voters in previously covered states.

self-reported turnout against voter files was inconsistent across states (Grimmer, Hersh, Meredith, Mummolo, & Nall, 2018). I also discard observations from Virginia before 2012, as validation was unavailable there in 2008 and 2010 (Ansolabehere, 2010, 2012). The resulting sample consists of 35,322 non-Hispanic Black and 239,597 non-Hispanic white respondents from six federal elections (2008–2018).

I present absolute turnout rates, expressed as the percentage of adults who cast a ballot, for Black and white voters in Panels A and B of Figure 4. Black turnout was lower in covered states than in uncovered states before *Shelby* and the same is true of white turnout during presidential elections. While Black turnout rates decreased between the 2012 and 2016 presidential elections, this decrease was less pronounced in covered states than in uncovered states. Similarly, in midterm elections, the increase in Black turnout from the 2010 election to 2014 and 2018 was more pronounced in covered states than in uncovered states. Turnout rates also decreased among white voters between the 2012 and 2016 elections, though the decrease was somewhat less pronounced in uncovered states. In midterm elections, white turnout increased in uncovered states from 2010 to 2018, but remained constant in fully covered states. Difference-in-differences comparisons of absolute turnout rates in covered and uncovered states before and after *Shelby* would suggest that the *Shelby* decision is associated with an increase in Black turnout and a decrease in white turnout.

I turn to relative turnout, expressed as the difference between Black and white turnout rates, in Panel C of Figure 4. Negative turnout differentials throughout the sample period indicate that the white turnout rate exceeded the Black turnout rate during each election. Downward trends in turnout differentials indicate that the Black-white turnout gap increased between the 2008 and 2012 presidential elections. While turnout differentials did not meaningfully change in covered states between the 2012

Figure 4. Voter turnout and registration by race and preclearance coverage
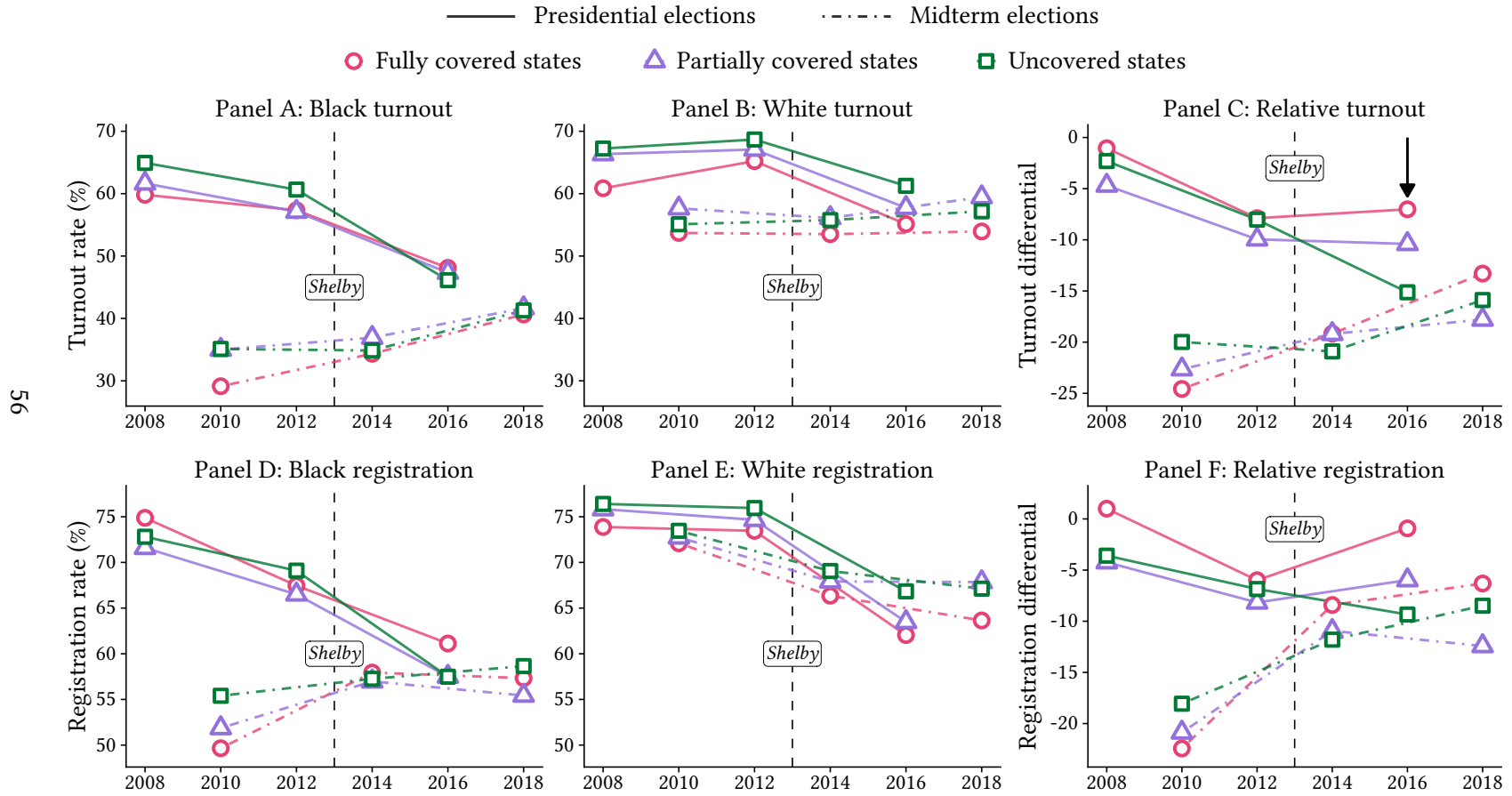
Notes: Implied turnout and registration rates derived from weighted averages of validated turnout responses of those in the analysis sample. The analysis sample consists of all Black and white CCES respondents, excluding those surveyed in Virginia during the 2008 presidential election and the 2010 midterms. Relative turnout (registration) refers to the difference in turnout (registration) rates between Black and white respondents.

and 2016 presidential elections, they continued to decrease in uncovered states. In other words, relative to the change in uncovered states, the Black-white turnout gap narrowed in covered states after *Shelby*. A relative narrowing of the turnout gap is less evident in midterm elections, though there is little indication that the gap widened in covered states after *Shelby*. Both patterns survive further scrutiny in Section 2.4.

Voter registration evidences a similar set of patterns. In Panel D, for instance, the decline in Black registration between 2012 and 2016 is smaller in covered states than in uncovered states, and in Panel E the decline in white registration is larger in covered states than in uncovered states. As with absolute turnout, difference-in-differences comparisons of absolute registration rates would suggest (albeit noisily) that *Shelby* increased Black registration and decreased white registration. Accordingly, in Panel F, Black-white registration gaps narrow in covered states between 2012 and 2016, but widen in uncovered states. With the possible exception of partially covered states in 2018, there is little indication that registration gaps widened in covered states after *Shelby*. As with relative turnout, the non-negative association of *Shelby* with relative registration survives further scrutiny in Section 2.4.

2.3.3 Empirical strategy

Naïve before-and-after comparisons of turnout or registration in covered states may confound the effects of *Shelby* with other factors that affect voter participation across elections, such as the presence of Black candidates for state or federal office (Washington, 2006), differences in *ex ante* expectations about the closeness of elections (Bursztyn, Cantoni, Funk, & Yuchtman, 2017), or changes in media consumption (DellaVigna & Kaplan, 2007; Gentzkow, 2006; Gerber, Karlan, & Bergan, 2009). Similarly, cross-sectional comparisons of covered states with uncovered states may confound the effects of *Shelby* with preexisting differences in participation—turnout

rates, for example, were lower in fully covered states than in uncovered states before *Shelby*. Even a difference-in-differences comparison could confound differential trends in turnout or registration between covered and uncovered states with the impact of *Shelby*.

To isolate the causal effect of *Shelby* on the relative turnout and registration of Black Americans, I employ a triple-difference design that leverages within-state variation in preclearance coverage. Specifically, I compare the Black-white turnout or registration differential within covered states to the differential within uncovered states, before and after the *Shelby* decision. I execute this comparison by estimating

$$\text{Participation}_{irst} = \beta \, \text{Black}_{ir} \times \text{fully covered}_s \times \textit{Shelby}_t$$

$$+ \, \delta \, \text{Black}_{ir} \times \text{partially covered}_s \times \textit{Shelby}_t \qquad (2.1)$$

$$+ \, \alpha_{rs} + \alpha_{rt} + \alpha_{st} + X'_{irst}\Gamma + \varepsilon_{irst} \, ,$$

where Participation$_{irst}$ is an indicator equal to one if respondent i of race r voted (or was registered to vote) in state s during election t or zero if the respondent did not vote (or was not registered to vote). Race-by-state fixed effects (α_{rs}) absorb time-invariant factors that affect the participation of Black or white voters within each state. Race-by-year fixed effects (α_{rt}) absorb election-specific characteristics that affect participation among all Black voters or all white voters across the country. State-by-year fixed effects (α_{st}) absorb election-specific characteristics that affect participation among all voters within each state, such as the presence of a Senate race in a midterm election. Respondent controls (X'_{irst}) adjust for differences in participation by age and gender. Each treatment interaction consists of a race indicator, a coverage indicator, and a post-treatment indicator.[20] The race indicator (Black$_{ir}$) equals one if respondent i identified as Black or zero if the respondent identified as white. Coverage indicators

---

[20] Fixed effects absorb all lower-order terms.

(fully covered$_s$ and partially covered$_s$) equal one if state s was covered before *Shelby* or zero if the state was uncovered. The post-treatment indicator (*Shelby$_t$*) equals one if election t was held after 2013 or zero if the election was held before 2013. The triple-difference parameters (β and δ) capture the effect of *Shelby* on the relative participation of eligible Black voters in fully covered and partially covered states after 2013. Following the conservative approach of Cameron and Miller (2015), I make inference using cluster-robust standard errors that account for clustering at the state level.

A causal interpretation of $\hat{\beta}$ and $\hat{\delta}$ rests on two identifying assumptions. The first asserts common trends—that is, conditional on the full set of fixed effects and respondent controls, the Black-white turnout or registration differential in covered states would have evolved similarly to the differential in uncovered states had the Supreme Court upheld the coverage formula. The second asserts that consequences of the *Shelby* decision did not spill over into uncovered states.

I gauge the plausibility of common trends by estimating an event study analog of Equation 2.1:

$$\text{Participation}_{irst} = \sum_{\tau \neq 2012} \beta_\tau \, \text{Black}_{ir} \times \text{fully covered}_s \times \mathbb{1}(t = \tau)_\tau$$

$$+ \sum_{\tau \neq 2012} \delta_\tau \, \text{Black}_{ir} \times \text{partially covered}_s \times \mathbb{1}(t = \tau)_\tau \qquad (2.2)$$

$$+ \alpha_{rs} + \alpha_{rt} + \alpha_{st} + X'_{irst}\Gamma + \varepsilon_{irst} \, ,$$

where β_τ represents the difference in Black-white turnout differentials between fully covered states and uncovered states in year τ relative to the difference in 2012 and δ_τ represents the difference between partially covered states and uncovered states. Estimates of β_τ or δ_τ that deviate from zero for elections before *Shelby* would suggest that covered and uncovered states do not share common trends in participation

differentials. If both identifying assumptions hold, then estimates of β_τ or δ_τ for elections after *Shelby* help illustrate how the effects of the decision have evolved over time.

2.4   Results

2.4.1   Event study

Estimates of the event study coefficients from Equation 2.2 lend support to the plausibility of common trends between covered and uncovered states. In Panel A of Figure 5, for instance, the differences in Black-white turnout differentials between fully covered states and uncovered states during the 2008 and 2010 elections are indistinguishable from the difference in 2012 at conventional significance levels ($p = 0.477$ in 2008 and $p = 0.562$ in 2010). The same is true of the relative differences in Black-white turnout differentials between partially covered states and uncovered states ($p = 0.89$ in 2008 and $p = 0.671$ in 2010). Furthermore, all four pre-2012 coefficients are jointly indistinguishable from zero ($F = 0.9$, $p = 0.463$). The analysis of voter registration in Panel B concludes similarly, as all four pre-2012 coefficients are individually indistinguishable from zero ($p = 0.255$ and $p = 0.353$ for fully covered states in 2008 and 2010, and $p = 0.889$ and $p = 0.798$ for partially covered states in 2008 and 2010) and jointly indistinguishable from zero ($F = 1.54$, $p = 0.189$). Taken together, these tests fail to reject common trends between covered and uncovered states before *Shelby*.

Rather than diminishing the Black share of the electorate, *Shelby* may have done the opposite—the positive-signed, post-*Shelby* event study estimates in Panel A of Figure 5 suggest that, if anything, Black turnout increased relative to white turnout within covered states. The estimate of the 2016 event study coefficient for fully covered states, for example, indicates that the difference in relative turnout between fully

60

Figure 5. Event study of the effect of *Shelby v. Holder* on relative turnout and voter registration
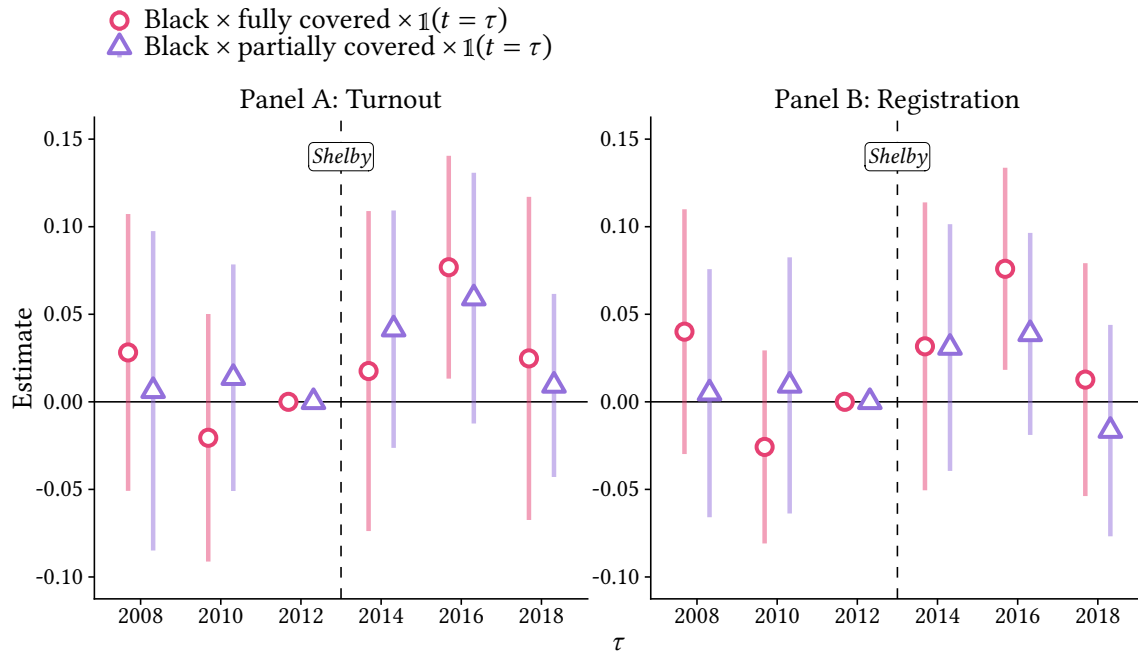
Notes: OLS estimates of event study coefficients from Equation 2.2. Each panel contains estimates from a single regression, and each regression includes race-by-state fixed effects, race-by-year fixed effects, state-by-year fixed effects, and controls for gender, age, and age squared. The sample includes Black and white respondents from six federal elections (2008–2018), weighted using CCES-provided sampling weights. Vertical bars outline pointwise 95% confidence intervals that are robust to clustering at the state level.

covered and uncovered states was 7.7 percentage points higher ($p = 0.019$), on average, in 2016 than in 2012. For partially covered states, the average relative increase between the 2012 and 2016 elections is similar in magnitude, though indistinguishable from zero at conventional significance levels (5.9 percentage points, $p = 0.103$). Estimates for the 2014 and 2018 midterm elections, while positive-signed, are smaller in magnitude and statistically insignificant. The event study estimates in Panel B evidence a similar story for voter registration, though the magnitude of the relative increase between the 2012 and 2016 elections is smaller for partially covered states (3.9 percentage points, $p = 0.183$) than for fully covered states (7.6 percentage points, $p = 0.011$). In any case,

Black voter participation does not appear to have decreased any more than white voter participation within covered states after *Shelby*.

2.4.2 Turnout

In Table 2, I present estimates of the effect of *Shelby* on relative turnout from triple-difference specifications based on Equation 2.1. Baseline estimates in column (1) suggest that *Shelby* had a positive effect on the relative turnout of Black voters, on average, in fully covered states (4.9 percentage-point increase, $p = 0.035$) and in partially covered states (3.8 percentage-point increase, $p = 0.001$). In other words, the *Shelby* decision is associated with a decrease in the Black-white turnout gap in favor of Black voters. While the estimates attenuate with the addition of controls for age and gender in column (2), they remain positive and statistically significant at the 10-percent level. The inclusion of state-by-race time trends in column (3) serves to probe whether the positive results are artifacts of differential trends within states. Although the increase in standard errors renders the triple-difference coefficients indistinguishable from zero, the point estimates remain positive and are somewhat larger in magnitude than those in column (1). Like the event study in Panel A of Figure 5, the estimates in columns (1)–(3) indicate that, if anything, *Shelby* differentially increased Black turnout.

If the net effect of *Shelby* on turnout equals the sum of a negative turnout response to differential increases in the cost of voting and a positive turnout response to (race-specific) counter-mobilization, then evidence of a relative increase in Black turnout suggests that counter-mobilization among Black voters outweighed any differential increase in the cost of voting, on average. One might also expect the relative importance of the cost response and the counter-mobilization response to vary by election type. For example, Panels A and B of Figure 4 show that the decrease in turnout between presidential and midterm elections is consistently more pronounced

Table 2. Effect of *Shelby v. Holder* on relative turnout

	Turnout					
	(1)	(2)	(3)	(4)	(5)	(6)
Black × fully covered × *Shelby*	0.049**	0.037*	0.052			
	(0.022)	(0.021)	(0.046)			
Black × fully covered × *Shelby* × presidential				0.089***	0.075***	0.089**
				(0.020)	(0.019)	(0.043)
Black × fully covered × *Shelby* × midterm				0.029	0.019	0.032
				(0.028)	(0.027)	(0.051)
Black × partially covered × *Shelby*	0.038***	0.030**	0.057			
	(0.011)	(0.012)	(0.042)			
Black × partially covered × *Shelby* × presidential				0.066***	0.052**	0.080
				(0.025)	(0.026)	(0.052)
Black × partially covered × *Shelby* × midterm				0.024	0.019	0.046
				(0.014)	(0.016)	(0.041)
Observations	274,919	274,919	274,919	274,919	274,919	274,919
Effective observations (race × state × year)	598	598	598	598	598	598
Race × state fixed effects	✓	✓	✓	✓	✓	✓
Race × year fixed effects	✓	✓	✓	✓	✓	✓
State × year fixed effects	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓
Race × state time trends			✓			✓

Notes: OLS estimates of triple-difference coefficients from Equation 2.1. Fixed effects absorb all lower-order terms. Demographic controls include gender, age, and age squared. The sample includes Black and white respondents from six federal elections (2008–2018), weighted using CCES-provided sampling weights. Standard errors (in parentheses) are robust to clustering at the state level.

* $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

for Black voters than for white voters. While race-by-year fixed effects absorb this variation, the contraction of Black turnout in midterm elections could nevertheless limit the potential reach of counter-mobilization efforts or otherwise improve the prospects of new discriminatory legislation, increasing the relative importance of differential increases in the costs of voting during midterm elections. Alternatively, it could be that individuals who vote in midterm elections are less responsive to voting restrictions than those who only vote in presidential elections, decreasing the relative importance of differential increases in the costs of voting during midterm elections. The pattern of event study coefficients in Panel A of Figure 5 suggests that the positive effects of the *Shelby* decision on relative turnout were smaller in magnitude for midterm elections, consistent with a weaker counter-mobilization response or a stronger cost response.

I explore this possibility in further detail in columns (4)–(6) of Table 2. Estimates from the preferred specification in column (5) indicate that the increases documented in the first three specifications mask differences between presidential and midterm elections. In presidential elections, the *Shelby* decision increased Black turnout relative to white turnout by 7.5 percentage points ($p < 0.001$), on average, in fully covered states and 5.2 percentage points ($p = 0.047$), on average, in partially covered states. These increases represent a full reduction of the Black-white turnout gap that existed in fully covered states during the 2012 presidential election and a 52-percent reduction of the gap that existed in partially covered states, and are consistent with the effects of counter-mobilization outweighing the effects of differential increases in the cost of voting. In midterm elections, *Shelby* had no significant effect on the relative turnout of Black voters in fully covered states (1.9 percentage points, $p = 0.479$) or partially covered states (1.9 percentage points, $p = 0.238$). Still, the null findings for midterm

64

elections remain consistent with counter-mobilization offsetting the effects of increases in the cost of voting.

Although the inclusion of race-by-state time trends in column (6) reduces precision, all triple-difference estimates remain positive and the coefficient for fully covered states in presidential elections remains statistically distinguishable from zero at the 5-percent level. The absence of negative turnout responses across several specifications, coupled with the relative increase in Black turnout during the 2016 presidential election, contradict the notion that the *Shelby* decision enabled previously covered states to reshape the electorate by suppressing the Black vote. More generally, these results highlight the potential for counter-mobilization to offset the effects of differential increases in the cost of voting.

While the data do not support that *Shelby* differentially reduced average turnout rates among Black voters, this does not imply that no individual was disenfranchised by the decision. A null effect of *Shelby* on Black turnout could reflect a scenario in which counter-mobilization increased turnout among some Black voters while increases in the cost of voting disenfranchised others. It also remains unclear whether counter-mobilization would continue to offset differential increases in the cost of voting in the long run. If mobilizing anger diminishes as differential increases in the costs of voting persist, then the *Shelby* decision could eventually suppress the Black vote. Still, the non-negative impact on the relative turnout of Black voters in the first three federal elections after *Shelby* is an encouraging result.

2.4.3   Registration

In Table 3, I repeat the same exercise for voter registration. Baseline estimates in column (1) suggest that *Shelby* had no discernible effect on the registration of Black voters relative to white voters in fully covered states (4.3 percentage points, $p = 0.102$)

Table 3. Effect of *Shelby v. Holder* on relative voter registration

	Registration					
	(1)	(2)	(3)	(4)	(5)	(6)
Black × fully covered × *Shelby*	0.043	0.036	0.078*			
	(0.026)	(0.025)	(0.046)			
Black × fully covered × *Shelby* × presidential				0.082***	0.072***	0.113**
				(0.025)	(0.024)	(0.047)
Black × fully covered × *Shelby* × midterm				0.025	0.018	0.059
				(0.031)	(0.031)	(0.048)
Black × partially covered × *Shelby*	0.019	0.013	0.052			
	(0.017)	(0.015)	(0.051)			
Black × partially covered × *Shelby* × presidential				0.044	0.034	0.074
				(0.028)	(0.027)	(0.059)
Black × partially covered × *Shelby* × midterm				0.006	0.003	0.041
				(0.018)	(0.016)	(0.049)
Observations	274,919	274,919	274,919	274,919	274,919	274,919
Effective observations (race × state × year)	598	598	598	598	598	598
Race × state fixed effects	✓	✓	✓	✓	✓	✓
Race × year fixed effects	✓	✓	✓	✓	✓	✓
State × year fixed effects	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓
Race × state time trends			✓			✓

Notes: OLS estimates of triple-difference coefficients from Equation 2.1. Fixed effects absorb all lower-order terms. Demographic controls include gender, age, and age squared. The sample includes Black and white respondents from six federal elections (2008–2018), weighted using CCES-provided sampling weights. Standard errors (in parentheses) are robust to clustering at the state level.

* $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

or in partially covered states (1.9 percentage points, $p = 0.278$), but estimates from the preferred specification in column (5) indicate that the baseline estimates mask heterogeneous effects by election type. In fully covered states, the *Shelby* decision increased Black registration relative to white registration during presidential elections (7.2 percentage points, $p = 0.005$), but not during midterm elections (1.8 percentage points, $p = 0.56$). Both coefficients for partially covered states are indistinguishable from zero, though the point estimate for presidential elections (3.4 percentage points, $p = 0.223$) is larger in magnitude than the point estimate for midterm elections (0.3 percentage points, $p = 0.865$). As with the analysis of turnout in Table 2, no specification in Table 3 supports that *Shelby* widened the Black-white registration gap in covered states. If anything, *Shelby* may have narrowed the gap in fully covered states during the 2016 presidential election.

In light of evidence that covered states responded to *Shelby* by purging voters (Brater et al., 2018; Feder & Miller, 2020), a non-negative effect of *Shelby* on relative registration further reinforces the potential role of counter-mobilization in maintaining Black voter participation. For example, if Black voters were more likely than white voters to be purged from the rolls in covered states after *Shelby*, then even a null effect of *Shelby* on relative registration would suggest that a significant number of Black voters re-registered after being purged, or that previously unregistered Black voters decided to register in response.

2.5 Conclusion

As part of the argument against the constitutionality of preclearance coverage, Chief Justice Roberts cited data from the November CPS as evidence that the Black-white turnout gap had diminished or even reversed in covered states since the Voting Rights Act was enacted. Researchers have documented, however, that self-reported CPS

turnout data mask significant racial disparities in voting (Ansolabehere et al., 2021). Validated turnout data from the CCES, for example, show that significant racial gaps in turnout existed in covered states before *Shelby* and have since continued to exist.

Despite well-founded fears to the contrary, the *Shelby* decision does not appear to have widened participation gaps between Black and white voters in previously covered states relative to those in uncovered states. Triple-difference comparisons of validated turnout and registration data suggest that *Shelby* had little effect on participation gaps during the first three federal elections that followed the 2013 decision, and, in the 2016 election, may have even increased the relative turnout of Black voters in covered states. These results are consistent with an accumulating body of evidence that suggests that voters and campaigns mobilize in response to increases in the cost of voting when those increases are perceived as threats to the franchise (Biggers & Smith, 2020; Cantoni & Pons, 2021; Valentino & Neuner, 2017).

The resilience of Black turnout in previously covered states provides some grounds for optimism. That said, voters of color do continue to experience systematic barriers to voting (Chen, Haggag, Pope, & Rohla, 2020) and questions remain about the impact of *Shelby* on state and local elections, which could pose important consequences for the provision of public resources and the enforcement of anti-discrimination laws.

APPENDIX

ALTERNATIVE ESTIMATES USING CURRENT POPULATION SURVEY DATA

I provide alternative triple-difference estimates of the effect of *Shelby County v. Holder* using data from the voting and registration supplement of the November Current Population Survey (CPS; Flood, King, Rodgers, Ruggles, & Warren, 2018). To facilitate comparisons with the CCES, I estimate the same specifications described in the main text for a weighted sample of Black and White voting-age adults surveyed between 2008 and 2018. Note, however, that the CPS does not validate self-reported turnout or registration against state voter files. As a result, CPS data feature higher turnout and registration rates and smaller Black-white participation gaps than those in the CCES (Ansolabehere et al., 2021).

There are multiple ways to define turnout using CPS self-reports, and each definition differs based on the classification of nonvoters. The US Census Bureau, for example, classifies any respondent who does not answer "yes" to the question as a nonvoter. This includes those that answer "no" or "don't know" to the turnout question, as well as those who refuse to answer and those who are 18 or older but were not asked whether they voted. In contrast, many political scientists prefer to classify nonvoters as those that answer "no" to the turnout question, treating all other non-"yes" responses as missing data (McDonald, 2021).

I construct three definitions of turnout and registration that gradually increase the number of missing cases. In the first, I adopt the Census definition, as described above, which does not treat any non-"yes" response as missing. In the second, I treat those who were not asked the turnout or registration question as missing cases, but otherwise keep those who refused to answer or did not know whether they voted or

registered. In the third, I restrict the sample to those who provided an explicit "yes" or "no" answer to the turnout or registration question, treating all others as missing.

I present event study estimates for each measure of turnout and registration in Figure A.6 and triple-difference estimates in Table A.4, Table A.5, Table A.6, and Table A.7. No measure of turnout or registration supports that *Shelby* significantly reduced relative participation of Black voters in previously covered states.

Figure A.6. Alternative event study estimates of the effect of *Shelby* on relative turnout and voter registration using CPS data
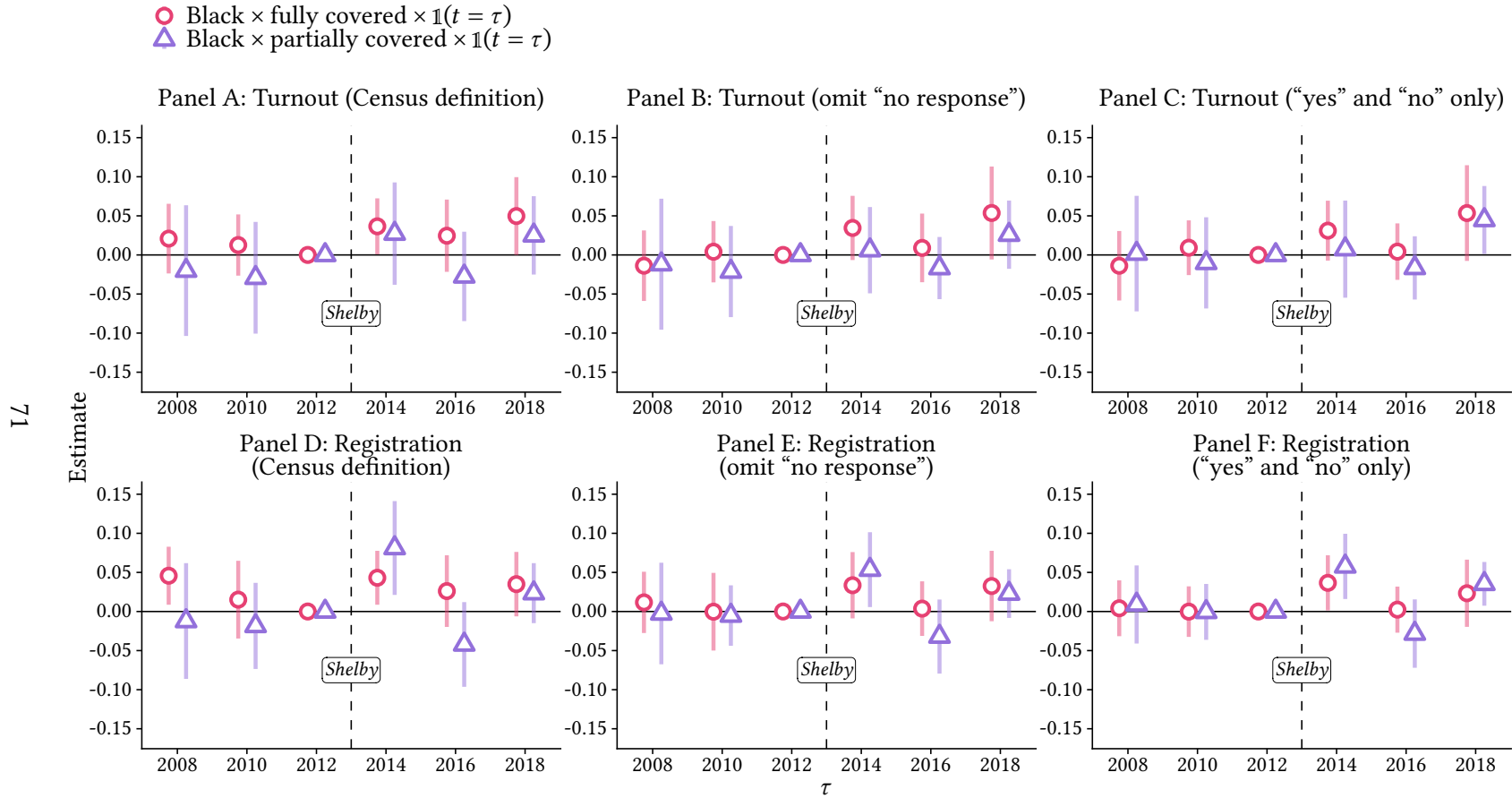
Notes: OLS estimates of event study coefficients. Each regression model contains race-by-state fixed effects, race-by-year fixed effects, state-by-year fixed effects, and controls for gender, age, and age squared. The sample includes Black and White respondents from six federal elections (2008–2018), weighted using CPS-provided sampling weights. Vertical bars outline pointwise 95% confidence intervals that are robust to clustering at the state level.

Table A.4. Alternative estimates of the effect of *Shelby* on relative turnout using CPS data

	Turnout (Census definition)			Turnout (omit "no response")			Turnout ("yes" or "no" only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black × fully covered × *Shelby*	0.026*	0.026*	0.031	0.035	0.043	0.012	0.031	0.036	0.004
	(0.014)	(0.015)	(0.025)	(0.026)	(0.026)	(0.033)	(0.024)	(0.023)	(0.028)
Black × partially covered × *Shelby*	0.022	0.024*	0.011	0.010	0.012	0.006	0.028	0.029	-0.028
	(0.015)	(0.014)	(0.040)	(0.028)	(0.024)	(0.055)	(0.029)	(0.026)	(0.085)
Observations	476,010	476,010	476,010	425,999	425,999	425,999	411,226	411,226	411,226
Effective observations (race × state × year)	611	611	611	611	611	611	611	611	611
Race × state fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Race × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
State × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓		✓	✓
Race × state time trends			✓			✓			✓

Notes: OLS estimates of triple-difference coefficients. Fixed effects absorb all lower-order terms. Demographic controls include gender, age, and age squared. The sample includes Black and White respondents from six federal elections (2008–2018), weighted using CPS-provided sampling weights. Standard errors (in parentheses) are robust to clustering at the state level.
* $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.5. Alternative estimates of the heterogeneous effects of *Shelby* on relative turnout using CPS data

	Turnout (Census definition)			Turnout (omit "no response")			Turnout ("yes" or "no" only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black × fully covered × *Shelby* × presidential	0.013	0.014	0.019	0.021	0.032	0.0003	-0.006	0.007	-0.026
	(0.020)	(0.019)	(0.034)	(0.030)	(0.027)	(0.040)	(0.022)	(0.022)	(0.038)
Black × fully covered × *Shelby* × midterm	0.032**	0.032**	0.037	0.044	0.049*	0.017	0.048*	0.050*	0.016
	(0.016)	(0.016)	(0.022)	(0.028)	(0.028)	(0.033)	(0.029)	(0.027)	(0.028)
Black × partially covered × *Shelby* × presidential	-0.011	-0.011	-0.024	-0.021	-0.010	-0.021	0.005	0.018	-0.040
	(0.024)	(0.024)	(0.043)	(0.044)	(0.040)	(0.056)	(0.038)	(0.036)	(0.086)
Black × partially covered × *Shelby* × midterm	0.039***	0.042***	0.029	0.026	0.023	0.015	0.040	0.034	-0.025
	(0.014)	(0.013)	(0.040)	(0.026)	(0.021)	(0.057)	(0.028)	(0.022)	(0.086)
Observations	476,010	476,010	476,010	425,999	425,999	425,999	411,226	411,226	411,226
Effective observations (race × state × year)	611	611	611	611	611	611	611	611	611
Race × state fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Race × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
State × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓		✓	✓
Race × state time trends			✓			✓			✓

Notes: OLS estimates of triple-difference coefficients for each election type. Fixed effects absorb all lower-order terms. Demographic controls include gender, age, and age squared. The sample includes Black and White respondents from six federal elections (2008–2018), weighted using CPS-provided sampling weights. Standard errors (in parentheses) are robust to clustering at the state level.
* $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.6. Alternative estimates of the effect of *Shelby* on relative voter registration using CPS data

	Registration (Census definition)			Registration (omit "no response")			Registration ("yes" or "no" only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black × fully covered × *Shelby*	0.015	0.015	0.054**	0.019	0.023	0.003	0.011	0.014	0.014
	(0.011)	(0.011)	(0.025)	(0.022)	(0.021)	(0.035)	(0.018)	(0.016)	(0.028)
Black × partially covered × *Shelby*	0.029**	0.030**	0.065*	0.003	0.004	0.022	0.017	0.017	0.019
	(0.012)	(0.011)	(0.035)	(0.024)	(0.023)	(0.049)	(0.023)	(0.021)	(0.036)
Observations	476,010	476,010	476,010	425,384	425,384	425,384	408,641	408,641	408,641
Effective observations (race × state × year)	611	611	611	611	611	611	611	611	611
Race × state fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Race × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
State × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓		✓	✓
Race × state time trends			✓			✓			✓

Notes: OLS estimates of triple-difference coefficients. Fixed effects absorb all lower-order terms. Demographic controls include gender, age, and age squared. The sample includes Black and White respondents from six federal elections (2008–2018), weighted using CPS-provided sampling weights. Standard errors (in parentheses) are robust to clustering at the state level.
* $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table A.7. Alternative estimates of the heterogeneous effects of *Shelby* on relative voter registration using CPS data

	Registration (Census definition)			Registration (omit "no response")			Registration ("yes" or "no" only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black × fully covered × *Shelby* × presidential	0.006	0.006	0.045	0.001	0.009	-0.011	-0.011	-0.005	-0.006
	(0.021)	(0.020)	(0.034)	(0.023)	(0.020)	(0.034)	(0.019)	(0.018)	(0.032)
Black × fully covered × *Shelby* × midterm	0.019	0.019	0.058**	0.029	0.031	0.008	0.022	0.023	0.021
	(0.012)	(0.012)	(0.023)	(0.025)	(0.024)	(0.039)	(0.024)	(0.021)	(0.030)
Black × partially covered × *Shelby* × presidential	-0.032	-0.032	0.004	-0.054	-0.046	-0.037	-0.040	-0.034	-0.029
	(0.027)	(0.027)	(0.036)	(0.052)	(0.051)	(0.052)	(0.056)	(0.057)	(0.047)
Black × partially covered × *Shelby* × midterm	0.059***	0.062***	0.098**	0.033*	0.030*	0.039	0.039*	0.037**	0.035
	(0.011)	(0.010)	(0.038)	(0.020)	(0.018)	(0.056)	(0.020)	(0.016)	(0.041)
Observations	476,010	476,010	476,010	425,384	425,384	425,384	408,641	408,641	408,641
Effective observations (race × state × year)	611	611	611	611	611	611	611	611	611
Race × state fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Race × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
State × year fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demographic controls		✓	✓		✓	✓		✓	✓
Race × state time trends			✓			✓			✓

Notes: OLS estimates of triple-difference coefficients for each election type. Fixed effects absorb all lower-order terms. Demographic controls include gender, age, and age squared. The sample includes Black and White respondents from six federal elections (2008–2018), weighted using CPS-provided sampling weights. Standard errors (in parentheses) are robust to clustering at the state level.
* $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

REFERENCES CITED

Anderson, K. P., & Ritter, G. W. (2017). Disparate use of exclusionary discipline: Evidence on inequities in school discipline from a US state. *Education Policy Analysis Archives*, *25*(9).

Anderson, K. P., & Ritter, G. W. (2020). Do school discipline policies treat students fairly? Evidence from Arkansas. *Educational Policy*, *34*(5), 707–734.

Anderson, K. P., Ritter, G. W., & Zamarro, G. (2019). Understanding a vicious cycle: The relationship between student discipline and student academic outcomes. *Educational Researcher*, *48*(5), 251–262.

Ang, D. (2019). Do 40-year-old facts still matter? Long-run effects of federal oversight under the Voting Rights Act. *American Economic Journal: Applied Economics*, *11*(3), 1–53.

Ansolabehere, S. (2010). *Guide to the 2008 Cooperative Congressional Election Survey* (Codebook Version 6). Harvard Dataverse. Retrieved from https://doi.org/10.7910/DVN/YUYIVB/ODKLI3

Ansolabehere, S. (2012). *Guide to the 2010 Cooperative Congressional Election Survey* (Codebook Version 3). Harvard Dataverse. Retrieved from https://doi.org/10.7910/DVN/VKKRWA/CO1QSD

Ansolabehere, S., Fraga, B. L., & Schaffner, B. F. (2021). The CPS Voting and Registration Supplement overstates minority turnout. *Journal of Politics*. Retrieved from https://www.bernardfraga.com/s/CPS_AFS_2021.pdf

Arcidiacono, P., Aucejo, E. M., & Spenner, K. (2012). What happens after enrollment? An analysis of the time path of racial differences in GPA and major choice. *IZA Journal of Labor Economics*, *1*(1).

Avenancio-León, C. F., & Aneja, A. (2019). *The effect of political power on labor market inequality: Evidence from the 1965 Voting Rights Act.* Retrieved from https://abhayaneja.files.wordpress.com/2019/09/vralabor_sept2019.pdf

Bacher-Hicks, A., Billings, S. B., & Deming, D. J. (2019). *The school to prison pipeline: Long-run impacts of school suspensions on adult crime* (NBER Working Paper No. 26257). National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w26257

Barrett, N., McEachin, A., Mills, J. N., & Valant, J. (2021). Disparities and discrimination in student discipline by race and family income. *Journal of Human Resources*, *56*(3), 711–748.

Bazerman, M. H., Moore, D. A., Tenbrunsel, A. E., Wade-Benzoni, K. A., & Blount, S. (1999). Explaining how preferences change across joint versus separate evaluation. *Journal of Economic Behavior & Organization*, *39*(1), 41–58.

Beck, A. N., & Muschkin, C. G. (2012). The enduring impact of race: Understanding disparities in student disciplinary infractions and achievement. *Sociological Perspectives*, *55*(4), 637–662.

Bernini, A., Facchini, G., & Testa, C. (2018). *Race, representation and local governments in the U.S. South: The effect of the Voting Rights Act* (CEPR Discussion Paper No. 12774). Centre for Economic Policy Research. Retrieved from https://cepr.org/active/publications/discussion_papers/dp.php?dpno=12774

Biggers, D. R., & Smith, D. A. (2020). Does threatening their franchise make registered voters more likely to participate? Evidence from an aborted voter purge. *British Journal of Political Science*, *50*(3), 933–954.

Braconnier, C., Dormagen, J.-Y., & Pons, V. (2017). Voter registration costs and disenfranchisement: Experimental evidence from France. *American Political Science Review*, *111*(3), 584–604.

Brater, J., Morris, K., Pérez, M., & Deluzio, C. (2018). *Purges: A growing threat to the right to vote* (Research Report). Brennan Center for Justice. Retrieved from https://www.brennancenter.org/our-work/research-reports/purges-growing-threat-right-vote

Bursztyn, L., Cantoni, D., Funk, P., & Yuchtman, N. (2017). *Polls, the press, and political participation: The effects of anticipated election closeness on voter turnout* (NBER Working Paper No. 23490). National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w23490

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372.

Cantoni, E. (2020). A precinct too far: Turnout and voting costs. *American Economic Journal: Applied Economics*, *12*(1), 61–85.

Cantoni, E., & Pons, V. (2021). Strict ID laws don't stop voters: Evidence from a U.S. nationwide panel, 2008–2018. *Quarterly Journal of Economics*, *136*(4), 2615–2660.

Card, D., & Payne, A. A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry*, *59*(1), 9–28.

Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, *2*(1), 211–228.

Cascio, E. U., & Washington, E. (2014). Valuing the vote: The redistribution of voting rights and state funds following the Voting Rights Act of 1965. *Quarterly Journal of Economics*, *129*(1), 379–433.

Chen, M. K., Haggag, K., Pope, D. G., & Rohla, R. (2020). Racial disparities in voting wait times: Evidence from smartphone data. *Review of Economics and Statistics*. Retrieved from https://doi.org/10.1162/rest_a_01012

Children's Defense Fund. (1975). *School suspensions: Are they helping children?* Cambridge, MA: Children's Defense Fund.

Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, *49*(8), 566–578.

Clinton, J. D., Eubank, N., Fresh, A., & Shepherd, M. E. (2021). Polling place changes and political participation: Evidence from North Carolina presidential elections, 2008–2016. *Political Science Research and Methods*, *9*(4), 800–817.

Cox, G. W., & Mccubbins, M. D. (1986). Electoral politics as a redistributive game. *Journal of Politics*, *48*(2), 370–389.

Craig, A. C., & Martin, D. C. (2019). *Discipline reform, school culture, and student achievement.* Retrieved from https://scholar.harvard.edu/files/david-martin/files/martin_jmp.pdf

Curran, F. C. (2016). Estimating the effect of state zero tolerance laws on exclusionary discipline, racial discipline gaps, and student behavior. *Educational Evaluation and Policy Analysis*, *38*(4), 647–668.

DellaVigna, S., & Kaplan, E. (2007). The Fox News effect: Media bias and voting. *Quarterly Journal of Economics*, *122*(3), 1187–1234.

Downs, A. (1957). *An economic theory of democracy.* New York: Harper and Row.

Facchini, G., Knight, B. G., & Testa, C. (2020). *The franchise, policing, and race: Evidence from arrests data and the Voting Rights Act* (NBER Working Paper No. 27463). National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w27463

Feder, C., & Miller, M. G. (2020). Voter purges after Shelby. *American Politics Research*, *48*(6), 687–692.

Filer, J. E., Kenny, L. W., & Morton, R. B. (1991). Voting laws, educational policies, and minority turnout. *The Journal of Law and Economics*, *34*(2), 371–393.

Flood, S., King, M., Rodgers, R., Ruggles, S., & Warren, J. R. (2018). *Current Population Survey: Version 6.0* (Integrated Public Use Microdata Series). Minnesota Population Center. Retrieved from https://doi.org/10.18128/D030.V6.0

Fresh, A. (2018). The effect of the Voting Rights Act on enfranchisement: Evidence from North Carolina. *Journal of Politics*, *80*(2).

Fudenberg, D., & Tirole, J. (1986). A "signal-jamming" theory of predation. *The RAND Journal of Economics*, *617*(3), 366-376.

Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: Evidence from Brazil. *Econometrica*, *83*(2), 423–464.

Gentzkow, M. (2006). Television and voter turnout. *Quarterly Journal of Economics*, *121*(3), 931–972.

Gerber, A. S., Karlan, D., & Bergan, D. (2009). Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics*, *1*(2), 35–52.

Gibson, N. S. (2020). Moving forward or backsliding: A causal inference analysis of the effects of the Shelby decision in North Carolina. *American Politics Research*, *48*(5), 649–662.

Glenn, B., Barrett, N., & Lightfoot, E. (2020). *The effects and implementation of restorative practices for discipline in New Orleans schools* (Technical Report). Education Research Alliance for New Orleans. Retrieved from https://educationresearchalliancenola.org/files/publications/20201218-Glenn-Barrett-Lightfoot-The-Effects-and-Implementation-of-Restorative-Practices-for-Discipline-in-New-Orleans-Schools.pdf

Gopalan, M., & Nelson, A. A. (2019). Understanding the racial discipline gap in schools. *AERA Open*, *5*(2).

Grimmer, J., Hersh, E., Meredith, M., Mummolo, J., & Nall, C. (2018). Obstacles to estimating voter ID laws' effect on turnout. *Journal of Politics*, *80*(3), 1045–1051.

Ho, D. E. (2017). Building an umbrella in a rainstorm: The new vote denial litigation since Shelby County. *Yale Law Journal Forum*, *127*, 799.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, *66*(1), 169-182.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, *125*(5), 576.

Husted, T. A., & Kenny, L. W. (1997). The effect of the expansion of the voting franchise on the size of government. *Journal of Political Economy*, *105*(1), 54–82.

Kinsler, J. (2011). Understanding the black–white school discipline gap. *Economics of Education Review*, *30*(6), 1370–1383.

Kinsler, J. (2013). School discipline: A source or salve for the racial achievement gap? *International Economic Review*, *54*(1), 355–383.

Kuriwaki, S. (2020). *Cumulative CCES Common Content (2006-2019)* (Dataset Version 5). Harvard Dataverse. Retrieved from https://doi.org/10.7910/DVN/II2DB6

Lacoe, J., & Steinberg, M. P. (2018). Rolling back zero tolerance: The effect of discipline policy reform on suspension usage and student outcomes. *Peabody Journal of Education*, *93*(2), 207–227.

Lindbeck, A., & Weibull, J. W. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, *52*(3), 273–297.

Liu, J., Hayes, M. S., & Gershenson, S. (2022). JUE Insight: From referrals to suspensions: New evidence on racial disparities in exclusionary discipline. *Journal of Urban Economics*. Retrieved from https://doi.org/10.1016/j.jue.2022.103453

Losen, D. J., Hodson, C. L., Keith II, M. A., Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap?* Los Angeles: The Center for Civil Rights Remedies, University of California.

McDonald, M. P. (2021). *CPS vote over-report and non-response bias correction* (blog post). United States Elections Project. Retrieved from http://www.electproject.org/home/voter-turnout/cps-methodology (accessed 5 January 2021)

Michelmore, K., & Dynarski, S. (2017). The gap within the gap: Using longitudinal data to understand income differences in educational outcomes. *AERA Open*, *3*(1).

Miller, G. (2008). Women's suffrage, political responsiveness, and child survival in American history. *Quarterly Journal of Economics*, *123*(3), 1287–1327.

Naidu, S. (2012). *Suffrage, schooling, and sorting in the post-bellum U.S. South* (NBER Working Paper No. 18129). National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w18129

Neiheisel, J. R., & Horner, R. (2019). Voter identification requirements and aggregate turnout in the U.S.: How campaigns offset the costs of turning out when voting is made more difficult. *Election Law Journal: Rules, Politics, and Policy*, *18*(3), 227–242.

Pope, N. G., & Zuo, G. W. (2020). *Suspending suspensions: The education production consequences of school suspension policies.* Retrieved from http://www.econweb.umd.edu/~pope/Suspensions_pope_zuo.pdf

Reykdal, C., Weaver-Randall, K., & Ireland, L. (2018). *Comprehensive Education Data and Research System (CEDARS) appendix manual* (Version 10.2). Washington State Office of Superintendent of Public Instruction. Retrieved from https://www.k12.wa.us/sites/default/files/public/cedars/pubdocs/2017-18cedarsappendices.pdf

Riker, W. H., & Ordeshook, P. C. (1968). A theory of the calculus of voting. *American Political Science Review*, *62*(1), 25–42.

Ritter, G. W., & Anderson, K. P. (2018). Examining disparities in student discipline: Mapping inequities from infractions to consequences. *Peabody Journal of Education*, *93*(2), 161–173.

Schuit, S., & Rogowski, J. C. (2017). Race, representation, and the Voting Rights Act. *American Journal of Political Science*, *61*(3), 513–526.

Shepherd, M. E., Fresh, A., Eubank, N., & Clinton, J. D. (2021). The politics of locating polling places: Race and partisanship in North Carolina election administration, 2008–2016. *Election Law Journal: Rules, Politics, and Policy*, *20*(2), 155–177.

Shi, Y., & Zhu, M. (2022). Equal time for equal crime? Racial bias in school discipline. *Economics of Education Review*, *88*. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0272775722000334

Smith, C. W., Kreitzer, R. J., & Suo, F. (2020). The dynamics of racial resentment across the 50 U.S. states. *Perspectives on Politics*, *18*(2), 527–538.

Sorensen, L. C., Bushway, S. D., & Gifford, E. J. (2022). Getting tough? The effects of discretionary principal discipline on student outcomes. *Education Finance and Policy*, *17*(2), 255–284.

Steinberg, M. P., & Lacoe, J. (2018). Reforming school discipline: School-level policy implementation and the consequences for suspended students and their peers. *American Journal of Education*, *125*(1), 29–77.

US Commission on Civil Rights. (2018). *An assessment of minority voting rights access in the United States* (Statutory Report). Retrieved from https://www.usccr.gov/pubs/2018/Minority_Voting_Access_2018.pdf

Valentino, N. A., & Neuner, F. G. (2017). Why the sky didn't fall: Mobilizing anger in reaction to voter ID laws. *Political Psychology*, *38*(2), 331–350.

Washington, E. (2006). How black candidates affect voter turnout. *Quarterly Journal of Economics*, *121*(3), 973–998.

Welsh, R. O., & Little, S. (2018). The school discipline dilemma: A comprehensive review of disparities and alternative approaches. *Review of Educational Research*, *88*(5), 752–794.