Crime and Corporal Punishment

Tabea Rudolf and Carmen Villa*

October, 2026

Abstract

This paper provides the first causal estimates of the effects of corporal punishment bans on crime. We exploit the staggered introduction of bans in U.S. public schools across 28 states between 1974 and 2019. The bans reduced juvenile arrest rates by nearly 20%, particularly among Black students and in states with high pre-ban corporal punishment rates. The decline in arrests extends across a broad range of offenses and is too large to be explained by directly punished students alone, suggesting important spillover effects. Our findings challenge deterrence-based justifications for physical discipline.

JEL codes: I21, I28, K42, H75

1 Introduction

Corporal punishment in schools is prevalent in many parts of the world and legal in 62 nations, including in 17 U.S. states (Gershoff, 2017; Heekes et al., 2022). In 2017/18, at least 70,000 pupils—disproportionately Black and disadvantaged—were corporally punished in American public schools (U.S. Department of Education, 2024). This contrasts against attitudes toward corporal punishment in other institutional settings. Most countries prohibit corporal punishment in prisons and juvenile detention, and only a few still allow courts to sentence adults to it.

Should corporal punishment in schools be banned where it is still legal? On the one hand, psychological and pediatric research has argued that children subjected to physical punishment are more likely to develop externalizing behavior, and that this

*Rudolf: Department of Economics at the University of Zurich, tabea.rudolf@econ.uzh.ch; Villa: Department of Economics and Jacobs Center for Productive Youth Development at the University of Zurich, and Institute for Fiscal Studies, carmen.villa@econ.uzh.ch. We are grateful to Anjali Adukia, Amanda Agan, Manuel Bagues, Ana Costa-Ramon, Zoe Cullen, Aureo de Paula, Jan Fasnacht, Hikaru Kawarazaki, Sandra McNally, Steve Mello, Zach Porreca, Imran Rasul, Ulf Zölitz and colleagues at the University of Zurich and the Institute for Fiscal Studies for helpful comments. We received very useful feedback at the EAYE Annual Meeting 2026, the 10th Workshop on the Economics of Organised Crime, the 5th Workshop on the Economics of Crime for Junior Scholars, the Women in Empirical Microeconomics Conference, and the FAIR research seminar at NHH. Ilina Logani provided excellent research assistance. Any errors are ours alone.

extends to later delinquency (Bandura, 1973; Dodge, 1980; Dodge et al., 1990; Widom, 1989). On the other hand, proponents of physical discipline might argue that the threat of punishment is necessary to maintain classroom order, as it could deter unwanted behaviors as in deterrence models à la Becker (1968), and facilitate learning (with potential returns in educational attainment and labor markets). The existing empirical literature documents correlations between experiences of punishment and crime (Gershoff, 2017; Gershoff, 2002), but disentangling causal effects is difficult due to potential selection and reverse causality. Understanding whether bans on corporal punishment might affect crime is ultimately an empirical question.

In this paper, we provide the first causal estimates of the effects of banning corporal punishment in schools on arrests. We exploit the staggered implementation of bans in public schools across 28 U.S. states between 1975 and 2011 in difference-in-differences designs, and leverage data spanning 1974 to 2019. After establishing that the bans effectively reduce corporal punishment use, we examine whether schools substitute toward other disciplinary measures, and whether banning physical punishment affects crime. We also assess effects in education and labor market outcomes.

Our results show that the bans abolish corporal punishment in schools, with rates declining by 89%. We do not find that schools increase suspensions as a substitute punishment. If anything, suspensions decline in the medium term, consistent with improved behavior in the classroom but not with sudden changes in punitive approaches.

Arrest rates at ages 10–17 fall by 1.2 percentage points after the bans (19% of the pre-ban arrest rate). Declines are larger among the groups most exposed to corporal punishment before the bans, namely Black pupils in states with high corporal punishment rates at baseline. In some areas, the effects persist into early adulthood (at ages 18–24). The magnitude is substantial, with only a few studies—often evaluating targeted interventions among disadvantaged groups, like summer jobs (Davis and Heller, 2020; Heller, 2014) and cognitive behavioral therapy (Heller et al., 2017)—documenting larger declines in arrests. The effect size is comparable to that of restorative justice approaches (Adukia et al., 2025), and larger than that of policies targeting educational attainment

(Bell et al., 2022; Hjalmarsson et al., 2015; Lochner and Moretti, 2004).[1] We estimate that the bans account for a fall of 4% in the all-age arrest rate between 1991 and 2001.[2]

We no not find any effects in educational attainment or in labor market outcomes up to age 30. This suggests that the declines in crime are not mediated by the bans leading young people to higher engagement with their studies (incapacitation). Similarly, we do not find evidence of the bans hampering attainment or labor market outcomes, rejecting that corporal punishment helps human capital acquisition.

We reject that the effects are confounded by increases in the dropout age or abortion laws (Bell et al., 2022; Levitt, 2004). The results are robust to controlling for contemporaneous changes in economic indicators and police force size. We do not find effects on adult arrest rates among cohorts unaffected by bans, ruling out state-level confounders affecting all cohorts. Triple-differences estimates using adults as a second comparison group confirm these results. The estimates are robust to a battery of other checks, including alternative samples, estimators, and controls.

To explore mechanisms, we assess whether the effects are driven by students' own experiences of corporal punishment (direct effects) or by impacts on the broader classroom (indirect effects). Because no data link individual experiences of corporal punishment to arrests, we cannot perfectly disentangle the two channels. Heterogeneity analyses by individual attributes show that students who were more likely to experience corporal punishment see larger arrest declines, consistent with own experiences increasing anti-social behavior and in line with the correlational literature in pediatrics and psychology. But, back-of-the-envelope calculations suggest that direct effects can not account for the full arrest decline: at most 6% of students were punished during their schooling pre-ban, and their arrest rates would have needed to fall by more than 300%. To provide a

[1] Summer jobs reduced violent-crime arrests by 33–43% among disadvantaged youth in Chicago, and a cognitive-behavioral program reduced total arrests by 28–35% (Heller et al., 2017). READI Chicago, which offered men at very high risk of gun violence an 18-month job with cognitive behavioral therapy, lowered arrests for shootings and homicides by 65% (Bhatt et al., 2024). Adukia et al. (2025) find that restorative-justice practices across Chicago public high schools reduced juvenile arrests by 18.8%, and Bacher-Hicks et al. (2024) document that a one-standard-deviation increase in middle-school strictness raised adult arrest probabilities by 15–20% in Charlotte-Mecklenburg schools in North Carolina. Bell et al. (2022) estimate that an additional year of compulsory schooling in the U.S. reduced juvenile arrests by about 6%. Similar magnitudes are found in other contexts (Anderson, 2014; Hjalmarsson et al., 2015; Lochner and Moretti, 2004).

[2] We follow the accounting that Levitt (2004) applies to reported crime. He attributes a 10% fall in crime to legalized abortion, 8–12% to the growth in prisons and 5.5% to more police.

more realistic bound, we assume that punished students were more likely to offend, using odds ratios of 1.5–2.0 from meta-analyses in the correlational literature on the corporal-punishment–offending link, and that the bans closed this gap. These suggest that own experiences account for at most 30% of the arrest decline. Therefore, at least 70% of the effect comes from students who would not have been physically punished themselves. Indirect effects might emerge from the threat of punishment, a less fair or more hostile classroom climate (Algan et al., 2013; Fissel et al., 2019; Way, 2011), witnessing violence by teachers (Bandura, 1977; Bandura et al., 1961), or peer networks (Billings et al., 2019; Billings and Hoekstra, 2023; Carneiro et al., forthcoming; Carrell et al., 2018; Carrell and Hoekstra, 2010; Kawarazaki, 2024; Le and Nguyen, 2019).

Heterogeneity analyses by crime type point at several non-competing psychological channels linking experiences of corporal punishment to externalizing behaviors. We observe a broad decline across crime types, consistent with punishment weakening the internalization of social norms and self-control (Gershoff, 2002; Gottfredson and Hirschi, 1990). The largest fall in arrests is in hard-drug possession (-79%), in line with physical punishment inducing trauma and coping through substance use (Dube et al., 2003; Khantzian, 1997). We also find large drops in disorderly conduct (-39%), assault (-33%), rape (-23%) and sex offenses (-18%), which support social-learning and cycle-of-violence theories whereby children who experience or witness violence learn that force is an acceptable response to conflict (Bandura, 1973; Lavy and Shayo, 2026; Widom, 1989).

Our paper contributes to three strands of the literature. First, we provide the first causal estimates of the effects of banning corporal punishment. Existing observational work documents correlations between experiences of physical discipline (both at home and in school) and worse later life outcomes, including poorer socio-emotional skill development and later criminal behavior (Cuartas, 2022; Cuartas et al., 2025; Gershoff and Grogan-Kaylor, 2016; Gershoff, 2002). However, comparisons of punished and non-punished children are likely to suffer from selection bias (Jaffee et al., 2004), even when researchers can control for family attributes (Currie and Tekin, 2012). From an experimental standpoint, two studies in Jamaica show that after positive parenting interventions, parents self-report lower use of corporal punishment and children show

better short-run socio-emotional outcomes (Dinarte-Diaz et al., 2026; Francis and Baker-Henningham, 2021). We identify effects from a policy change using data from 46 states, measure offending using arrest data, and follow cohorts from ages 10 to 29, providing short- and long-run estimates.

Second, we contribute to the literature on the causal effects of punitive approaches in schools. Previous work in specific U.S. school districts has focused on suspensions, showing that stricter disciplinary environments increase later arrests and that restorative justice approaches reduce them (Adukia et al., 2025; Bacher-Hicks et al., 2024). A key mechanism underlying the relationship between suspensions and crime, as well a the broader literature on the role of schooling concludes that incapacitation is a key mediating channel (Bell et al., 2022). We show that teacher-inflicted punishment is an additional determinant of criminal engagement. Forgoing physical discipline does not come at a cost to educational attainment, adding to evidence favoring less punitive disciplinary environments.

Third, we contribute to understanding the causal effects of experiences of violence, which previous work have studied in the context of crime victimization.[3] The existing literature has shown effects of direct victimization—and particularly of assault—on physical and mental health, birth outcomes, labor market outcomes and precautionary behavior (see Bindler et al., 2020, for a review).[4] Some papers have also documented spillovers to non-victims.[5] One open question is the relationship between victimization and the onset of criminal behavior, or the victim-offender overlap, for which there is descriptive support (see Jennings et al., 2012, for a review). We show that experiences of violence inflicted by teachers can lead to offending which extend beyond the children who are punished.

Our findings have implications for policy reform where corporal punishment in schools remains permitted, both in U.S. states and in other countries. We conclude that teacher-

[3]Corporal punishment inflicted by teachers can be prosecuted as assault where it is not legally permitted.

[4]Ornstein (2017) and Bindler and Ketel (2022) show detrimental effects of assault on physical health and mortality, and Cornaglia et al. (2014), Mahuteau and Zhu (2016) and Johnston et al. (2018) on mental health and well-being.

[5]Exposure to crime in the neighborhood or near school affects residents' mental health (Cornaglia et al., 2014; Dustmann and Fasani, 2016), birth outcomes (Foureaux Koppensteiner and Manacorda, 2016), students' test scores, attendance and dropout (Foureaux Koppensteiner and Menezes, 2021; Monteiro and Rocha, 2017), economic mobility in adulthood (Sharkey and Torrats-Espinosa, 2017) and the decision to move (Xie and McDowall, 2008).

inflicted corporal punishment is not an effective educational tool. Although the effects of corporal punishment might differ depending on who administers and how, these findings may also inform efforts to limit physical discipline beyond schools. This is relevant in a context where parental use of corporal punishment is prevalent (Egyir et al., 2026), legal in most of the world, and where norms around its acceptability are divided.[6]

2 Institutional background

This section describes the prevalence of corporal punishment in U.S. schools and the implementation of corporal punishment bans. We draw on publicly available information and insights from a survey administered to U.S. school teachers (see Appendix B).

2.1 Corporal punishment prevalence in U.S. schools

Corporal punishment in U.S. schools refers to "paddling, spanking, or other forms of physical punishment" imposed on students as a disciplinary measure (U.S. Department of Education, 2024). It typically involves a member of the school staff striking a student on the buttocks with a flat wooden paddle. Documented triggers include tardiness, talking in class, dress-code violations, playing with a pen, chewing gum, or "disrespect" (Human Rights Watch and American Civil Liberties Union, 2008).

The practice was once permitted across the U.S. and 63.6% of teachers in our survey report having experienced or witnessed physical discipline back when they were pupils. In 1976, the first year where recorded data exist, it was lawful in every state but four, and public schools reported close to 1.5 million students punished during the school year, or 3.5% of enrolled pupils (U.S. Department of Education, 2024). Use was concentrated in Southern states, which accounted for a third of enrollment but two thirds of paddling. The highest rates were 11.8% of students in Florida, 10.6% in Mississippi, 10.0% in Georgia and 9.5% in Texas. Among teachers in our survey whose school permitted corporal punishment for at least part of their career, 54.3% (19 of 35) state that they or their colleagues used it.

[6] In the U.S. in 2024, 54% of respondents to the General Social Survey agreed with the statement that child discipline through spanking was sometimes necessary (Davern et al., 2025). Half of the respondents to the World Values Survey in 2018 stated that under some circumstances it is acceptable for a parent to beat their child; with most of the variation in attitudes existing within countries.

This figure is 75% (9 of 12) among teachers whose school permitted corporal punishment throughout their entire career.

The use of physical discipline declined steadily in the 1980s and 1990s (Figure A1), in part as many bans on its use in public schools were introduced across school districts and states, as we describe in the next section. By the turn of the millennium the rate of corporal punishment was 2%. The decline continued in the 2000s and 2010s. As of 2023 it affects ∼3 in 1,000 students enrolled in 17 states where it is still permitted. Four states (Mississippi, Texas, Alabama and Kansas) account for about 80% of the students punished. A map of the states by ban status at the time of writing is in Figure A2.

Corporal punishment is disproportionately used on Black students (15% of the national student population but 37% of those paddled), boys (80% of those paddled) and students with learning difficulties (Gershoff and Font, 2016; U.S. Department of Education, 2024). Poorer, more rural counties also have higher rates of corporal punishment (Figure A3).

2.2 Corporal punishment bans

The first state to ban corporal punishment in public schools was New Jersey, which did so in 1867. The ban reflected parents' distrust of teachers (Raichle, 1974), but the reform was heavily contested at the time. Teachers across the state warned that classroom order could not be maintained without physical discipline. The first years after the ban were met with several attempts to re-instill corporal punishment rights. Legislative action in other states did not begin until over a century later.[7]

In 1977, the issue reached national attention in *Ingraham v. Wright*. James Ingraham, a 14-year-old student, was paddled for being slow to respond to a teacher, required medical treatment, missed eleven days of school, and developed a hematoma. He sued the school, arguing the punishment violated the Eighth Amendment. By a 5-to-4 margin, the Supreme Court ruled that corporal punishment in schools does not constitute cruel and unusual punishment. In turn, regulation was left to each state (Supreme Court of the United States, 1977).

The ruling catalyzed opposition. The District of Columbia and Rhode Island banned

[7] Massachusetts banned the practice in 1971, Hawaii in 1973, and Maine in 1975.

the practice in 1977, and a further 26 states enacted bans between 1980 and 2019, with the majority of the bans (14 states) taking place in the 1980s (Figure A1). Colorado and Idaho enacted bans in 2023, although some of those recent bans have arrived onto states where CP was effectively zero for several years. This is in part due to some school districts voluntarily banning corporal punishment before their state enacted legislation.[8] Only two states extended the ban to private schools, where students may have continued to experience corporal punishment. Figure A4 shows the evolution of the number of states with and without bans from 1970 to 2025.

The legislative process behind these bans was typically protracted. Bills were often introduced multiple times before passage, sometimes over a decade apart. The bills were sometimes driven by advocacy from civil liberties and pediatric organizations, or incited by public outcries following particularly salient cases. Both Democratic and Republican states passed bans, often in bipartisan coalitions.

States that banned corporal punishment tend to be richer than states where it is still legal. Comparing states banning in the 1980s with the never banned states shows fewer Black residents in the banning states (6.6% against 17.6%), fewer adults without a high-school diploma (32.7% against 43.1%) and fewer households receiving food stamps (6.3% against 9.8%) — see Appendix Table A3. Juvenile offense rates were not significantly different. States banning corporal punishment in the 1990s also tend to be richer than the never treated states.

In states where the ban is in place, corporal punishment can be prosecuted as assault or child abuse. There were a few legal disputes relating to corporal punishment following the bans, but we did not find evidence of systematic ban violations.

[8]District policies are set at the school board level, change over time and are not centrally compiled. Kentucky and North Carolina have banned the use of corporal punishment de facto, as the last school district that still permitted corporal punishment in Kentucky adopted a policy banning it in 2023; and in North Carolina in 2018. Ohio first introduced a partial ban in 1994, allowing districts to opt-out. By the time the full ban was instilled, 97 of the 100 largest districts had already abolished corporal punishment.

3 Data

This section details the various data sources, most of which are publicly available. Summary statistics for our main outcomes are in Table A2.

Corporal Punishment Bans. Information on bans on corporal punishment in schools is from Global Initiative to End All Corporal Punishment of Children (2024) and is available at the state-level.[9] We use this source to denote treatment timing, so as to compare pupils in states where bans were enacted to pupils in areas which permit corporal punishment, before and after the bans. We denote a state as having implemented a ban if it did so for the entire public school student body. For Ohio, which introduced a partial (opt-in) ban in 1994 and a full ban in 2009, we use the year of the first ban. Table A1 presents the ban year by state.

Corporal Punishment Rates. We use the Civil Rights Data Collection (CRDC) to obtain proxies on punishment rates in schools at the county and year level. The CRDC is a mandatory biennial survey of public schools in grades K–12 (ages 5 to 18) run by the Office for Civil Rights of the U.S. Department of Education since 1976. The CRDC samples school districts, and every school in that district completes the survey. Each wave covers between 3,000 and 5,800 districts and between 35,000 and 52,000 schools, located in 1,500 to 2,400 counties (about 60% of U.S. counties on average).[10]

Our main variable of interest is the corporal punishment rate computed as the number of students who received corporal punishment at least once during the school year per 100 enrolled students. Schools report students, not incidents, so the rate measures the extensive margin of corporal punishment.[11] We also compute suspension rates to assess

[9] We were not able to obtain information on the timing of bans at the school-district level. This implies that in some states corporal punishment may have already been banned de-facto in a state if it was banned in a majority of school districts.

[10] Data for 1984 are not available. The CRDC continues after 1994, but the 1997, 1998, and 2009–10 public files do not include sampling weights, so we cannot produce representative estimates. For 2000 to 2006, the files available to us report corporal punishment only as state-level projections.

[11] The CRDC did not record the number of incidents until the 2023–24 wave. A limitation of the CRDC is that the data are self-reported by schools. Misreporting of corporal punishment is possible, particularly underreporting in states where bans were already enacted, although we did not find evidence suggesting misreporting.

substitution toward other punitive approaches, and as a proxy for misbehavior in the classroom. Suspensions are the number of students suspended from school for at least one full school day during the previous school year, counting each student once regardless of the number of suspensions. Schools report corporal punishment and suspensions for all enrolled students, without a breakdown by age.

We assign school districts to counties and aggregate the data at the county-year level. Since multiple schools operate within a district and multiple districts within a county, we use school and district sampling weights to ensure representative coverage. Sampling weights are only available from 1976–1994, our effective sample.

Arrest Rates. To measure delinquency, we use arrest data from the U.S. Federal Bureau of Investigation Uniform Crime Reports (UCR) Program covering 1974 to 2019. The UCR is a nationwide cooperative initiative in which law enforcement agencies voluntarily report their annual arrests. The count of arrests is available for each agency by age, sex, race and offense type. Data on arrests by sex and crime type exist disaggregated by specific age up to age 24, and by five year age bands at ages 25 and older. Ethnicity-disaggregated arrests only exist by crime type and broad age group (juvenile or adult). A unique individual may be arrested more than once. When a person is arrested on several charges, the agency records only the most serious offense.

We aggregate agency-level records to the county year level and match them with population data from the Surveillance, Epidemiology, and End Results (SEER*Stat) Program (National Cancer Institute, 2025) to compute arrest rates.[12] A caveat of the UCR is there are substantial gaps in data reported across agencies. This motivates our use of county instead of state-level data. We also produce a separate dataset keeping only the counties that report arrests in every year of their event window (350 counties), which we use in a robustness check.

[12]The UCR only has aggregate population counts while the SEER data allow to compute age-, gender- or race-specific rates. Since the UCR coverage is incomplete, age or race-specific arrest rates using official SEER population counts would under-estimate true crime incidence. We keep county-years in which reporting agencies cover more than 95% of the population, so that missing agencies do not show up as falls in crime. Following Bell et al. (2022), we construct a yearly covering ratio of the UCR by comparing agency-level aggregate populations from the UCR with official aggregate SEER estimates. We use this ratio to re-weight the age or race-specific population when constructing county-level arrest rates.

Our main dependent variable is the arrest rate at ages 10–17. In our preferred estimates, we winsorize the right tail at the 95th percentile to mitigate the influence of outliers.[13] From 1974 to 2019, arrest rates at ages 10–17 were 5.3 per 100 (7.7 for males). We calculate the arrest rate at ages 18–24 in a county-year to measure medium term effects, and arrest rates at ages 40–64 as a placebo.

We compute total crime rates, and rates by gender, ethnicity, age and crime type to assess heterogeneous effects, and to evaluate plausible mechanisms. We define violent crimes as murder, rape, robbery, aggravated assault, negligent manslaughter and simple assault following Bell et al., 2022.[14] Other crime definitions use the UCR classification.

Education and Labor-Market Outcomes. We measure long-run education and labor-market outcomes using the 5% samples of the 1990 and 2000 Decennial Census and the American Community Survey (ACS) for 2005–2024 (Ruggles et al., 2025). We focus on educational attainment and labor market prospects at ages 25–29, when most individuals are likely to have finished studying. This covers birth cohorts 1961 to 1999. For conciseness, we present estimates on total years of schooling and employment status (which includes self-employment) in the main results. Additional outcomes include high-school completion, attaining a bachelor's degree or more, and log real wages, and are discussed in the Appendix.

We construct a county by birth-year panel pooling all survey years in which each cohort is observed at ages 25–29. The county is the county of residence, and IPUMS identifies it only for counties large enough to be separated in the public-use geography (550 counties, home to 56% of U.S.-born adults aged 25–29). We assign exposure to a ban using state and year of birth. We restrict this analysis to U.S.-born individuals who live in their state of birth (68% of the sample). For them, the state of birth is more likely to be the state where they attended school, reducing error in treatment status.

Survey among U.S. school teachers We conducted an online survey among teachers to learn about the perceptions around the use of corporal punishment and its effectiveness.

[13] We show results winsorizing at the 99th percentile and on the raw arrest rate in robustness checks.
[14] The UCR definition for violent crime does not include negligent manslaughter and simple assault.

We recruited participants on Prolific and sampled respondents who are or had been K-12 teachers in the U.S. The sample included 162 individuals. Details are in Appendix B.

Additional Data. Some analyses include economic and political control variables. The National Welfare Data of University of Kentucky Center for Poverty Research (2026) provides state-level SNAP recipients from 1980 to 2020, GDP per capita and Senate and House partisanship; state unemployment rates come from U.S. Bureau of Labor Statistics (2026). County characteristics are measured in the 1980 census: population size, racial composition, and urbanization from IPUMS NHGIS (2026), and educational attainment from USDA ERS (2025). To control for the size of the county-level police force, we use the Law Enforcement Employees Data (FBI Crime Data Explorer, 2026).

4 Empirical Strategy

Estimating the causal effects of corporal punishment is empirically challenging. Correlations between experiences of punishment and outcomes could reflect selection or reverse causality, as children with behavioral difficulties might be more likely to be punished, and the causal link could run from behavior to punishment, rather than punishment to behavior. Meanwhile, there are evident ethical barriers to exploring the question experimentally.

We use the staggered adoption of statewide bans on corporal punishment in schools as a shock to exposure, and compare cohorts whose schooling fell under a ban to those whose schooling did not. We present first-stage evidence of the effects on the use of corporal punishment, and reduced form estimates on the effects on years of schooling, employment rates, and arrests along the life-cycle. Corporal punishment and suspension rates (every two years) and arrest rates (every year) are recorded by county and year for all pupils aged 5 to 18. Each county therefore contributes multiple observations, with the cohorts of interest progressively cycling through the observed age band. Educational and labor-market outcomes are observed once per cohort at a single age (a cross-section at the county-cohort level).

Our parameter of interest is the average treatment effect of having spent at least one school year under a ban on treated cohorts (ATT). Let ban_s denote the year in which state s adopted its ban ($\text{ban}_s = \infty$ for never-treated states), and define the first treated birth cohort as $g_s = \text{ban}_s - 18$. Cohorts born in g_s or later were aged 17 or younger when the ban took effect. Treatment intensity rises with the share of school years spent under the ban, from a single year for cohort g_s to all years for cohorts $c \geq g_s + 12$. We refer to pupils aged 6 to 17 at ban as partly treated and to those aged 5 or younger as fully. We present estimates across the whole student population, and present heterogeneity analyses by pupil attributes.[15]

To provide first-stage evidence on the effects on corporal punishment and suspension rates, we estimate stacked models (Cengiz et al., 2019; Deshpande and Mueller-Smith, 2022; Wooldridge, 2023) using Poisson quasi-maximum likelihood (PQMLE) given that the outcome at the county-level is a rare non-negative event. We build one stack per ban cohort, with never-treated and not-yet-treated counties as controls, and include stack-by-county and stack-by-year fixed effects. Estimates are weighted by enrollment, and standard errors are clustered at the state level. Because the data are biennial, the reference period in the event studies is two years before the ban.

To provide reduced-form evidence on the effects on education, employment and arrest rates, we use the staggered difference in differences estimator developed by Callaway and Sant'Anna, 2021:

$$Y_{ip} = \alpha_i + \lambda_p + \sum_{g \neq \infty} \sum_{p \geq g} \text{ATT}(g, p-g) \cdot \mathbb{1}\{G_i = g\} \cdot \mathbb{1}\{P = p\} + \varepsilon_{ip}, \tag{1}$$

where Y is an outcome of interest, α_i are county fixed effects, λ_p are period fixed effects, p is the analytic period and G_i is the first treated period in county i. Period p takes one of two interpretations. For *cross-sectional* outcomes each cohort contributes one observation per county; $P = c$ (birth year) and $G_i = g_s = \text{ban}_s - 18$. For *panel* outcomes each county contributes multiple observations as cohorts cycle through the relevant age range; $P = t$ (calendar year) and $G_i = \text{ban}_s$.

We use not-yet-treated states as the comparison group; compute standard errors with

[15] Because we do not have information by month of birth, age at ban has an error of up to one year.

a wild bootstrap clustered at the state level (Bertrand et al., 2004); and weight by cell population (the population of the age band for arrests, and of the county-cohort for education and employment).[16] All estimates use an event window from -10 to $+20$.[17]

To interpret the estimated ATT as causal we rely on two assumptions. The first is parallel trends. We assume that in the absence of treatment, outcomes in cohorts (or county-years) first treated at g would have evolved similarly to those in not-yet-treated cohorts. For the PQMLE specifications, the assumption applies to ratios of expected outcomes across periods rather than to levels (Wooldridge, 2023). The second is no anticipation. We assume potential outcomes are unaffected by treatment in pre-treatment periods. We present event studies in support of both assumptions and discuss how potential violations of these assumptions affect the estimates.

Our estimators are robust to heterogeneous treatment effects across cohorts and avoid the "forbidden comparisons" that can bias two-way fixed effects estimates (Baker et al., 2026; Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021). This is pertinent in our context, since about 60% of the states in our samples eventually ban corporal punishment (26 of 45 states in the corporal punishment data and 28 of 47 in the arrest data). States that banned before the first year observed in each dataset are "always treated" and excluded: Hawaii, Massachusetts and New Jersey, and in the corporal punishment data also the District of Columbia, Maine and Rhode Island.

5 Findings

In this section we present the estimates for our main outcomes of interest: corporal punishment rates, suspension rates, arrest rates during childhood, arrest rates during early adulthood, years of schooling and employment rates. We first present estimates for all pupils and then discuss heterogeneous effects by pupil attributes in section **??**.

The bans appear to nearly abolish reported use of corporal punishment in schools (Table 1, column 1). Rates fell by 89% relative to the pre-ban mean of 0.52 students

[16] Counties differ largely in size, with the largest 143 counties (5%) accounting for about 50% of the population. Unweighted estimates, which give equal importance to each county, are in the Appendix.

[17] Our results are robust to different choices on control groups and event windows.

Table 1: Effects of Corporal Punishment Bans

	(1)	(2)	(3)	(4)	(5)	(6)
	School discipline		Crime		Education and labor market	
	CP rate (K–12)	Suspensions (K–12)	Arrests (10–17)	Arrests (18–24)	Yrs schooling (25–29)	Employed (25–29)
ATT	-2.238***	-0.143**	-1.227***	-1.690*	-0.014	-0.003
	(0.290)	(0.062)	(0.324)	(0.948)	(0.032)	(0.009)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.022	0.000	0.075	0.657	0.756
Magnitude (%)	-89.33	-13.29	-19.04	-15.47	-0.11	-0.36
Mean (pre-ban)	0.522	5.333	6.445	10.926	13.315	0.771
N	132,766	141,558	85,344	85,341	8,383	8,383

Notes: Columns 1–2 report stacked PQMLE estimates for corporal punishment and suspension rates, weighted by enrollment. Columns 3–6 report Callaway and Sant'Anna (2021) estimates (CS DiD). Columns 3–4 use the county-year design (p = calendar year, G = ban year) for arrest rates at ages 10–17 and 18–24, weighted by age-band population and winsorized at the 95th percentile. Columns 5–6 use the cohort design (p = birth year, G = ban year -18) for years of schooling and employment at ages 25–29, weighted by cell population. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. "Mean" is the pre-ban mean of the dependent variable in treated units. "Magnitude" is the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta})-1)\times 100$ in columns 1–2 and as the ATT divided by the mean in columns 3–6. Standard errors, clustered at the state level, are in parentheses (analytic in columns 1–2, wild cluster bootstrap in columns 3–6). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

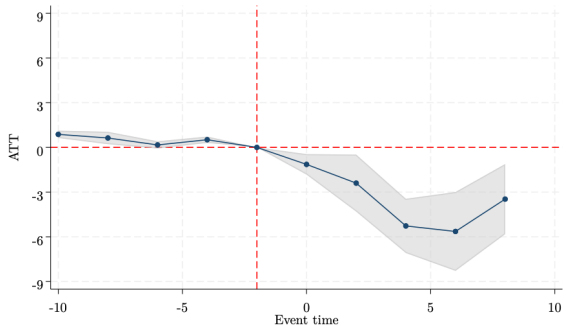
per 100 enrolled ($p < 0.001$). The event study (Figure 1, Panel a) shows that pre-trends are positive, but very small in magnitude, suggesting that corporal punishment use was slightly in decline before the state ban (and consistent with some school districts banning corporal punishment before the state-wide regulation). The event study shows a very sharp drop at ban and a decline that persists up to 8 years, the longest horizon we can study for this outcome. The findings support a first-stage effect of the bans on corporal punishment.

We find some declines in school suspensions (13%, p-value 0.02). Trends are indistinguishable from zero from t-10 up to t+6. Estimates become statistically significant 8 years after the ban. The event study doesn't suggest a sudden decline in teacher use of suspensions due to the bans. Instead, progressive drops are consistent with improved behavior in the classroom, particularly among cohorts who spent more years of schooling in a corporal punishment free environment.

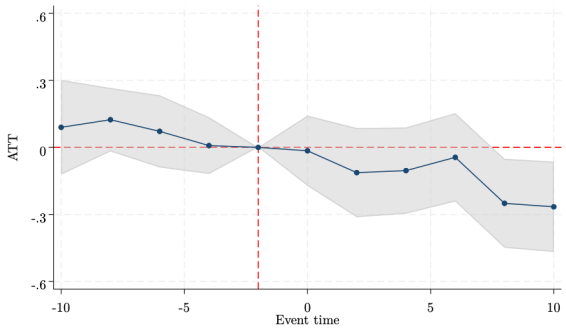
We find large and significant declines on juvenile arrests. At ages 10–17, the arrest rate fell by 1.2 percentage points, a magnitude of 19% of the pre-ban mean of 6.4 per

Figure 1: Event-studies of the Effects of Corporal Punishment Bans

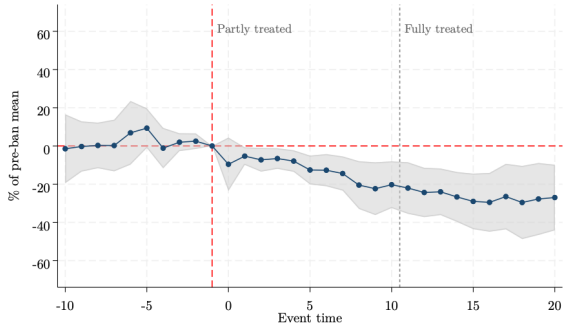
(a) Corporal punishment rate (5–18)



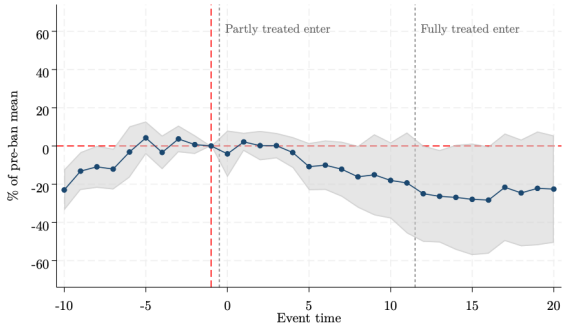
(b) Suspension rate (5–18)



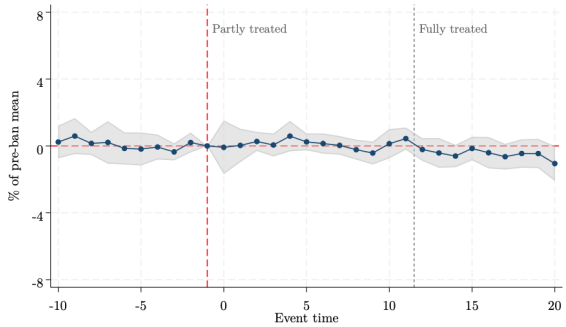
(c) Arrest (10–17)



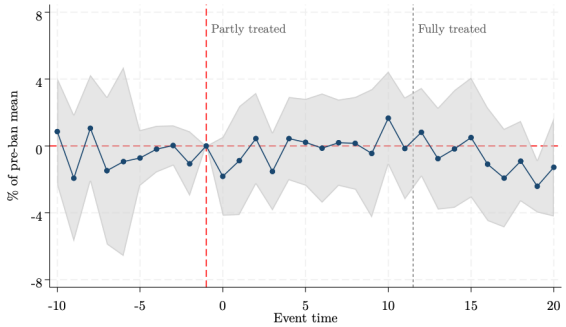
(d) Arrest (18–24)



(e) Years of schooling (25–29)



(f) Employed (25–29)



Notes: Event-study estimates of the effects of corporal punishment bans on the six main outcomes. Panels (a)–(b) are stacked PQMLE estimates for the corporal punishment and suspension rates, in log points; the reference period is -2 because the data are biennial. Panels (c)–(d) are Callaway and Sant'Anna (2021) estimates on the county-year design (p = calendar year, G = ban year) for arrest rates at ages 10–17 and 18–24; panels (e)–(f) are Callaway–Sant'Anna estimates on the cohort design (p = birth year, G = ban year $-$ 18, the first cohort aged 18 or younger in the year of the ban). Panels (c)–(f) are expressed as a percentage of the pre-ban mean in treated units. The dotted line marks the event time from which all cohorts observed went through school entirely under the ban: event time 11 in panel (c) and event time 12 in panels (e) and (f). In panel (d), the dotted lines mark when the first partly treated cohorts (event time 0) and the first fully treated cohorts (event time 12) reach age 18. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by enrollment (panels a–b), by age-band population (panels c–d) and by cell population (panels e–f). Arrest rates are winsorized at the 95th percentile. Shaded areas are 95% confidence intervals from standard errors clustered at the state level: analytic in panels (a)–(b) and a wild cluster bootstrap in panels (c)–(f).

100 residents ($p < 0.001$). The event study (Figure 1 panel c) shows pre-trends centered at zero, an immediate decline, and a persistent effect that strengthens over time at least until 20 years after the bans. The effects grow as successive cohorts spent a larger fraction of their juvenile years under the ban and fully-treated cohorts make up a growing share of this age group.

We also estimate whether effects on arrests persist up to ages 18–24. Since cohorts progressively move through the age band, estimates at longer lags involve cohorts who would have spent more years under the ban. There is suggestive evidence of a decline, with fewer 1.7 arrests per 100 residents, 15% of the pre-ban mean of 10.9. This effect is however, not statistically significant at conventional levels ($p = 0.07$; Table 1, column 4). The event study (Figure 1, panel d) shows no effect in the first four years after the bans, when the treated cohorts in this age band had spent at most four school years under a ban, and a steady decline afterwards, reaching 25 to 28% of the pre-ban mean 12 to 16 years after the bans. Sustained effects may be concentrated in a few cohorts and areas.

How much of the decline in arrests in the U.S. can the bans account for? We compare observed changes in arrest rates with those predicted in the absence of bans following the accounting that Levitt (2004) applies to reported crime. Between 1991 and 2001, the bans account for a 3.4–4.4% fall in the all-age arrest rate, which declined by 8%, and an 8.6–10.2% fall in the arrest rate at ages 10–17, which declined by 12%.[18]

We do not find that the effects on arrests are mediated by improved educational and labor market outcomes. We find a null effect on years of schooling (Table 1, column 5). The 95% confidence interval ranges from −0.076 to +0.048 years of schooling, roughly −4 to +2.5 weeks (economically negligible).[19] The event study is centered at zero, with small negative coefficients (at most 0.14 years, 20 years after the ban) for the fully treated cohorts. Alternative education indicators show mixed but similarly small effects: high-school completion falls by 1.4% ($p = 0.075$) but attaining a bachelor's degree or above rises by 2.6% ($p = 0.11$) (Appendix Table A4, columns 1 and 2). The estimate on employment

[18] The lower and upper bounds assume that the estimated effects remain constant as a number of arrests per 100 residents or as a share of the arrest rate, respectively. The two coincide when arrest rates are close to their pre-ban level and diverge as arrests fall. Effects more than 20 years after a ban are held at the year-20 estimate, and arrests at ages 25 and over are assumed unaffected.

[19] Assuming a return of 10% per additional year of schooling (Card, 1999; Gunderson and Oreopoulos, 2020; Psacharopoulos and Patrinos, 2018), a loss of 0.076 years would reduce earnings by less than 1%.

rates at ages 25–29 (Table 1, column 6) is null, with a confidence interval of −2.7 to +1.9% of the pre-ban mean. The event study is centered at zero. We do not find changes in log wages either (Appendix Table A4, column 3).

5.1 Heterogeneity

Where data permit, we explore heterogeneity in effects by gender, race, by fully or partly treated, and by high or low corporal punishment at baseline.[20] Full or partly treated is defined using the age at ban. Individuals are partly treated if they were 7 to 18 at the ban, and fully if they were 6 or younger. High corporal punishment states have above-median rates prior to a ban.[21]

Effects on corporal punishment are similar across gender and ethnicity, but are more precisely estimated in high-pre ban states.[22] The effects on suspension rates are statistically indistinguishable across subgroups (Figure 2, panel b).

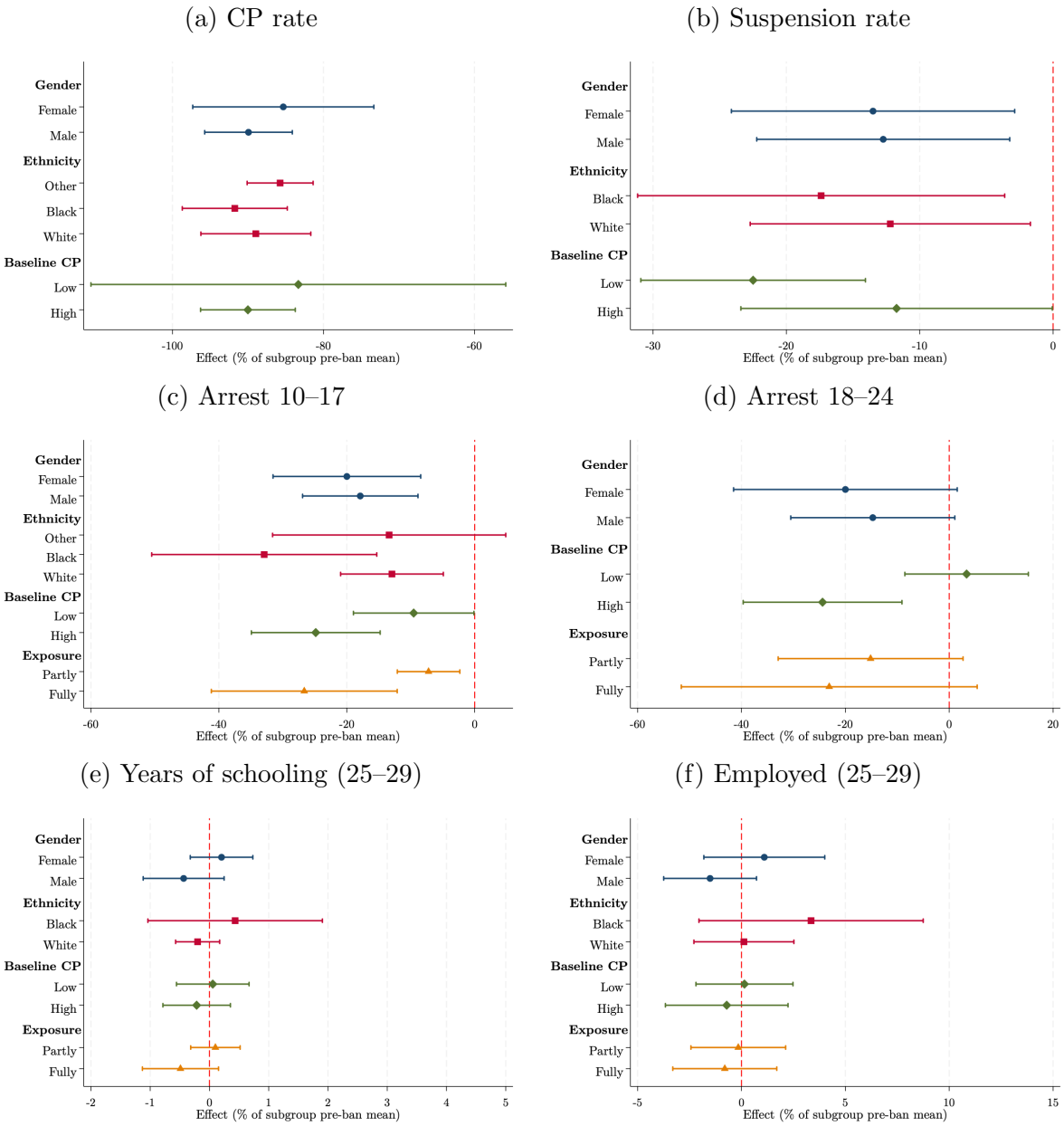
Heterogeneous effects in arrests are consistent with groups with higher exposure to corporal punishment experiencing larger declines (Figure 2, panels c–d). Juvenile arrests at ages 10–17 fall across all ethnic groups. The decline for Black youth (−33%, $p < 0.001$) is more than twice that for White youth (−13%, $p = 0.002$), and at baseline Black pupils faced corporal punishment rates 1.5 times higher than White pupils. The estimate for other ethnic groups is also negative but imprecise (−13%, $p = 0.15$). The decline is statistically significantly larger in states with high pre-ban corporal punishment (−25%, $p < 0.001$) than in low-punishment states (−10%, $p = 0.047$). The effect is also larger for fully exposed cohorts (−27%, $p < 0.001$) than for partially exposed (−7%, $p = 0.004$). Combining the three dimensions supports the interpretation that the bans reduce arrests through reduced exposure to corporal punishment (Appendix Figure A7). The decline is largest for Black youth in high-baseline states who were fully treated (−54%, $p < 0.001$).

[20] We cannot compute adult arrest by race, since the UCR data do not have cohort-level information by ethnicity. We have little power to assess employment and education outcomes for people not categorized as either "Black" or "White" and thus opt not to show uninformative null coefficients on the "Other" ethnicity variable, although they can be provided on request.

[21] High corporal punishment states are Alaska, California, Delaware, Illinois, Michigan, Montana, Nebraska, Nevada, New Mexico, Ohio, Oregon, Virginia, Washington, and West Virginia

[22] Because the corporal punishment data end in 1994 and group all pupils aged 5 to 18, we can only assess effects on partly treated cohorts. At no point in the post-period are all pupils in the age band fully treated (aged 6 or younger in the year of the ban).

Figure 2: Heterogeneity in the Effects of Corporal Punishment Bans



Notes: Heterogeneity in the effects of corporal punishment bans by gender, ethnicity, baseline corporal punishment and exposure. Baseline compares states above and below the median pre-ban corporal punishment rate among ever-ban states. Exposure averages the event-study coefficients over event times at which all cohorts in the outcome were partly treated (aged 7 to 18 in the year of the ban) or fully treated (aged 6 or younger). These are event times 0–3 and 11–20 in panel (c), 6–11 and 18–20 in panel (d), and 0–11 and 12–20 in panels (e)–(f). Panels (a)–(b) have no exposure rows because the CRDC data end in 1994, before all enrolled cohorts are fully treated. "Other" (Asian and Native American, about 3% of the population) is dropped from panels (b), (e) and (f) as too imprecise. UCR records Hispanic offenders as White. Panel (d) has no ethnicity rows because UCR records race by age only for the 18+ aggregate. Estimators, designs, weights, comparison group, event window and standard errors are as in the corresponding columns of Table 1. Estimates are expressed as a percentage of the subgroup pre-ban mean in treated units. Markers are point estimates and bars are 95% confidence intervals.

The same gradient by baseline intensity appears in arrests at ages 18–24, with the decline being driven byhigh-baseline-corporal-punishment states (-25%, $p = 0.002$). The estimates for fully (-23%, $p = 0.11$) and partly treated cohorts (-15%, $p = 0.10$) are similar. This age band contains only fully treated cohorts at event times 18 to 20.

We do not find meaningful heterogeneity in years of schooling (Figure 2, panel e). No estimate is statistically significant at the 10% level. For white pupils the estimate is very small and negative (-0.2%, $p = 0.29$), and its confidence interval rules out declines beyond 0.08 years, roughly one month of schooling. Even accepting a decline of that size as a consequence of the bans, it would not be a reasonable mediator for the observed arrest declines based on existing estimates on the relationship between education and deterrence. Black pupils might experience small gains, though these are not statistically significant and amount to at most 1.9% (about three months). We also find no meaningful heterogeneity in employment (Figure 2, panel f).[23]

6 Robustness

We assess the robustness of our main findings to confounding policies, different estimators, samples, different assumptions over outliers, controls groups, and control variables. The corporal punishment first stage is invariant across specifications. The juvenile arrest decline remains negative and statistically significant throughout. The null effects on years of schooling and employment are equally robust. The sustained effect on arrests later in life is not robust to changes in sample, and it is only significant where high-pre ban corporal punishment states are in the sample.

Alternative treatment timing The main results estimate the average effect of the bans, and we can show a first-stage effect for bans implemented up to 1994. Our key estimand of interest is however the effect of corporal punishment on outcomes. If bans enacted after 1994 arrived once social norms around the use of physical punishment had already changed, or if de-facto most school districts had already enacted bans, our

[23] Using other proxies for educational attainment, such as high-school completion, attaining a bachelors' degree or more, and yearly earnings show a similar pattern (Figure A6).

treatment timing might contain noise and attenuate our reduced form coefficients. We do a robustness check where we use as treatment timing the first year in which sustained zeroes in corporal punishment data at the state level appear in the CRDC. In four states those sustained zeroes arrive earlier than the ban.[24] The estimated effects on arrests are only slightly larger in magnitude but within the confidence interval of the main estimates: arrests at ages 10–17 fall by 20% ($p < 0.001$) and at ages 18–24 by 16% ($p = 0.063$) (Table A5). Estimates on education and labor markets remain null.

Placebo and triple differences estimates on arrests Our main result are declines in arrests, particularly among juveniles. To assess whether these effects might be confounded by state-level policy changes —particularly policies affecting crime incentives— we estimate effects on arrests of adults aged 40 to 64, who were at least 19 years old when the bans were enacted. We find that arrest rates did not change for those older adults after the bans (Table A6, column 1). The estimate is 0.01 arrests per 100 residents (0.6% of the pre-ban mean, $p = 0.92$). Using adults aged 40–64 as a second control group in a triple-differences design, which allows us to control for state-year, state-age group, and year-age group fixed effects, confirms a differential decline in arrests among age groups affected by bans. Arrests at ages 10–17 fall by 1.26 per 100 relative to adults aged 40 to 64 (20% of the pre-ban mean, $p < 0.001$). The decline at ages 18–24 is also similar to the main estimate, at 16% ($p = 0.049$; column 3).

Confounding policies. We test whether the effects are confounded by other well-researched policies shown to have influenced criminal behavior among youth. First, we assess whether changes in the minimum dropout age (as studied by Bell et al., 2022) play a part. Some reassuring evidence is the two policies are not perfectly correlated; of the 28 treated states in our sample, 14 changed the dropout age and 14 did not.[25] Among the 19 never-treated states, 10 changed the dropout age and 9 did not. We run a robustness check re-estimating the effects of the corporal punishment bans on the 22

[24] Those are Colorado, Idaho, Kansas, and Wyoming, who reach sustained zero corporal punishment in 2013 or 2017.

[25] On average, the change in dropout ages in states that implemented corporal punishment bans arrived about 9 years after the bans. Only New Mexico changed the dropout age *before* the CP ban.

states that never changed their dropout age during our sample period (13 treated and 9 never-treated). These states are not representative, with lower corporal punishment rates at baseline. We find a 77% decline in corporal punishment rates (smaller than the main result), and a drop on suspensions of similar magnitude as in the main estimates (−15%, $p = 0.058$). The fall in juvenile arrests is substantial, at around 14% ($p = 0.001$), and there are null effects on years of schooling and employment (Table A7, Panel A). The estimated effect on arrests at ages 18–24 instead becomes indistinguishable from zero.

Second, we assess whether the drop in juvenile arrests could be confounded by early exposure to legalized abortion. Five states liberalized abortion access in 1969–1970, before national legislation in 1973, and existing research shows that these changes explain large crime drops in the 1990s (Donohue and Levitt, 2001). We reject that the effects of corporal punishment bans are confounded by this policy in a robustness check where we exclude the early-exposure states.[26] The first-stage effect on corporal punishment rates remains unchanged, the juvenile arrest decline remains negative (−15%, $p = 0.03$) and within the 95% confidence interval of the main estimates, as does the estimated effect on suspensions (−16%, $p = 0.017$). The effects on schooling and employment are unchanged (Table A7, Panel B). The estimated decline on arrests at ages 18–24 is imprecisely estimated.

Last, we assess the robustness of our estimates to controlling for police office numbers. Fewer arrests could be explained both by higher numbers of police (leading to deterrence), or by lower numbers (leading to fewer arrests). Although the number of police officers may be a bad control, since it could respond to the policy, our estimated effect on juveniles remains negative and statistically significant at the 5% level, at −13% against −17% without the control on the same sample (Table A8). The estimated effect at ages 18–24 is not statistically significant with this additional control.

Winsorization of arrest rates In our preferred specification we winsorize arrests at the 95$^{\text{th}}$ percentile —common in crime applications as extreme right-tail outliers are frequent. Relaxing winsorization to the 99$^{\text{th}}$ percentile, or not winsorizing shows comparable declines, around 20% ($p = 0.001$ and $p = 0.002$). Arrests at ages 18–24 are

[26] The five states which facilitated legalized abortion earlier are Alaska, California, Hawaii, New York, and Washington.

not sensitive to this choice either and estimated at around 15% ($p = 0.083$ and $p = 0.079$) (Table A9).

Control group Our main estimates use not-yet-treated states as the comparison group. Restricting the comparison to never-treated states leaves virtually unchanged coefficients. Juvenile arrests fall by about 19% ($p < 0.001$), arrests at ages 18–24 by about 15% (p-value 0.08), while the effects on years of schooling and employment stay null (Table A10).

Event window Our preferred estimates restrict the sample to observations within 10 years before and 20 years after the ban. The results are not sensitive to this choice (Table A11). Across windows of $[-5, +15]$, $[-8, +15]$ and $[-15, +25]$, juvenile arrests fall by 16 to 20% (all p-values below 0.01), and arrests at ages 18–24 are declines between 12 and 17% (with p-values below 0.1). Corporal punishment falls by about 90%, and the effects on years of schooling and employment remain null. The effect of suspensions is only statistically significant when including longer leads in the estimates (-13% in both the $[-8, +15]$ and $[-15, +25]$ windows, $p = 0.025$ and $p = 0.020$, against -11%, $p = 0.18$, in the $[-5, +15]$ window). Restricting the analysis to bans enacted in the period for which we have reliable data on the use of corporal punishment (1975–1994) yields virtually identical results to the main estimates (Table A12).

Estimators We assess the robustness to different estimators for dynamic difference in differences (Table A13). For the corporal punishment and suspension rates, we estimate Poisson two-way fixed effects and the Poisson Wooldridge (2025) estimator. For the remaining outcomes we use linear two-way fixed effects, the Wooldridge (2025) estimator and the imputation estimator of Borusyak et al. (2024). The corporal punishment effect stays near -90%; the suspensions rate drop at about 14% (16% with the Wooldridge estimator; p-values 0.026 and below 0.001), and juvenile arrests fall by 14 to 20%. Two-way fixed effects give the smallest decline (about -14%, p-value 0.041), consistent with the bias documented in the recent literature —particularly in this setting when many states are eventually treated and there are long staggered gaps in bans (Goodman-Bacon, 2021); the heterogeneity-robust Wooldridge and imputation estimators both yield declines of

about 20% (p-values below 0.001). At ages 18–24, the drop in arrests is estimated at −7% with the Wooldridge and imputation estimators (p-values 0.051 and 0.037). Education and labor-market estimates remain economically negligible.

Control variables Our baseline estimates do not control for state-level socio-economic indicators, because these may themselves be affected by the bans and would then be bad controls. At the same time, if these change around the time of the bans and also affect crime, omitting them could bias our estimates. We therefore estimate additional results controlling for GDP per capita, SNAP receipt, the Black population share and the share of adults without a high-school degree (Table A14). Declines in juvenile arrests estimated are unchanged, corporal punishment falls by about 85%, juvenile arrests by about 16% (p-value 0.006). The fall in suspensions is not statistically significant (−6%, $p = 0.31$), suggesting it proxies for slow-moving changes in demographics. The effects on years of schooling and employment remain null. The estimated effect on young adults is negative (−9%) but not statistically significant ($p = 0.49$).

Sample We assess whether the effects are driven by a specific area of the U.S. excluding each census region at a time (Table A15). Corporal punishment rates are consistently estimated to decline by 80 to 91% (all $p < 0.001$). The fall in suspensions is not statistically significant when excluding Southern states, although the point estimate remains 9% (p-value 0.25). The decline in arrests at ages 10-17 is estimated at 11 to 20%, with the smallest estimate (−11%, p-value 0.077) in the regressions excluding the South, where corporal punishment was—and still is—historically most concentrated. Estimated declines in arrests during young adulthood are estimated at 17 to 19% (p-values between 0.037 and 0.096) and are driven by states in the West. The effect is not statistically significant when excluding Western states —which include very populated areas like California— falling to −5% ($p = 0.45$). Estimates on years of schooling and employment are never statistically significant and consistently small in magnitude estimates are not driven by a single state across any outcome (Figure A8).

Because not all counties report arrest data in the UCR in all years, we run a robustness check restricting the sample to counties that report in every year within the event window

(350 counties in 31 states). We estimate a decline in juvenile arrests of 16% (p-value 0.030) (Appendix Table A16) in line with our preferred estimates. Arrests at ages 18–24 fall by 20%, but the estimate is imprecise (p-value 0.105).

The IPUMS data identify only larger counties, so our education and labor-market estimates come from a restricted set of counties than the arrests estimates. Restricting the arrest analysis to the 484 counties observed in both data sets yields similar results: juvenile arrests fall by 18% (p-value < 0.001) and arrests at ages 18–24 by 18% (p-value 0.029) (Appendix Table A16).

Spillovers across state borders. If corporal punishment bans in a state also change behavior in neighboring no-ban areas the estimates in the main results could be biased (in either direction, depending on the behavioral response). To assess the robustness of the estimates to plausible spillovers, we identify counties that are adjacent to a state with a different treatment status. Excluding the border counties confirms our main conclusions (Table A7, Panel C).

Weighs Our preferred estimates are weighted, to give more importance to larger counties. This is important in a context where many counties have very small populations (i.e. the smallest 20% account for only about 1.4% of the total population). Nonetheless, without weights —which gives each county equal importance— the pattern is unchanged: corporal punishment falls by 93% and suspensions by 21% (both $p < 0.001$), juvenile arrests decline by about 14% ($p = 0.003$), and the education and labor-market effects remain null (Table A18). The effects at ages 18–24 are not significant, likely because large Western states with high pre-ban corporal punishment rates as California no longer have as much importance.

Level of aggregation Our preferred estimates use county-year data due high under-reporting in the UCR. We verify that the results hold when aggregating to the state-year level (Table A17). This also allows us to use more years of CRDC data on corporal punishment. Corporal punishment rates decline by 90%, suspensions fall by 14% (p-value 0.018), and juvenile arrests fall by 22% (p-value < 0.001). Arrests at ages 18–24 fall by

13%, although the coefficient is not significantly (p-value 0.11). We estimate negligible effects on years of schooling and employment.

7 Mechanisms

Bans led to large declines in juvenile arrest rates which were not accompanied by longer years in schooling or employment gains. This suggests that the effects on crime are due to channels other than incapacitation or human capital acquisition. In this section, we explore whether direct or indirect experiences of corporal punishment could explain the estimated effects and discuss psychological theories that find support in our analyses.

Direct vs Indirect Experiences of Corporal Punishment. The heterogeneity analyses in Figure 2 suggest a causal link between direct experiences of corporal punishment and delinquency which is consistent with the correlational evidence (Gershoff, 2017; Gershoff, 2002, among others). Arrest rates decline the most among pupils who were disproportionately exposed to corporal punishment before the bans, i.e. Black pupils and pupils in states with high pre-ban use (Figure 2, panels c–d), and among pupils who spent their entire schooling under a ban rather than only part of it.

Can direct experiences of corporal punishment account for the observed decline in arrests? We use our estimates and the existing meta-analytical literature to provide an upper bound on the share of the decline in arrests that could be due to the direct channel. In treated states before the ban, 0.52 in every 100 students were corporally punished each year. Assuming that distinct pupils were corporally punished each year allows us to have a maximum plausible number of directly punished pupils: at most 6.1% of students completing 12 years of schooling. If the entire ATT on arrests (a decline from 6.45 to 5.22 per 100) operated through this group, each punished student's arrest rate would need to fall by 313%— an implausible magnitude as it would require punished pupils to be arrested at more than three times the average rate and to stop offending entirely.[27] To obtain a realistic bound, we apply odds ratios from the meta-analytic literature linking

[27]The cumulative 12-year probability is $1 - (1 - 0.52/100)^{12} \approx 6.1\%$. If only this group changed behavior, the per-person decline would be $\text{ATT}/p = 20.2$ per 100, or 313% of the 6.45 per 100 pre-ban arrest rate.

corporal punishment to antisocial behavior. Using a conservative range of 1.5–2.0 higher likelihood of anti-social behavior for the punished, direct effects explain between 16% and 30% of the effect. Estimates under different assumptions are in Table A19.

One possible explanation for large peer effects is that own experiences of corporal punishment may have made punished pupils more disruptive and crime-prone. Existing literature has shown large effects both of disruptive peers (Carrell et al., 2018; Carrell and Hoekstra, 2010), and of criminal colleagues in the classroom (Billings and Hoekstra, 2023; Carneiro et al., forthcoming; Kawarazaki, 2024). Decreasing the likelihood of offending for a pupil who is central in a network can lead to the multiplier effects that we document.

Psychological Foundations. The psychological literature offers several theories connecting exposure to violence to anti-social behavior.

First, students who perceive rules as fair are less disruptive and more likely to comply with norms (Fissel et al., 2019; Way, 2011). Removing physical punishment and its threat may make school discipline seem fairer, particularly to those students more likely to experience corporal punishment. Harsh punishment may instead hinder the internalization of social norms (Gershoff, 2002).

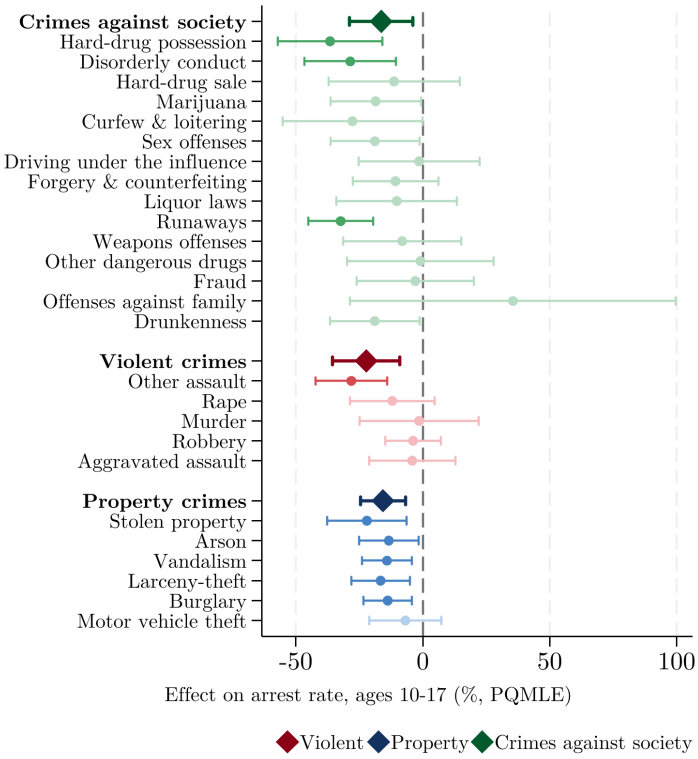
Second, when teachers use physical punishment, they may teach not only that a behavior is sanctioned but also that force is a legitimate response to conflict (Bandura, 1973). Children imitate adult behavior (Bandura, 1977; Bandura et al., 1961), and teachers' conduct shapes students' social attitudes and behavior (Algan et al., 2013; Lavy and Shayo, 2026). Once teachers no longer use force, pupils may be less likely to adopt it, even those who would never have been punished themselves.

Third, beyond this modeling effect, being struck may itself breed aggression. The cycle-of-violence hypothesis posits that victimization begets later aggression (Widom, 1989), and social information-processing models add that early victimization leads children to read others' behavior as hostile, fostering reactive aggression (Dodge, 1980; Dodge et al., 1990).

Fourth, physical punishment may impair self-regulation and mental health, leading to maladaptive coping such as substance use (Dube et al., 2003; Khantzian, 1997).

Heterogeneity analyses by crime type in Figure 3 show that arrests fall across most

Figure 3: Effects of Corporal Punishment Bans on Arrests by Crime Type (ages 10–17)



Notes: Stacked PQMLE estimates of the effect of corporal punishment bans on arrest rates at ages 10–17, by crime type, using county-year data. We build one stack per ban cohort, with never-treated and not-yet-treated counties as controls, and include stack-by-county and stack-by-year fixed effects. Broad UCR categories (diamonds, bold labels) head each block and are followed by their specific offenses; colors denote the broad family (violent, property, and crimes against society). The event window runs from −10 to +20. Estimates are weighted by the population aged 10–17. Arrest rates are winsorized at the 95th percentile. Estimates are expressed as percentage changes, $(e^{\hat{\beta}} - 1) \times 100$. Markers are point estimates and horizontal bars are 95% confidence intervals, from standard errors clustered at the state level (delta method); faded markers are not significant at the 5% level. The dashed line marks zero.

categories, with no statistically distinguishable difference between violent crimes, property crimes, and crimes against society. This general rather than offense-specific response is consistent with corporal punishment affecting norm internalization and self-control, which underlies offending of all types (Gottfredson and Hirschi, 1990). The largest declines are in hard-drug possession, followed by runaways, disorderly conduct, and violent assaults. The link between specific offenses and psychological channels is likely complex and intertwined. Taken at face value, however, the drops in drug possession and runaways are consistent with the bans reducing mental health problems and the coping they induce, and the drop in violent crime with aggression-based and cycle-of-violence theories.

8 Conclusion

This paper studies the effects of bans on corporal punishment in U.S. public schools. We exploit the staggered introduction of bans across 28 states between 1975 and 2011, and compare cohorts in states that banned corporal punishment with cohorts in states that had not (yet) banned it, up to age 29, in difference-in-differences designs. Our results show that the bans eliminated corporal punishment in schools, and that arrests at ages 10–17 fell by 19%. The effects are strongest among pupils more likely to be corporally punished at baseline. In some states, the arrest declines persist into young adulthood, at ages 18–24. The effects are not only due to declines among students who would have been punished. At least 70% of the arrest decline comes from spillovers to students who would not have been punished themselves. Removing physical punishment from schools hampered neither educational attainment nor labor market prospects.

These findings are consistent with a large body of medical and psychological research documenting associations between corporal punishment, aggression and delinquency (Cuartas et al., 2025; Gershoff and Grogan-Kaylor, 2016; Gershoff, 2002, among others). We are the first to show that at least part of this link is causal. We cannot disentangle the specific channels linking physical discipline to externalizing behavior, but the declines across crime types are consistent with weaker norm internalization, trauma and coping, and the cycle of violence. Witnessing violence, classroom climate, and peer spillovers are also plausible mechanisms.

While experiences of physical discipline in the classroom can differ from other types of violence, our results might also be informative on how experiences of violence lead to later aggression in a more general sense (a link that the crime literature deems the victim-offender overlap). They are informative in the context of large disparities in incidence of corporal punishment, whereby Black pupils and pupils in poorer counties are more likely to experience punishment. Part of the racial gap in youth arrests across socioeconomic groups may to an extent reflect how schools treat different children.

The results align with the existing literature favoring less punitive approaches in schools (Adukia et al., 2025; Bacher-Hicks et al., 2024). They are relevant for the 17

U.S. states where corporal punishment in schools remains legal, and for the 62 countries that still permit it. They may also inform debates on banning corporal punishment by parents, a policy adopted in about 70 countries that remains controversial elsewhere.

References

Adukia, A., B. Feigenberg, and F. Momeni (2025). "From Retributive to Restorative: An Alternative Approach to Justice in Schools". *American Economic Review* 115.8, pp. 2722–2754.

Algan, Y., P. Cahuc, and A. Shleifer (2013). "Teaching Practices and Social Capital". *American Economic Journal: Applied Economics* 5.3, pp. 189–210.

Anderson, D. M. (2014). "In School and Out of Trouble? The Minimum Dropout Age and Juvenile Crime". *Review of Economics and Statistics* 96.2, pp. 318–331.

Bacher-Hicks, A., S. B. Billings, and D. J. Deming (Nov. 2024). "The School-to-Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime". *American Economic Journal: Economic Policy* 16.4, pp. 165–193.

Baker, A., B. Callaway, S. Cunningham, A. Goodman-Bacon, and P. H. C. Sant'Anna (2026). "Difference-in-Differences Designs: A Practitioner's Guide". *Journal of Economic Literature* 64.2, pp. 498–557.

Bandura, A. (1973). *Aggression: A Social Learning Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

— (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., D. Ross, and S. A. Ross (1961). "Transmission of Aggression Through Imitation of Aggressive Models". *The Journal of Abnormal and Social Psychology* 63.3, pp. 575–582.

Becker, G. S. (1968). "Crime and Punishment: An Economic Approach". *Journal of Political Economy* 76.2, pp. 169–217.

Bell, B., R. Costa, and S. Machin (2022). "Why Does Education Reduce Crime?" *Journal of Political Economy* 130.3, pp. 732–765.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119.1, pp. 249–275.

Bhatt, M. P., S. B. Heller, M. Kapustin, M. Bertrand, and C. Blattman (2024). "Predicting and Preventing Gun Violence: An Experimental Evaluation of READI Chicago". *Quarterly Journal of Economics* 139.1, pp. 1–56.

Billings, S. B., D. J. Deming, and S. L. Ross (2019). "Partners in Crime". *American Economic Journal: Applied Economics* 11.1, pp. 126–150.

Billings, S. B. and M. Hoekstra (2023). "The Effect of School and Neighborhood Peers on Achievement, Misbehavior, and Adult Crime". *Journal of Labor Economics* 41.3, pp. 643–685.

Bindler, A. and N. Ketel (2022). "Scaring or Scarring? Labor Market Effects of Criminal Victimization". *Journal of Labor Economics* 40.4, pp. 939–970.

Bindler, A., N. Ketel, and R. Hjalmarsson (2020). "Costs of Victimization". *Handbook of Labor, Human Resources and Population Economics*. Ed. by K. F. Zimmermann. Cham: Springer, pp. 1–31.

Borusyak, K., X. Jaravel, and J. Spiess (Feb. 2024). "Revisiting Event-Study Designs: Robust and Efficient Estimation". *Review of Economic Studies* 91.6, pp. 3253–3285.

Callaway, B. and P. H. Sant'Anna (2021). "Difference-in-Differences with multiple time periods". *Journal of Econometrics* 225.2, pp. 200–230.

Card, D. (1999). "The Causal Effect of Education on Earnings". *Handbook of Labor Economics*. Ed. by O. Ashenfelter and D. Card. Vol. 3A. Amsterdam: Elsevier, pp. 1801–1863.

Carneiro, P., Y. Cruz-Aguayo, F. Salvati, and N. Schady (forthcoming). "Disruption in the Classroom: Experimental Evidence from Ecuador". *American Economic Journal: Applied Economics*.

Carrell, S. E., M. Hoekstra, and E. Kuka (Nov. 2018). "The Long-Run Effects of Disruptive Peers". *American Economic Review* 108.11, pp. 3377–3415.

Carrell, S. E. and M. L. Hoekstra (2010). "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids". *American Economic Journal: Applied Economics* 2.1, pp. 211–228.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). "The Effect of Minimum Wages on Low-Wage Jobs". *Quarterly Journal of Economics* 134.3, pp. 1405–1454.

Chaisemartin, C. de and X. D'Haultfœuille (2020). "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects". *American Economic Review* 110.9, pp. 2964–2996.

Cornaglia, F., N. E. Feldman, and A. Leigh (2014). "Crime and Mental Well-Being". *Journal of Human Resources* 49.1, pp. 110–140.

Cuartas, J. (2022). "The effect of spanking on early social-emotional skills". *Child Development* 93.1, pp. 180–193.

Cuartas, J., E. T. Gershoff, D. H. Bailey, M. A. Gutiérrez, and D. C. McCoy (2025). "Physical punishment and lifelong outcomes in low- and middle-income countries: a systematic review and multilevel meta-analysis". *Nature Human Behaviour* 9.7, pp. 1365–1379.

Currie, J. and E. Tekin (2012). "Understanding the Cycle: Childhood Maltreatment and Future Crime". *Journal of Human Resources* 47.2, pp. 509–549.

Davern, M., R. Bautista, J. Freese, P. Herd, and S. L. Morgan (2025). *General Social Survey 1972–2024*. Data accessed from the GSS Data Explorer website at https://gssdataexplorer.norc.org. Sponsored by the National Science Foundation. NORC ed.

Davis, J. M. and S. B. Heller (2020). "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs". *Review of Economics and Statistics* 102.4, pp. 664–677.

Deshpande, M. and M. Mueller-Smith (2022). "Does Welfare Prevent Crime? The Criminal Justice Outcomes of Youth Removed from SSI". *Quarterly Journal of Economics* 137.4, pp. 2263–2307.

Dinarte-Diaz, L., S. Ravindran, M. Shah, S. Powers, and H. Baker-Henningham (2026). "Violent discipline and parental behavior: Short- and medium-term effects of digital parenting support to caregivers". *Journal of Public Economics* 256, p. 105590.

Dodge, K. A. (1980). "Social Cognition and Children's Aggressive Behavior". *Child Development* 51.1, pp. 162–170.

Dodge, K. A., J. E. Bates, and G. S. Pettit (1990). "Mechanisms in the Cycle of Violence". *Science* 250.4988, pp. 1678–1683.

Donohue, J. J. and S. D. Levitt (2001). "The Impact of Legalized Abortion on Crime". *Quarterly Journal of Economics* 116.2, pp. 379–420.

Dube, S. R., V. J. Felitti, M. Dong, D. P. Chapman, W. H. Giles, and R. F. Anda (2003). "Childhood Abuse, Neglect, and Household Dysfunction and the Risk of Illicit Drug Use: The Adverse Childhood Experiences Study". *Pediatrics* 111.3, pp. 564–572.

Dustmann, C. and F. Fasani (2016). "The Effect of Local Area Crime on Mental Health". *The Economic Journal* 126.593, pp. 978–1017.

Egyir, J., E. De Cao, K. Thomas, and E. Aurino (2026). *Global Levels and Trends in Child Discipline: Evidence from 88 Countries, 2005–2023*. medRxiv preprint.

FBI Crime Data Explorer (2026). *Law Enforcement Employees Data, 1960–2025*. Data set. last modified May 2026. Washington, DC: Federal Bureau of Investigation, Uniform Crime Reporting (UCR) Program.

Fissel, E. R., P. Wilcox, and M. S. Tillyer (2019). "School Discipline Policies, Perceptions of Justice, and In-School Delinquency". *Crime & Delinquency* 65.10, pp. 1343–1370.

Foureaux Koppensteiner, M. and M. Manacorda (2016). "Violence and birth outcomes: Evidence from homicides in Brazil". *Journal of Development Economics* 119, pp. 16–33.

Foureaux Koppensteiner, M. and L. Menezes (2021). "Violence and Human Capital Investments". *Journal of Labor Economics* 39.3, pp. 787–823.

Francis, T. and H. Baker-Henningham (2021). "The Irie Homes Toolbox: A cluster randomized controlled trial of an early childhood parenting program to prevent violence against children in Jamaica". *Children and Youth Services Review* 126, p. 106060.

Gershoff, E. T. (2017). "School corporal punishment in global perspective: prevalence, outcomes, and efforts at intervention". *Psychology, Health & Medicine* 22.sup1, pp. 224–239.

Gershoff, E. T. and S. A. Font (2016). "Corporal Punishment in U.S. Public Schools: Prevalence, Disparities in Use, and Status in State and Federal Policy". *Social Policy Report* 30.1, pp. 1–26.

Gershoff, E. T. and A. Grogan-Kaylor (2016). "Spanking and Child Outcomes: Old Controversies and New Meta-Analyses". *Journal of Family Psychology* 30.4, pp. 453–469.

Gershoff, E. T. (2002). "Corporal Punishment by Parents and Associated Child Behaviors and Experiences: A Meta-Analytic and Theoretical Review". *Psychological Bulletin* 128.4, pp. 539–579.

Global Initiative to End All Corporal Punishment of Children (2024). *Corporal punishment of children in the USA*. Accessed: 2025-03-28.

Goodman-Bacon, A. (2021). "Difference-in-differences with variation in treatment timing". *Journal of Econometrics* 225.2, pp. 254–277.

Gottfredson, M. R. and T. Hirschi (1990). *A General Theory of Crime*. Stanford, CA: Stanford University Press.

Gunderson, M. and P. Oreopoulos (2020). "Returns to education in developed countries". *The Economics of Education: A Comprehensive Overview*. Ed. by S. Bradley and C. Green. 2nd ed. London: Academic Press, pp. 39–51.

Heekes, S.-L., C. B. Kruger, S. N. Lester, and C. L. Ward (2022). "A systematic review of corporal punishment in schools: Global prevalence and correlates". *Trauma, Violence, & Abuse* 23.1, pp. 52–72.

Heller, S. B. (2014). "Summer jobs reduce violence among disadvantaged youth". *Science* 346.6214, pp. 1219–1223.

Heller, S. B., A. K. Shah, J. Guryan, J. Ludwig, S. Mullainathan, and H. A. Pollack (2017). "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago". *Quarterly Journal of Economics* 132.1, pp. 1–54.

Hjalmarsson, R., H. Holmlund, and M. J. Lindquist (Sept. 2015). "The Effect of Education on Criminal Convictions and Incarceration: Causal Evidence from Micro-data". *The Economic Journal* 125.587, pp. 1290–1326.

Human Rights Watch and American Civil Liberties Union (2008). *A Violent Education: Corporal Punishment of Children in US Public Schools*. Tech. rep. New York: Human Rights Watch and American Civil Liberties Union.

Jaffee, S. R., A. Caspi, T. E. Moffitt, M. Polo-Tomas, T. S. Price, and A. Taylor (2004). "The Limits of Child Effects: Evidence for Genetically Mediated Child Effects on Corporal Punishment but Not on Physical Maltreatment". *Developmental Psychology* 40.6, pp. 1047–1058.

Jennings, W. G., A. R. Piquero, and J. M. Reingle (2012). "On the overlap between victimization and offending: A review of the literature". *Aggression and Violent Behavior* 17.1, pp. 16–26.

Johnston, D. W., M. A. Shields, and A. Suziedelyte (2018). "Victimisation, Well-being and Compensation: Using Panel Data to Estimate the Costs of Violent Crime". *The Economic Journal* 128.611, pp. 1545–1569.

Kawarazaki, H. (2024). "When Trouble Travels: Peer Effects of Criminal Exposure in Schools". Working paper, University College London and Institute for Fiscal Studies.

Khantzian, E. J. (1997). "The Self-Medication Hypothesis of Substance Use Disorders: A Reconsideration and Recent Applications". *Harvard Review of Psychiatry* 4.5, pp. 231–244.

Lavy, V. and M. Shayo (Mar. 2026). *Imitation, Reciprocity, and Community Structure in Norm Transmission: Evidence from Teachers and Students*. Working Paper 34931. National Bureau of Economic Research.

Le, K. and M. Nguyen (2019). "'Bad Apple' peer effects in elementary classrooms: the case of corporal punishment in the home". *Education Economics* 27.6, pp. 557–572.

Levitt, S. D. (2004). "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not". *Journal of Economic Perspectives* 18.1, pp. 163–190.

Lochner, L. and E. Moretti (2004). "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports". *American Economic Review* 94.1, pp. 155–189.

Mahuteau, S. and R. Zhu (2016). "Crime Victimisation and Subjective Well-Being: Panel Evidence From Australia". *Health Economics* 25.11, pp. 1448–1463.

Monteiro, J. and R. Rocha (2017). "Drug Battles and School Achievement: Evidence from Rio de Janeiro's Favelas". *Review of Economics and Statistics* 99.2, pp. 213–228.

National Cancer Institute (2025). *Single Year of Age County Population Estimates, 1969–2023*. County-level data; Accessed 2025-02-02. Surveillance, Epidemiology, and End Results (SEER) Program.

Ornstein, P. (2017). *The price of violence: Consequences of violent crime in Sweden*. Working Paper 2017:22. Uppsala: Institute for Evaluation of Labour Market and Education Policy (IFAU).

Psacharopoulos, G. and H. A. Patrinos (2018). "Returns to investment in education: a decennial review of the global literature". *Education Economics* 26.5, pp. 445–458.

Raichle, D. R. (1974). "The Abolition of Corporal Punishment in New Jersey Schools". *History of Childhood Quarterly* 2.1, pp. 53–78.

Ruggles, S., S. Flood, M. Sobek, D. Backman, G. Cooper, J. A. Rivera Drew, S. Richards, R. Rodgers, J. Schroeder, and K. C. W. Williams (2025). *IPUMS USA: Version 16.0*. dataset. American Community Survey (2001–2024) and Decennial Census samples (1960–2000). Minneapolis, MN.

Schroeder, J., D. Van Riper, S. Manson, G. Cooper, Z. Krause, T. Kugler, T. Zhu, and S. Ruggles (2026). *IPUMS National Historical Geographic Information System: Version 21.0*. Version 21.0. Minneapolis, MN: IPUMS.

Sharkey, P. and G. Torrats-Espinosa (2017). "The effect of violent crime on economic mobility". *Journal of Urban Economics* 102, pp. 22–33.

Supreme Court of the United States (1977). *U.S. Reports: Ingraham v. Wright, 430 U.S. 651 (1977)*. U.S. Reports. Retrieved from the Library of Congress.

U.S. Bureau of Labor Statistics (2026). *Annual Unemployment Rates by State*. Local Area Unemployment Statistics.

U.S. Department of Agriculture, Economic Research Service (2025). *Educational attainment for adults age 25 and older for the United States, States, and counties, 1970–2023. County-Level Data Sets*.

U.S. Department of Education (2024). *Civil Rights Data Collection*. Includes current data from https://civilrightsdata.ed.gov/publications and historical data from https://civilrightsdata.ed.gov/archive. Accessed: 2025-07-16.

U.S. Federal Bureau of Investigation (2024). *Uniform Crime Reporting Program Data: Arrests by Age, Sex, and Race, Summarized Yearly, United States, 2021*.

University of Kentucky Center for Poverty Research (2026). *UKCPR National Welfare Data, 1980–2024*. https://ukcpr.uky.edu/resources/national-welfare-data. Accessed: 25 February 2026. Lexington, KY.

Way, S. M. (2011). "School Discipline and Disruptive Classroom Behavior: The Moderating Effects of Student Perceptions". *The Sociological Quarterly* 52.3, pp. 346–375.

Widom, C. S. (1989). "The Cycle of Violence". *Science* 244.4901, pp. 160–166.

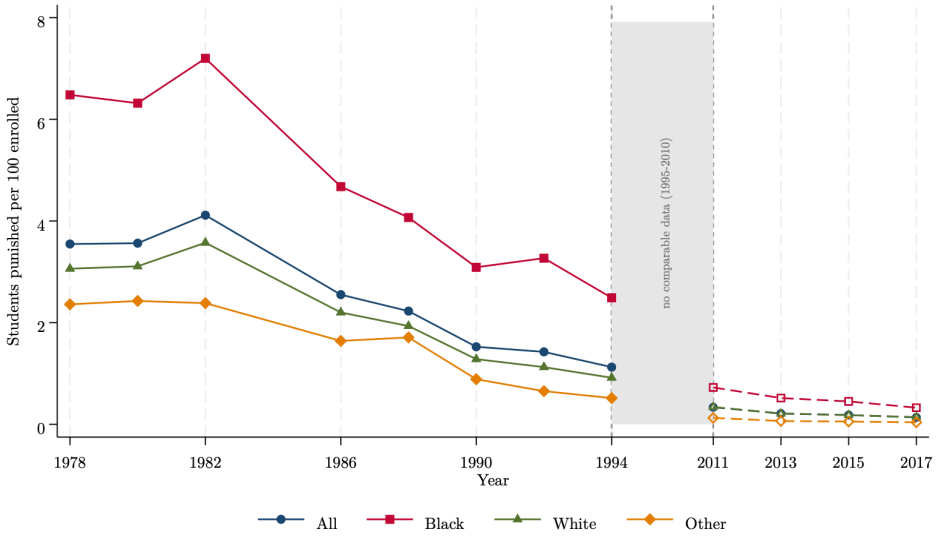
Wooldridge, J. M. (2023). "Simple approaches to nonlinear difference-in-differences with panel data". *The Econometrics Journal* 26.3, pp. C31–C66.

— (2025). "Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators". *Empirical Economics* 69.5, pp. 2545–2587.

Xie, M. and D. McDowall (2008). "Escaping Crime: The Effects of Direct and Indirect Victimization on Moving". *Criminology* 46.4, pp. 809–840.

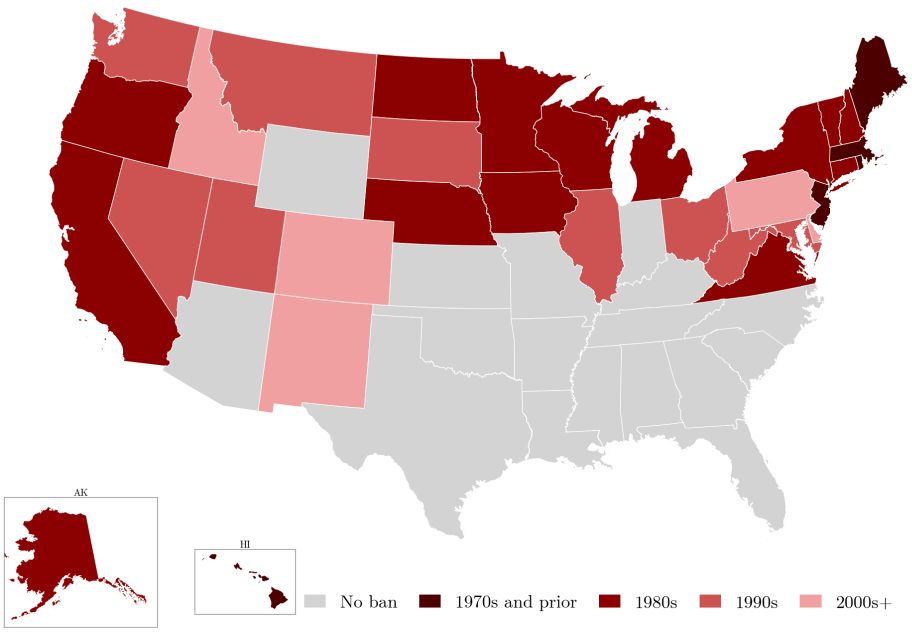
A Additional Figures and Tables

Figure A1: Corporal Punishment Rate over Time, by Ethnicity



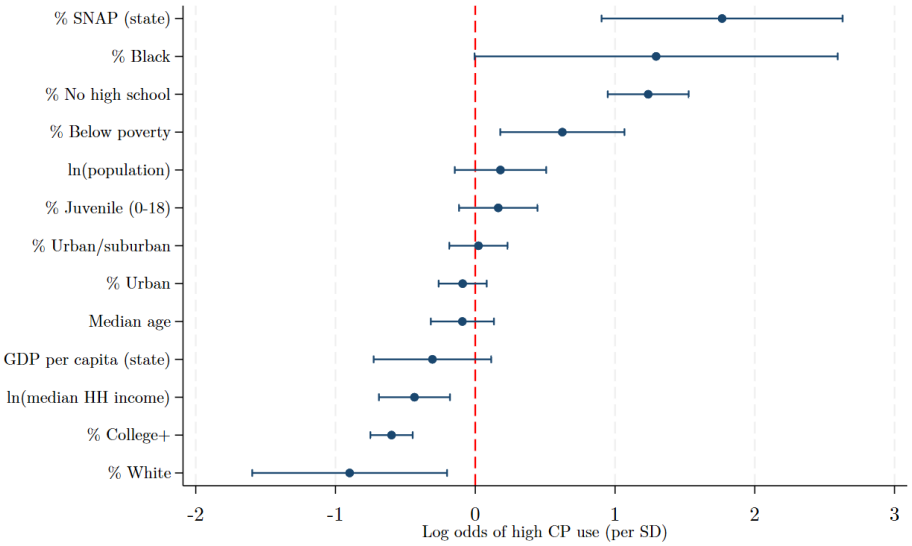
Notes: National rate of students corporally punished per 100 enrolled, weighted by enrollment. Solid lines show county-level CRDC data, 1978–1994 (our analysis sample). Dashed lines show state-level CRDC waves, 2011–2017 (descriptive only). Each series is weighted by its own enrollment count. The shaded band marks 1995–2010, for which no comparable data exist.

Figure A2: Map of the United States by Corporal Punishment Ban Status



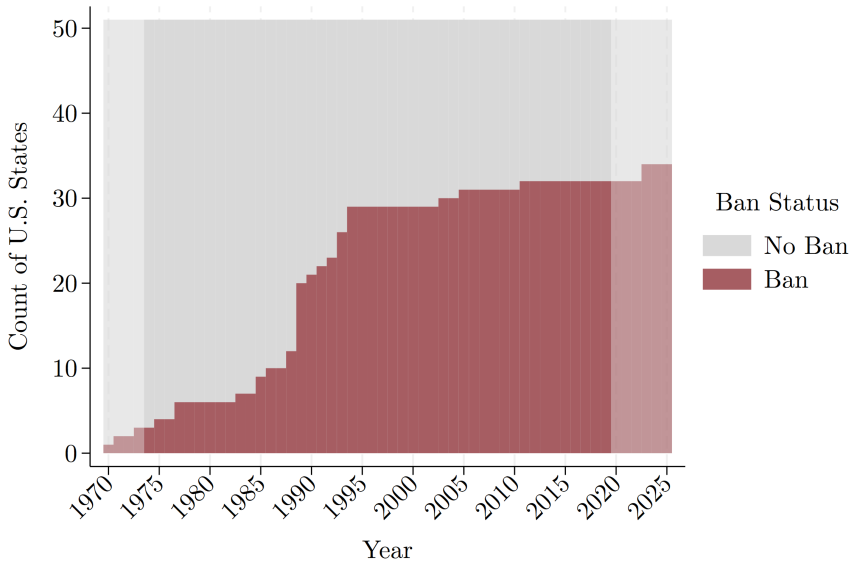
Notes: The map displays U.S. state-level bans on corporal punishment in public schools for the entire student body, including bans enacted after our data end (Colorado and Idaho, 2023). Color intensity reflects the decade of enactment: darkest red indicates bans passed in the 1970s or earlier, followed by progressively lighter shades for the 1980s, 1990s, and 2000s or later. Ohio is dated by its 1994 partial ban. Gray states had not enacted a statewide ban as of 2025.

Figure A3: County attributes and corporal punishment use



Notes: Using bivariate log odds, each marker is the coefficient from a separate logit of an indicator for high county corporal punishment use—the county mean over 1978–1994, split at the median among ever-ban counties—on one standardized county attribute, so that the coefficient is the change in log odds per standard deviation. County attributes are measured in the 1980 Census (poverty, education, Black and White population shares, urban and urban-or-suburban shares, log population, log median household income, median age, and the juvenile population share); the SNAP share and GDP per capita are measured at the state level. Markers are point estimates and bars are 95% confidence intervals, from analytic standard errors clustered at the state level; the dashed line marks zero.

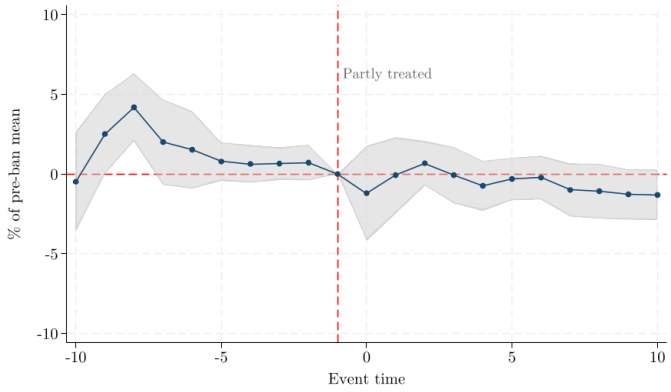
Figure A4: Count of states with and without bans over time



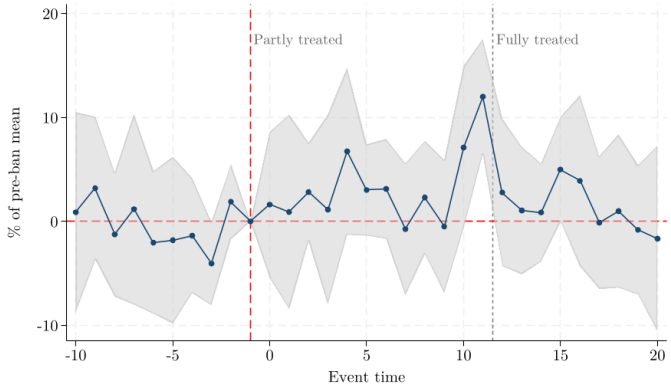
Notes: Number of U.S. states with and without a ban on corporal punishment in public K–12 schools, 1970 to 2025. Gray bars count states without a ban; red bars count states with a ban.

Figure A5: Event-study estimates of the effects of corporal punishment bans — additional outcomes

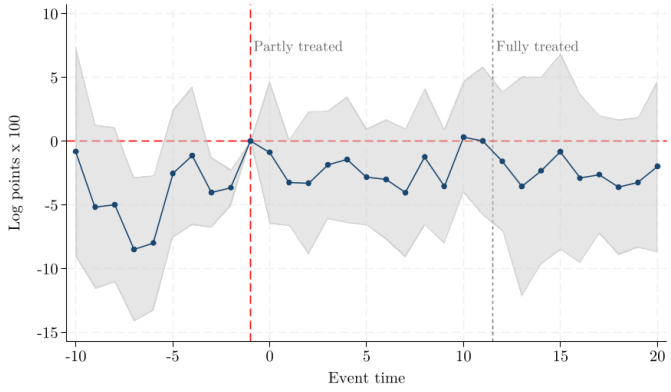
(a) HS completion (ages 25–29)



(b) BA or more (ages 25–29)



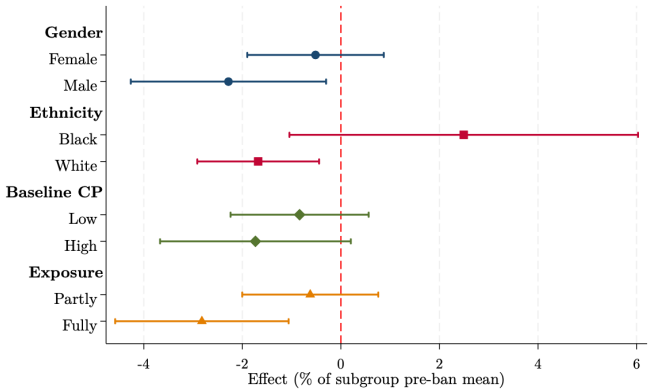
(c) Log real wages (ages 25–29)



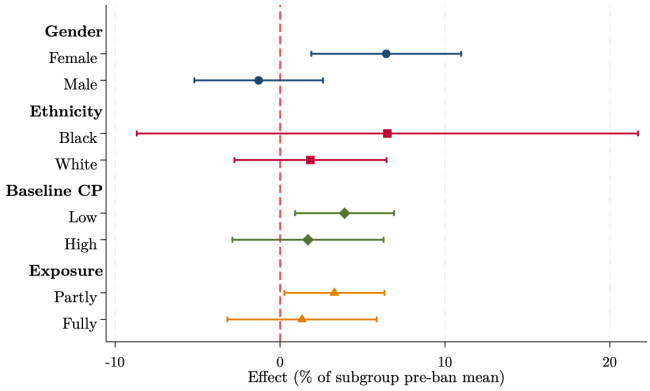
Notes: Event-study estimates of the effects of corporal punishment bans on the education and labor-market outcomes not shown in Figure 1. All panels are Callaway and Sant'Anna (2021) estimates for outcomes measured at ages 25–29 for U.S.-born individuals living in their state of birth, using the cohort design (p = birth year, G = ban year -18). The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by cell population and expressed as a percentage of the pre-ban mean in treated units, and in log points $\times 100$ for log wages. Shaded areas are 95% confidence intervals from standard errors clustered at the state level using a wild cluster bootstrap.

Figure A6: Heterogeneity in the effects of corporal punishment bans — additional outcomes

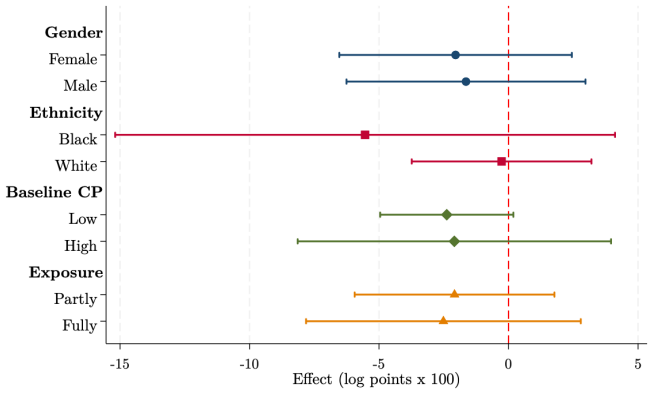
(a) HS completion (ages 25–29)



(b) BA or more (ages 25–29)

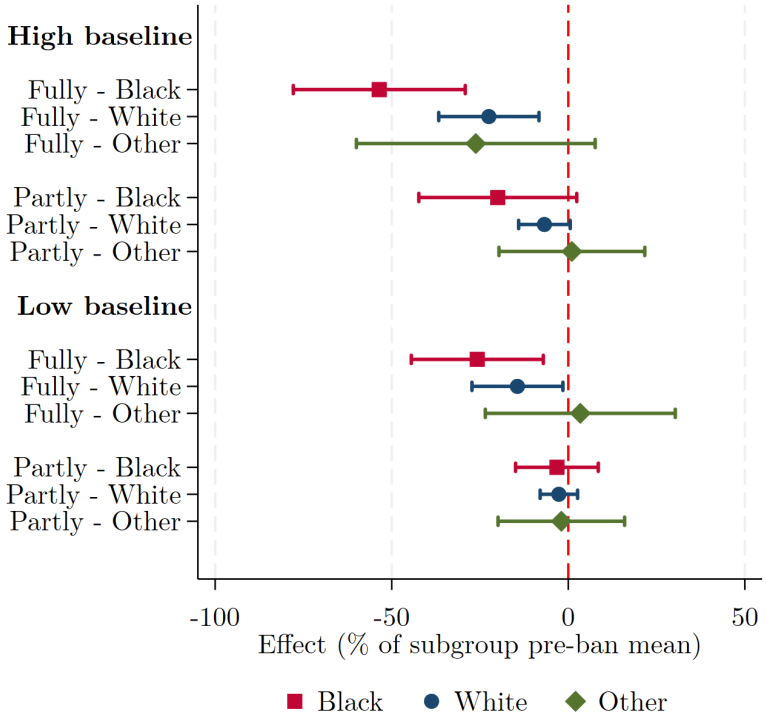


(c) Log real wages (ages 25–29)



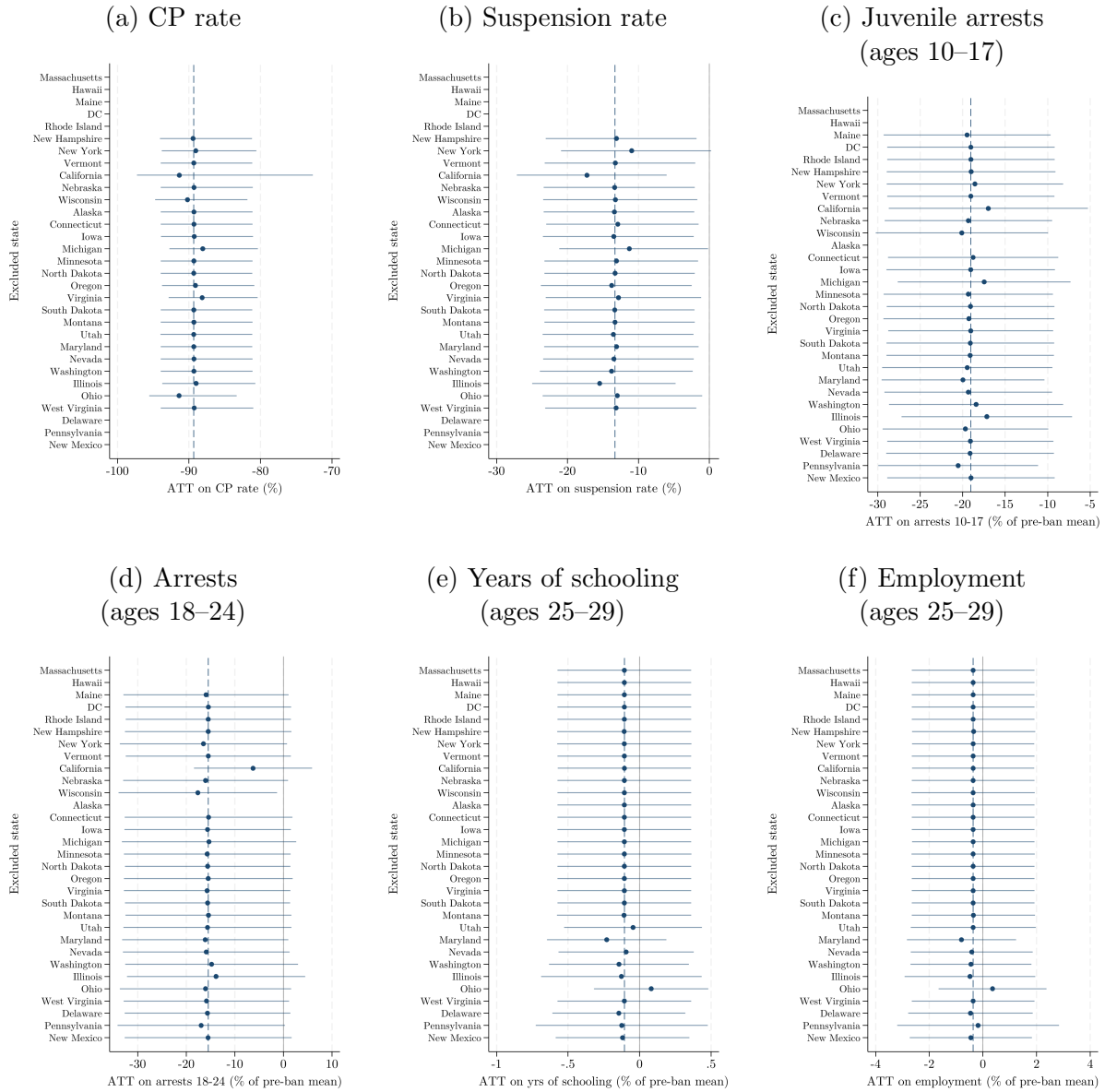
Notes: Heterogeneity in the effects of corporal punishment bans on the education and labor-market outcomes not shown in Figure 2, by gender, ethnicity, baseline corporal punishment and exposure, as defined there. "Other" is dropped, as too imprecise to be informative. All panels report Callaway and Sant'Anna (2021) estimates using the cohort design (p = birth year, G = ban year -18), weighted by cell population. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are expressed as a percentage of the subgroup pre-ban mean in treated units, and in log points $\times 100$ for log wages. Markers are point estimates and horizontal bars are 95% confidence intervals, from standard errors clustered at the state level using a wild cluster bootstrap.

Figure A7: Heterogeneity in juvenile arrest declines —additional subcategories



Notes: Callaway and Sant'Anna (2021) estimates of the effect of corporal punishment bans on the juvenile (ages 10–17) arrest rate, for each baseline corporal punishment × exposure × ethnicity cell, using the county-year design (p = calendar year, G = ban year). Effects are averaged over event years 0–4 ("partly treated") or 12–20 ("fully treated"). High (low) baseline states are those above (below) the median pre-ban corporal punishment rate among ever-ban states; fully treated cohorts were aged 5 or younger at the ban. The comparison group is not-yet-treated units and the event window runs from −10 to +20. Estimates use race-specific county-year arrest rates and are weighted by the race-specific age-band population. Arrest rates are winsorized at the 95th percentile. Effects are expressed as a percentage of the subgroup pre-ban mean in treated units. Markers are point estimates and horizontal bars are 95% confidence intervals, from standard errors clustered at the state level using a wild cluster bootstrap; the dashed line marks zero.

Figure A8: Robustness: leave-one-state-out estimates for the main outcomes



Notes: Estimated effects of corporal punishment bans on the six main outcomes, after dropping one ever-ban state at a time. States are ordered by year of ban introduction and occupy the same row in every panel; a row without an estimate marks a state outside that outcome's estimation sample. The corporal punishment and suspension rates are stacked PQMLE estimates; the remaining outcomes are Callaway and Sant'Anna (2021) estimates. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by enrollment (corporal punishment and suspension rates), by age-band population (arrests) and by cell population (cohort outcomes). Arrest rates are winsorized at the 95th percentile. Estimates are expressed as a percentage of the pre-ban mean in treated units, computed as $(\exp(\hat{\beta}) - 1) \times 100$ for the corporal punishment and suspension rates and as the ATT divided by the mean for the other four outcomes. Markers are point estimates and horizontal bars are 95% confidence intervals; the dashed line marks the full-sample estimate. Standard errors are clustered at the state level.

Table A1: Corporal Punishment Bans by State

State	Ban year	State	Ban year
Alabama	–	Montana	1991
Alaska	1989	Nebraska	1988
Arizona	–	Nevada	1993
Arkansas	–	New Hampshire	1983
California	1986	New Jersey	1867*
Colorado	2023*	New Mexico	2011
Connecticut	1989	New York	1985
Delaware	2003	North Carolina	–
District of Columbia	1977	North Dakota	1989
Florida	–	Ohio	1994[a]
Georgia	–	Oklahoma	–
Hawaii	1973*	Oregon	1989
Idaho	2023*	Pennsylvania	2005
Illinois	1994	Rhode Island	1977
Indiana	–	South Carolina	–
Iowa	1989	South Dakota	1990
Kansas	–	Tennessee	–
Kentucky	–	Texas	–
Louisiana	–	Utah	1992
Maine	1975	Vermont	1985
Maryland	1993	Virginia	1989
Massachusetts	1971*	Washington	1993
Michigan	1989	West Virginia	1994
Minnesota	1989	Wisconsin	1988
Mississippi	–	Wyoming	–
Missouri	–		

Notes: Year in which each state banned corporal punishment in public K–12 schools. "–" indicates the state has not enacted a ban. Source: Global Initiative to End All Corporal Punishment of Children (2024). [a] Ohio first introduced a partial ban in 1994, allowing districts to opt out; the full statewide ban followed in 2009. * We do not leverage these bans as they took effect before our data begin (New Jersey, Massachusetts and Hawaii) or after our data end (Colorado and Idaho).

Table A2: Summary Statistics

	Mean	SD	Years	Units	N
Panel A. County-year outcomes					
Corporal punishment rate (per 100 students)	4.997	6.938	1978–1994	2,540	13,018
Male	7.786	10.373	1978–1994	2,540	13,018
Female	2.020	3.563	1978–1994	2,540	13,018
Black	7.561	12.146	1978–1994	2,408	11,677
Other	2.810	10.237	1978–1994	2,491	12,183
White	4.499	6.369	1978–1994	2,540	13,010
Suspension rate (per 100 students)	4.147	3.737	1978–1994	2,888	14,406
Male	5.788	5.155	1978–1994	2,888	14,406
Female	2.393	2.422	1978–1994	2,888	14,406
Black	7.355	10.367	1978–1994	2,683	12,594
Other	3.283	31.505	1978–1994	2,832	13,449
White	3.561	3.287	1978–1994	2,888	14,394
Arrest rate, ages 10–17 (per 100)	3.869	3.294	1974–2019	2,915	85,344
Male	5.597	4.652	1974–2019	2,915	85,344
Female	2.023	2.005	1974–2019	2,915	85,344
Black	7.370	8.871	1974–2019	2,756	70,717
Other	1.680	3.467	1974–2019	2,826	74,112
White	3.598	3.164	1974–2019	2,915	85,339
Arrest rate, ages 18–24 (per 100)	11.567	6.723	1974–2019	2,915	85,341
Male	18.304	10.718	1974–2019	2,915	85,341
Female	4.417	3.247	1974–2019	2,915	85,341
Panel B. County-cohort outcomes (by year of birth)					
Years of schooling, ages 25–29	13.304	0.691	1961–1999	434	8,383
Male	13.083	0.786	1961–1999	434	8,372
Female	13.532	0.852	1961–1999	434	8,375
Black	12.708	1.261	1961–1999	421	7,100
Other	13.759	2.328	1961–1999	413	3,960
White	13.502	0.811	1961–1999	434	8,382
Employed, ages 25–29 (share)	0.773	0.096	1961–1999	434	8,383
Male	0.809	0.122	1961–1999	434	8,372
Female	0.737	0.132	1961–1999	434	8,375
Black	0.651	0.280	1961–1999	421	7,100
Other	0.734	0.350	1961–1999	413	3,960
White	0.797	0.102	1961–1999	434	8,382
Panel C. State-year outcomes					
Corporal punishment rate (per 100 students)	2.439	3.527	1976–1994	44	394
Suspension rate (per 100 students)	5.116	2.322	1976–1994	44	394
Arrest rate, ages 10–17 (per 100)	5.741	2.624	1974–2019	45	1,448
Arrest rate, ages 18–24 (per 100)	12.011	3.474	1974–2019	47	1,603

Notes: Unweighted means and standard deviations of the outcomes in the main regressions, on their estimation samples. Panel A covers county-year outcomes: corporal punishment and suspension rates per 100 enrolled students from the Civil Rights Data Collection (biennial waves, estimation sample of Table 1, columns 1–2) and arrest rates per 100 residents of each age band from the UCR (Table 1, columns 3–4). Panel B covers county-cohort outcomes from the Decennial Census and the American Community Survey, measured at ages 25–29 for U.S.-born individuals living in their state of birth (Table 1, columns 5–6); "Years" refers to years of birth. Panel C covers the state-year outcomes of Table A17. Indented rows report the subgroups used in the heterogeneity analysis (Figure 2), on the same estimation sample as the total and for the units reporting each group; corporal punishment and suspension rates by subgroup are per 100 enrolled students of that group. Arrests at ages 18–24 are not available by ethnicity. "Units" is the number of counties (Panels A and B) or states (Panel C). Arrest rates are winsorized at the 95th percentile.

Table A3: Baseline Characteristics by Decade of Corporal Punishment Ban

	Never	1970s		1980s		1990s		Post 2000s	
	Banned	Mean	Diff.	Mean	Diff.	Mean	Diff.	Mean	Diff.
Offense Rate (10–17)	4.9	6.5	1.5	6.0	1.0	5.9	1.0	5.6	0.7
	(1.9)	(2.2)	(1.2)	(2.4)	(0.8)	(2.5)	(0.9)	(1.8)	(1.2)
% Corp. Punishment	7.3	0.0	-7.3**	0.5	-6.9***	2.1	-5.2***	4.1	-3.2
	(4.5)	(0.0)	(2.7)	(0.9)	(1.2)	(2.0)	(1.6)	(2.1)	(2.7)
% Population (10–17)	13.7	12.8	-0.9	13.6	-0.1	13.6	-0.1	14.0	0.3
	(0.7)	(1.5)	(0.5)	(0.5)	(0.2)	(0.4)	(0.3)	(0.8)	(0.5)
% Black	17.6	30.0	12.4	6.6	-11.0***	8.2	-9.4*	11.2	-6.4
	(13.3)	(48.2)	(12.3)	(7.2)	(3.9)	(9.4)	(5.0)	(9.3)	(8.1)
% Other	1.9	1.0	-1.0	3.6	1.7	3.8	1.9	4.3	2.4
	(2.6)	(0.2)	(1.5)	(6.1)	(1.6)	(3.2)	(1.1)	(6.0)	(1.9)
% No High School	43.1	32.9	-10.2	32.7	-10.4***	33.8	-9.3**	36.2	-6.9
	(10.3)	(1.0)	(6.1)	(5.8)	(3.1)	(7.6)	(3.9)	(1.1)	(6.1)
% Urban	10.7	62.2	51.5***	16.5	5.8	13.5	2.8	20.8	10.1**
	(5.9)	(46.2)	(9.7)	(14.0)	(3.6)	(8.6)	(2.8)	(11.5)	(4.1)
% Unemployment	6.6	7.4	0.8	6.7	0.1	7.1	0.5	7.7	1.1
	(1.5)	(0.2)	(0.9)	(2.2)	(0.6)	(1.5)	(0.6)	(0.2)	(0.9)
% SNAP	9.8	9.9	0.0	6.3	-3.5***	6.6	-3.2**	13.2	3.4
	(4.3)	(2.2)	(2.6)	(2.0)	(1.2)	(2.2)	(1.5)	(4.6)	(2.7)
GDP per capita	11487	16801	5314*	13600	2112	11505	18	12028	540
	(3174)	(12407)	(3073)	(7133)	(1836)	(1604)	(1128)	(1285)	(1888)
Senate Dem.	0.72	0.65	-0.07	0.57	-0.15*	0.58	-0.14	0.65	-0.07
	(0.23)	(0.35)	(0.18)	(0.17)	(0.08)	(0.18)	(0.09)	(0.13)	(0.14)
House Dem.	0.69	0.68	-0.01	0.56	-0.13*	0.57	-0.12	0.53	-0.16
	(0.24)	(0.23)	(0.18)	(0.12)	(0.07)	(0.18)	(0.09)	(0.05)	(0.14)
N States	19	3		14		9		3	

Notes: Baseline characteristics of ever-ban states, by decade of adoption, and of never-ban states. Columns report pre-ban means with standard deviations in parentheses, and differences between each adoption cohort and never-ban states with standard errors in parentheses. All variables are measured at baseline in 1980. Offense rates are per 100 population aged 10–17. % Corp. Punishment is the share of students corporally punished, and Senate Dem. is an indicator for Democratic control of the state senate. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Effects on additional educational and labor-market outcomes

	(1) HS completion (ages 25–29)	(2) BA or more (ages 25–29)	(3) Log real wages (ages 25–29)
ATT	-0.012*	0.007	-0.022
	(0.007)	(0.004)	(0.021)
p-value	0.075	0.111	0.285
Magnitude (%)	-1.38	2.61	-2.22
Mean (pre-ban)	0.880	0.253	10.325
N	8,383	8,383	8,382

Notes: Callaway and Sant'Anna (2021) estimates of the effects of corporal punishment bans on high-school completion, BA or more, and log real wages at ages 25–29, using the cohort design (p = birth year, G = ban year -18). The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by cell population. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean in columns 1–2 and in log points ($\hat{\beta} \times 100$) in column 3. Standard errors, in parentheses, are clustered at the state level using a wild cluster bootstrap. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Robustness: treatment defined as first year of sustained zero corporal punishment

	(1) Arrest rate (ages 10–17)	(2) Arrest rate (ages 18–24)	(3) Yrs schooling (ages 25–29)	(4) Employed (ages 25–29)
ATT	-1.271***	-1.753*	-0.014	-0.003
	(0.331)	(0.944)	(0.031)	(0.009)
p-value	0.000	0.063	0.666	0.710
Magnitude (%)	-19.87	-15.92	-0.10	-0.43
Mean (pre-ban)	6.396	11.014	13.333	0.773
N	80,196	80,193	8,314	8,314

Notes: Treatment is dated by the earlier of the statutory ban year and the first Civil Rights Data Collection wave after which the state reports only zero students receiving corporal punishment (a de facto ban); where the sustained zero comes after the ban, the ban year is kept. This moves Colorado and Idaho, which banned in 2023, to 2013 and 2017, and treats Kansas and Wyoming, which have no ban, from 2017 and 2013. All other states keep their statutory ban year. Columns 1–2 use the county-year design (p = calendar year, G = ban year) and columns 3–4 the cohort design (p = birth year, G = ban year -18), as in Table 1. All estimates are Callaway and Sant'Anna (2021) estimates with not-yet-treated units as the comparison group and an event window from -10 to $+20$, weighted by age-band population (columns 1–2) and cell population (columns 3–4). Arrest rates are winsorized at the 95th percentile. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" is the estimate as a percentage of that mean. Standard errors, clustered at the state level with a wild bootstrap, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Robustness: placebo and triple difference-in-differences using untreated adults (ages 40–64)

	(1)	(2)	(3)
	Arrest rate	Triple difference	
	(ages 40–64)	Ages 10–17	Ages 18–24
ATT	0.011	-1.261***	-1.712**
	(0.109)	(0.274)	(0.870)
p-value	0.919	0.000	0.049
Magnitude (%)	0.57	-19.55	-15.66
Mean (pre-ban)	1.922	6.448	10.935
N	85,344	85,344	85,341

Notes: Column 1 estimates the effect of corporal punishment bans on the arrest rate at ages 40–64, a group that was at least 19 years old when the ban was enacted throughout the event window, and therefore untreated. Columns 2–3 estimate triple differences: the outcome is the difference between the arrest rate of the treated age group (10–17 or 18–24) and the arrest rate at ages 40–64 in the same county and year, which removes any county-year shock common to both age groups. All columns are Callaway and Sant'Anna (2021) estimates on the county-year design (p = calendar year, G = ban year), with not-yet-treated units as the comparison group and an event window from -10 to $+20$. Estimates are weighted by the population of the age group in column 1 and of the treated age group in columns 2–3. Arrest rates are winsorized at the 95th percentile. "Pre-ban average" is the average of the event-study coefficients at event times -10 to -2 (reference -1). "Mean" is the pre-ban mean arrest rate of the age group in column 1 and of the treated age group in columns 2–3, in treated units; "Magnitude" is the estimate as a percentage of that mean. Standard errors, clustered at the state level, are in parentheses (wild cluster bootstrap for the ATT, analytic for the pre-ban average). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Robustness: additional sample restrictions

	(1)	(2)	(3)	(4)	(5)	(6)
	School discipline		Arrest rates		Education and labor market	
	CP rate (K–12, in school)	Suspensions (K–12, in school)	Arrest rate (ages 10–17)	Arrest rate (ages 18–24)	Yrs schooling (ages 25–29)	Employed (ages 25–29)
Panel A. States that never changed the dropout age						
ATT	-1.502***	-0.163*	-0.844***	-0.043	-0.004	0.004
	(0.399)	(0.086)	(0.243)	(0.561)	(0.044)	(0.013)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.058	0.001	0.939	0.930	0.731
Magnitude (%)	-77.74	-15.02	-13.90	-0.45	-0.03	0.56
Mean (pre-ban)	0.593	5.084	6.069	9.419	13.456	0.778
N	63,420	68,490	37,715	37,715	4,372	4,372
Panel B. Excluding pre-Roe abortion-reform states						
ATT	-2.385***	-0.172**	-0.898**	-0.493	-0.019	-0.003
	(0.623)	(0.072)	(0.418)	(0.677)	(0.033)	(0.009)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.017	0.031	0.466	0.564	0.699
Magnitude (%)	-90.80	-15.76	-14.63	-4.81	-0.14	-0.44
Mean (pre-ban)	0.639	5.436	6.139	10.246	13.380	0.779
N	128,483	136,899	80,492	80,489	7,859	7,859
Panel C. Excluding counties adjacent to differently-treated states						
ATT	-2.225***	-0.132**	-1.026***	-1.344	-0.023	-0.008
	(0.299)	(0.057)	(0.364)	(1.030)	(0.042)	(0.010)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.021	0.005	0.192	0.579	0.435
Magnitude (%)	-89.20	-12.41	-15.99	-12.40	-0.17	-1.02
Mean (pre-ban)	0.515	5.309	6.416	10.843	13.315	0.774
N	118,031	125,760	75,102	75,102	7,547	7,547

Notes: The six main outcomes of Table 1, each re-estimated on a restricted sample. Columns 1–2 are the corporal punishment and suspension rates per 100 enrolled students, stacked PQMLE estimates; columns 3–4 are the arrest rates at ages 10–17 and 18–24, Callaway and Sant'Anna (2021) estimates on the county-year design (p = calendar year, G = ban year); columns 5–6 are years of schooling and employment at ages 25–29, Callaway–Sant'Anna estimates on the cohort design (p = birth year, G = ban year -18). Panel A keeps the 22 states that never changed their compulsory schooling (dropout) age; for the county-year outcomes (columns 1–4) the sample is further restricted to years through 2015, while the cohort outcomes (columns 5–6) are restricted by state only. Panel B drops the five states that liberalized abortion access before *Roe v. Wade* (1973): Alaska, California, Hawaii, New York, and Washington (Donohue and Levitt, 2001). Panel C drops counties adjacent to a county in a state with a different treatment status, using the Census County Adjacency File. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by enrollment (columns 1–2), by age-band population (columns 3–4) and by cell population (columns 5–6). Arrest rates are winsorized at the 95th percentile. "Estimator" gives the estimator used in each column, where "CS DiD" denotes Callaway and Sant'Anna (2021); "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta}) - 1) \times 100$ in columns 1–2 and as the ATT divided by the mean in columns 3–6. Standard errors, in parentheses, are clustered at the state level: analytic in columns 1–2 and a wild cluster bootstrap in columns 3–6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Robustness: arrest effects controlling for police force size

	(1)	(2)	(3)	(4)
	Arrest rate (ages 10–17)		Arrest rate (ages 18–24)	
Police control	No	Yes	No	Yes
ATT	-1.070***	-0.833**	-0.139	0.215
	(0.357)	(0.330)	(0.790)	(0.770)
Officers per 1,000		1.390***		2.154***
		(0.217)		(0.447)
Magnitude (%)	-16.59	-12.92	-1.29	2.01
Mean (pre-ban)	6.451	6.451	10.725	10.725
N	565,739	565,739	565,702	565,702

Notes: Stacked difference-in-differences estimates of the effect of corporal punishment bans on arrest rates per 100 individuals at ages 10–17 (columns 1–2) and 18–24 (columns 3–4). Each ban cohort forms a stack with never-treated counties and not-yet-treated counties observed before their own ban, over event times −10 to +20. All specifications are estimated by OLS with stack-by-county and stack-by-year fixed effects, weighted by age-band population. Columns 2 and 4 add sworn officers per 1,000 residents in the same county and year (FBI Law Enforcement Employees data, restricted to counties where reporting agencies cover more than 95% of the population, winsorized at the 95th percentile). All columns use the county-years with police data, so the control is the only difference between columns 1 and 2 and between columns 3 and 4. Arrest rates are winsorized at the 95th percentile. "Mean" is the pre-ban mean of the dependent variable in treated counties; "Magnitude" is the ATT as a percentage of that mean. N counts stacked county-years. Standard errors, in parentheses, are clustered at the state level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Robustness: arrest effects with different winsorization

	(1)	(2)	(3)	(4)
	Arrest rate (ages 10–17)		Arrest rate (ages 18–24)	
	Winsorized p99	Raw	Winsorized p99	Raw
ATT	-1.388***	-1.345***	-1.652*	-1.669*
	(0.399)	(0.427)	(0.952)	(0.951)
Magnitude (%)	-20.69	-19.94	-15.04	-15.16
p-value	0.001	0.002	0.083	0.079
Mean (pre-ban)	6.707	6.747	10.986	11.009
N	85,344	85,344	85,341	85,341

Notes: Callaway and Sant'Anna (2021) estimates of the effects of corporal punishment bans on arrest rates per 100 individuals at ages 10–17 (columns 1–2) and 18–24 (columns 3–4), under alternative treatments of the dependent variable, using the county-year design (p = calendar year, G = ban year). The arrest rate is winsorized at the 99th percentile in columns 1 and 3 and left unwinsorized in columns 2 and 4; the main specification winsorizes at the 95th percentile. As in the main specification, the rate at ages 10–17 is winsorized over the pooled sample and the rate at ages 18–24 within each year. The comparison group is not-yet-treated units and the event window runs from −10 to +20. Estimates are weighted by age-band population. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean. Standard errors, in parentheses, are clustered at the state level using a wild cluster bootstrap. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Robustness: effects using the never-treated comparison group

	(1) Arrest rate (ages 10–17)	(2) Arrest rate (ages 18–24)	(3) Yrs schooling (ages 25–29)	(4) Employed (ages 25–29)
ATT	-1.240***	-1.677*	-0.015	-0.003
	(0.331)	(0.963)	(0.032)	(0.009)
Estimator	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.081	0.643	0.749
Magnitude (%)	-18.98	-15.20	-0.11	-0.37
Mean (pre-ban)	6.534	11.034	13.315	0.771
N	84,181	84,178	7,785	7,785

Notes: Callaway and Sant'Anna (2021) estimates of the effects of corporal punishment bans on the four outcomes estimated with the Callaway–Sant'Anna estimator, using only never-treated states as the comparison group; the main estimates use not-yet-treated units, so those results are not repeated here. Columns 1–2 are the arrest rates at ages 10–17 and 18–24, on the county-year design (p = calendar year, G = ban year); columns 3–4 are years of schooling and employment at ages 25–29, on the cohort design (p = birth year, G = ban year $-$ 18). The event window runs from -10 to $+20$. Estimates are weighted by age-band population (columns 1–2) and by cell population (columns 3–4). Arrest rates are winsorized at the 95th percentile. "Estimator" gives the estimator used in each column, where "CS DiD" denotes Callaway and Sant'Anna (2021); "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean. Standard errors, in parentheses, are clustered at the state level using a wild cluster bootstrap. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Robustness: effects under varying event windows

	(1)	(2)	(3)	(4)	(5)	(6)
	School discipline		Arrest rates		Education and labor market	
	CP rate (K–12, in school)	Suspensions (K–12, in school)	Yrs schooling (ages 25–29)	Employed (ages 25–29)	Arrest rate (ages 10–17)	Arrest rate (ages 18–24)
Panel A. Event window [-5,+15]						
ATT	-2.260***	-0.113	0.001	-0.001	-1.075***	-1.424*
	(0.274)	(0.083)	(0.027)	(0.009)	(0.287)	(0.804)
Event window	[-5,+15]	[-5,+15]	[-5,+15]	[-5,+15]	[-5,+15]	[-5,+15]
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.176	0.979	0.911	0.000	0.076
Magnitude (%)	-89.56	-10.68	0.01	-0.13	-16.49	-11.99
Mean (pre-ban)	0.400	5.252	13.367	0.763	6.519	11.876
N	97,297	106,872	7,545	7,545	74,948	74,945
Panel B. Event window [-8,+15]						
ATT	-2.166***	-0.139**	0.001	-0.001	-1.052***	-1.404*
	(0.259)	(0.062)	(0.027)	(0.009)	(0.279)	(0.797)
Event window	[-8,+15]	[-8,+15]	[-8,+15]	[-8,+15]	[-8,+15]	[-8,+15]
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.025	0.982	0.913	0.000	0.078
Magnitude (%)	-88.54	-12.94	0.00	-0.13	-16.23	-12.39
Mean (pre-ban)	0.519	5.322	13.330	0.769	6.484	11.331
N	114,868	123,623	7,916	7,916	78,698	78,695
Panel C. Event window [-15,+25]						
ATT	-2.445***	-0.145**	-0.028	-0.001	-1.296***	-1.810*
	(0.313)	(0.062)	(0.032)	(0.008)	(0.360)	(1.063)
Event window	[-15,+25]	[-15,+25]	[-15,+25]	[-15,+25]	[-15,+25]	[-15,+25]
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.020	0.386	0.866	0.000	0.089
Magnitude (%)	-91.33	-13.46	-0.21	-0.18	-19.91	-17.04
Mean (pre-ban)	0.811	5.361	13.279	0.774	6.508	10.623
N	159,933	168,132	8,981	8,981	94,363	94,360

Notes: Table 1 re-estimated with varying event windows. Columns 1–2 report stacked PQMLE estimates for the corporal punishment and suspension rates per 100 enrolled students; columns 3–4 report Callaway and Sant'Anna (2021) estimates for years of schooling and employment at ages 25–29, using the cohort design; columns 5–6 report Callaway and Sant'Anna (2021) estimates for arrest rates per 100 individuals, using the county-year design. Estimates are weighted by enrollment (columns 1–2), by cell population (columns 3–4) and by age-band population (columns 5–6). Arrest rates are winsorized at the 95th percentile. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta}) - 1) \times 100$ in columns 1–2 and as the ATT divided by the mean in columns 3–6. Standard errors, in parentheses, are clustered at the state level: analytic in columns 1–2 and a wild cluster bootstrap in columns 3–6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Robustness: effects for the period 1978 to 1994

	(1)	(2)	(3)	(4)	(5)	(6)
	School discipline		Arrest rates		Education and labor market	
	CP rate (K–12, in school)	Suspensions (K–12, in school)	Yrs schooling (ages 25–29)	Employed (ages 25–29)	Arrest rate (ages 10–17)	Arrest rate (ages 18–24)
ATT	-2.249***	-0.166***	-0.025	-0.003	-1.342***	-1.880*
	(0.292)	(0.062)	(0.041)	(0.012)	(0.313)	(0.976)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.007	0.541	0.799	0.000	0.054
Magnitude (%)	-89.45	-15.30	-0.19	-0.40	-20.61	-17.18
Mean (pre-ban)	0.522	5.322	13.233	0.774	6.510	10.943
N	106,073	112,942	6,460	6,460	75,674	75,668

Notes: Table 1 re-estimated for the period 1978–1994, matching the available data of column (1). Columns 1–2 report stacked PQMLE estimates for the corporal punishment and suspension rates per 100 enrolled students; columns 3–4 report Callaway and Sant'Anna (2021) estimates for years of schooling and employment at ages 25–29, using the cohort design (p = birth year, G = ban year $-$ 18); columns 5–6 report Callaway–Sant'Anna estimates for arrest rates per 100 individuals, using the county-year design, where column 5 is the juvenile arrest rate at ages 10–17 and column 6 is the arrest rate at ages 18–24. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by enrollment (columns 1–2), by age-band population (columns 5–6) and by cell population (columns 3–4). Arrest rates are winsorized at the 95th percentile. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta}) - 1) \times 100$ in columns 1–2 and as the ATT divided by the mean in columns 3–6. Standard errors, in parentheses, are clustered at the state level: analytic in columns 1–2 and a wild cluster bootstrap in columns 3–6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Robustness: effects using alternative estimators

	(1)	(2)	(3)	(4)	(5)	(6)
	School discipline		Arrest rates		Education and labor market	
	CP rate (K–12, in school)	Suspensions (K–12, in school)	Yrs schooling (ages 25–29)	Employed (ages 25–29)	Arrest rate (ages 10–17)	Arrest rate (ages 18–24)
Panel A: Two-way fixed effects						
ATT	-2.302***	-0.145**	-0.034	0.005	-0.886**	-0.119
	(0.280)	(0.065)	(0.027)	(0.004)	(0.433)	(0.743)
Magnitude (%)	-90.0	-13.5	-0.3	0.6	-13.7	-1.1
N	11,110	12,518	11,168	11,168	87,737	87,734
Panel B: Heterogeneity-robust (Wooldridge)						
ATT	-0.202***	-1.046***	-0.047**	0.005	-1.283***	-0.734*
	(0.004)	(0.258)	(0.019)	(0.006)	(0.227)	(0.376)
Magnitude (%)	-87.9	-15.7	-0.4	0.7	-19.9	-6.7
N	11,043	12,608	10,461	10,461	87,244	87,241
Panel C: Borusyak, Jaravel & Spiess imputation						
ATT			-0.049***	0.006	-1.289***	-0.787**
			(0.017)	(0.006)	(0.221)	(0.378)
Magnitude (%)			-0.4	0.8	-20.0	-7.2
N			10,461	10,461	87,244	87,241

Notes: The six main outcomes of Table 1, re-estimated with alternative difference-in-differences estimators. Panel A reports two-way fixed effects, Poisson for the corporal punishment and suspension rates and linear for the remaining outcomes. Panel B reports the heterogeneity-robust Wooldridge (2025) estimator, again Poisson for the corporal punishment and suspension rates and linear otherwise. Panel C reports the imputation estimator of Borusyak et al. (2024), which applies only to the linear outcomes. The corporal punishment and suspension rates and both arrest outcomes use the county-year design (p = calendar year, treatment = post-ban indicator); years of schooling and employment use the cohort design (p = birth year, G = ban year -18). Column 6 is the arrest rate at ages 18–24. The event window runs from -10 to $+20$, and cohorts treated before the start of the window are dropped. Estimates are weighted by enrollment (corporal punishment and suspension rates), by age-band population (both arrest outcomes) and by cell population (cohort outcomes). Arrest rates are winsorized at the 95th percentile. "Magnitude" expresses the estimate as a percentage of the pre-ban mean in treated units, computed as $(\exp(\hat{\beta}) - 1) \times 100$ for Poisson TWFE, as the average partial effect relative to the implied counterfactual mean for Poisson Wooldridge, and as the ATT divided by the mean for the linear outcomes. Standard errors, in parentheses, are analytic and clustered at the state level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Robustness: effects including socio-economic controls

	(1)	(2)	(3)	(4)	(5)	(6)
	School discipline		Arrest rates		Education and labor market	
	CP rate (K–12, in school)	Suspensions (K–12, in school)	Arrest rate (ages 10–17)	Arrest rate (ages 18–24)	Yrs schooling (ages 25–29)	Employed (ages 25–29)
ATT	-1.899***	-0.062	-1.025***	-0.943	-0.018	0.002
	(0.342)	(0.061)	(0.376)	(1.376)	(0.056)	(0.009)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.311	0.006	0.493	0.747	0.814
Magnitude (%)	-85.03	-5.99	-16.36	-8.62	-0.14	0.28
Mean (pre-ban)	0.498	5.457	6.264	10.948	13.315	0.771
N	72,193	76,564	78,419	77,745	7,297	7,297

Notes: The six main outcomes of Table 1, re-estimated with socio-economic controls. Columns 1–2 report stacked PQMLE estimates for the corporal punishment and suspension rates per 100 enrolled students; columns 3–4 report Callaway and Sant'Anna (2021) estimates for arrest rates per 100 individuals, using the county-year design, where column 3 is the arrest rate at ages 10–17 and column 4 is the arrest rate at ages 18–24; columns 5–6 report Callaway–Sant'Anna estimates for years of schooling and employment at ages 25–29, using the cohort design (p = birth year, G = ban year $-$ 18). The controls are GDP per capita, the SNAP recipient share, the Black population aged 0–18 and the 1980 share without a high-school degree; the three time-varying controls enter at their value in the year before the ban in columns 1–4, and as a 1980–1985 county average in columns 5–6. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by enrollment (columns 1–2), by age-band population (columns 3–4) and by cell population (columns 5–6). Arrest rates are winsorized at the 95th percentile. "Estimator" gives the estimator used in each column, where "CS DiD" denotes Callaway and Sant'Anna (2021); "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta}) - 1) \times 100$ in columns 1–2 and as the ATT divided by the mean in columns 3–6. Standard errors, in parentheses, are clustered at the state level: analytic in columns 1–2 and a wild cluster bootstrap in columns 3–6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Robustness: effects excluding one region at a time

	(1)	(2)	(3)	(4)
Panel A. Corporal punishment rate				
ATT	-2.224***	-2.401***	-1.610***	-2.384***
	(0.295)	(0.300)	(0.305)	(0.591)
Excl.	Northeast	Midwest	South	West
Magnitude (%)	-89.19	-90.94	-80.01	-90.78
p-value	0.000	0.000	0.000	0.000
Mean (pre-ban)	0.579	0.435	0.469	0.586
N	126,805	99,315	54,335	117,843
Panel C. Arrest rates (ages 10–17)				
ATT	-1.372***	-1.166***	-0.737*	-1.090***
	(0.348)	(0.397)	(0.416)	(0.399)
Excl.	Northeast	Midwest	South	West
Magnitude (%)	-20.30	-17.70	-11.09	-19.09
p-value	0.000	0.003	0.077	0.006
Mean (pre-ban)	6.757	6.589	6.644	5.710
N	80,634	60,039	40,546	73,430
Panel E. Years of schooling (ages 25–29)				
ATT	-0.017	0.014	-0.070*	-0.005
	(0.041)	(0.034)	(0.040)	(0.036)
Excl.	Northeast	Midwest	South	West
Magnitude (%)	-0.13	0.11	-0.53	-0.04
p-value	0.682	0.683	0.080	0.878
Mean (pre-ban)	13.193	13.317	13.332	13.398
N	6,874	6,725	3,276	7,439

	(1)	(2)	(3)	(4)
Panel B. Suspension rate				
ATT	-0.132**	-0.137*	-0.096	-0.208***
	(0.061)	(0.070)	(0.083)	(0.066)
Excl.	Northeast	Midwest	South	West
Magnitude (%)	-12.34	-12.77	-9.15	-18.81
p-value	0.030	0.052	0.250	0.002
Mean (pre-ban)	5.433	5.375	5.112	5.453
N	134,920	102,608	62,263	124,883
Panel D. Arrest rates (ages 18–24)				
ATT	-2.134**	-1.938*	-2.115*	-0.511
	(1.022)	(1.165)	(1.210)	(0.675)
Excl.	Northeast	Midwest	South	West
Magnitude (%)	-18.69	-16.67	-19.26	-5.30
p-value	0.037	0.096	0.081	0.449
Mean (pre-ban)	11.416	11.627	10.981	9.642
N	80,631	60,036	40,546	73,427
Panel F. Employment (ages 25–29)				
ATT	-0.001	0.003	-0.008	-0.003
	(0.012)	(0.009)	(0.010)	(0.010)
Excl.	Northeast	Midwest	South	West
Magnitude (%)	-0.15	0.38	-1.10	-0.44
p-value	0.921	0.743	0.406	0.720
Mean (pre-ban)	0.773	0.765	0.770	0.777
N	6,874	6,725	3,276	7,439

Notes: The six main outcomes of Table 1, re-estimated excluding one census region at a time, with regions as defined by the U.S. Census Bureau and the excluded region given in the "Excl." row. Panels A and B are the corporal punishment and suspension rates per 100 enrolled students, stacked PQMLE estimates; Panels C and D are the arrest rates per 100 individuals at ages 10–17 and 18–24, using the county-year design; Panels E and F are years of schooling and employment at ages 25–29, using the cohort design. Panels C–F are Callaway and Sant'Anna (2021) estimates. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by enrollment (Panels A–B), by age-band population (Panels C–D) and by cell population (Panels E–F). Arrest rates are winsorized at the 95th percentile. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta}) - 1) \times 100$ in Panels A–B and as the ATT divided by the mean in Panels C–F. Standard errors, in parentheses, are clustered at the state level: analytic in Panels A–B and a wild cluster bootstrap in Panels C–F. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A16: Robustness: arrest effects in a balanced panel and in the counties available in IPUMS data

	(1) Arrest rate (ages 10–17)	(2) Arrest rate (ages 18–24)	(3) Arrest rate (ages 10–17)	(4) Arrest rate (ages 18–24)
ATT	-1.039**	-2.288	-1.275***	-2.077**
	(0.480)	(1.412)	(0.333)	(0.952)
Sample	Balanced UCR	Balanced UCR	IPUMS counties	IPUMS counties
p-value	0.030	0.105	0.000	0.029
Magnitude (%)	-15.64	-20.26	-18.46	-18.47
Mean (pre-ban)	6.645	11.295	6.907	11.249
Counties	350	350	484	484
N	11,265	11,265	14,292	14,292

Notes: Callaway and Sant'Anna (2021) estimates of the effects of corporal punishment bans on arrest rates per 100 individuals, using the county-year design (p = calendar year, G = ban year), on two restricted samples given in the "Sample" row. Columns 1–2 ("Balanced UCR") keep counties with arrest data in every year of their estimation window: every year from 10 years before to 20 years after the ban (within 1974–2019) for counties in banning states, and every year from 1974 to 2019 for counties in never-banning states. This sample has 350 counties in 31 states (302 treated counties in 21 states and 48 never-treated counties in 10 states), and the same counties are used in both columns. Columns 3–4 ("IPUMS counties") keep the counties identified in the IPUMS data, which we use to estimate the effects on years of schooling and employment. Columns 1 and 3 are arrests at ages 10–17; columns 2 and 4 are arrests at ages 18–24. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by age-band population. Arrest rates are winsorized at the 95th percentile. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean. "Counties" is the number of counties in the estimation sample. Standard errors, in parentheses, are clustered at the state level using a wild cluster bootstrap. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A17: Robustness: state-level aggregation

	(1)	(2)	(3)	(4)	(5)	(6)
	School discipline		Arrest rates		Education and labor market	
	CP rate (K–12, in school)	Suspensions (K–12, in school)	Arrest rate (ages 10–17)	Arrest rate (ages 18–24)	Yrs schooling (ages 25–29)	Employed (ages 25–29)
ATT	-2.323***	-0.148**	-1.405***	-1.376	-0.056*	-0.005
	(0.295)	(0.063)	(0.358)	(0.857)	(0.030)	(0.009)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.018	0.000	0.108	0.063	0.592
Magnitude (%)	-90.20	-13.75	-21.83	-12.52	-0.42	-0.63
Mean (pre-ban)	0.431	5.910	6.436	10.996	13.315	0.771
N	2,802	2,803	1,448	1,603	921	921

Notes: Table 1 re-estimated at the state-year level instead of the county-year level. Columns 1–2 report stacked PQMLE estimates for the corporal punishment and suspension rates per 100 enrolled students, using the state CRDC series for 1976–1994 and bans enacted in 1974–1994; columns 3–4 report Callaway and Sant'Anna (2021) estimates for arrest rates per 100 individuals, using the state-year design, where column 3 is the arrest rate at ages 10–17 and column 4 is the arrest rate at ages 18–24 (county arrests and populations summed to the state-year); columns 5–6 report Callaway–Sant'Anna estimates for years of schooling and employment at ages 25–29, using the cohort design (p = birth year, G = ban year $-$ 18). The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Estimates are weighted by enrollment (columns 1–2), by age-band population (columns 3–4) and by cell population (columns 5–6). Arrest rates are winsorized at the 95th percentile. "Mean" is the pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta}) - 1) \times 100$ in columns 1–2 and as the ATT divided by the mean in columns 3–6. Standard errors, in parentheses, are clustered at the state level: analytic in columns 1–2 and a wild cluster bootstrap in columns 3–6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Robustness: effects without population weights

	School discipline		Arrest rates		Education and labor market	
	(1) CP rate (K–12)	(2) Suspensions (K–12)	(3) Ages (10–17)	(4) Ages (18–24)	(5) Yrs schooling (25–29)	(6) Employed (25–29)
ATT	-2.692***	-0.241***	-0.637***	0.228	-0.043	-0.005
	(0.600)	(0.054)	(0.212)	(0.665)	(0.034)	(0.013)
Estimator	PQMLE	PQMLE	CS DiD	CS DiD	CS DiD	CS DiD
p-value	0.000	0.000	0.003	0.731	0.201	0.665
Magnitude (%)	-93.23	-21.44	-13.81	2.35	-0.33	-0.70
Mean (pre-ban)	0.746	3.859	4.612	9.720	13.188	0.777
N	132,779	141,558	85,344	85,341	8,383	8,383

Notes: Table 1 re-estimated without population or enrollment weights; estimators and samples are otherwise identical. Columns 1–2 are stacked PQMLE estimates for the corporal punishment and suspension rates per 100 enrolled students. Columns 3–6 are Callaway and Sant'Anna (2021) estimates: columns 3–4 use the county-year design for arrest rates at ages 10–17 and 18–24, and columns 5–6 use the cohort design for years of schooling and employment at ages 25–29. The comparison group is not-yet-treated units and the event window runs from -10 to $+20$. Arrest rates are winsorized at the 95th percentile. "Estimator" gives the estimator used in each column, where "CS DiD" denotes Callaway and Sant'Anna (2021); "Mean" is the unweighted pre-ban mean of the dependent variable in treated units; "Magnitude" expresses the estimate as a percentage of that mean, computed as $(\exp(\hat{\beta}) - 1) \times 100$ in columns 1–2 and as the ATT divided by the mean in columns 3–6. Standard errors, in parentheses, are clustered at the state level: analytic in columns 1–2 and a wild cluster bootstrap in columns 3–6. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A19: Upper bound on the direct effect of corporal punishment on juvenile arrests, by assumed odds ratio

		Pre-ban arrest rate (per 100)		Direct effect	Share of the effect (%)	
Odds ratio	Source	Not punished	Punished	(per 100)	Direct	Indirect
1.50	Main text, lower bound	6.25	9.38	-0.190	16	84
1.92	Gershoff (2002), aggression	6.10	11.72	-0.342	28	72
2.00	Main text, upper bound	6.08	12.15	-0.370	30	70
2.14	Gershoff (2002), antisocial behavior	6.03	12.91	-0.419	34	66
3.00	Large effect	5.75	17.24	-0.699	57	43

Notes: Decomposition of the effect of corporal punishment bans on arrests at ages 10–17 (ATT of -1.227 per 100, Table 1, column 3) into a direct effect on pupils who were themselves corporally punished and an indirect effect on all other pupils. In treated states before the ban, 0.52 in every 100 students were corporally punished each year; assuming distinct pupils are punished each year, a share $p = 1 - (1 - 0.52/100)^{12} \approx 6.1\%$ is punished at least once over 12 school years, which maximizes the direct channel. Punished pupils are assumed to be arrested at a rate equal to the odds ratio times that of unpunished pupils, holding the pre-ban mean of 6.45 arrests per 100 fixed: the unpunished rate is $\bar{y}/(1 + p(\text{OR} - 1))$ and the punished rate is OR times that. The direct effect assumes the ban closes this gap for the punished share, $p \times (\text{rate}_{\text{not punished}} - \text{rate}_{\text{punished}})$; the direct share divides it by the ATT, and the indirect share is the remainder. The odds ratios of 1.92 and 2.14 convert the standardized effects of corporal punishment on aggression ($d = 0.36$) and antisocial behavior ($d = 0.42$) in Gershoff (2002) with $\text{OR} = \exp(d\pi/\sqrt{3})$. Treating the odds ratio as a ratio of arrest rates slightly overstates the direct share, consistent with an upper bound.

B Insights from surveying teachers

To further understand corporal punishment in U.S. schools and the implementation of corporal punishment bans, we ran a survey among 162 current and former U.S. K–12 teachers. The survey asked respondents about their teaching career, the corporal punishment regime, and their perceptions on the effectiveness of corporal punishment, among other questions. The design and sampling frame are described in Section 3. A transcript of the survey is available upon request.

Table C1 describes the sample and our main variables of interest. Column 1 covers all 162 respondents; 122 taught in a single state and can therefore be matched to a single ban status, covering 23 states. Columns 2–4 split single-state respondents by whether they worked in states that never banned corporal punishment, states that banned in the 1990s, and states that banned from 2000 onwards. Our sample is in its majority female and the average age is 57.

First, the survey confirms that corporal punishment is reasonably prevalent where it is still permitted. Among respondents working in no-ban states, 21% respond that their colleagues have used corporal punishment, and 6% that they have used it themselves. When asked what proportion of their colleagues might use corporal punishment, teachers state about 12%.

Second, the empirical result that bans virtually abolished corporal punishment use finds support in the survey. Among 24 teachers who worked while bans were implemented only 3 reported that they witnessed corporal punishment after the bans, reassuring us in the validity of the first stage. The survey also suggests there were some changes in behavior in response to the bans as most teachers state they used alternative forms of punishment. Detention (28.6%) and loss of recess or privileges (31.0%) were named most often, followed by suspension (23.8%). Extra work (7.1%), expulsion (4.8%) and isolation (4.8%) are rare.

Third, comparing teachers in ban and no ban states with regards to perceptions on corporal punishment shows substantial differences. Among teachers who taught only after the ban, 60.8% strongly disagree that corporal punishment helps students. The opposite

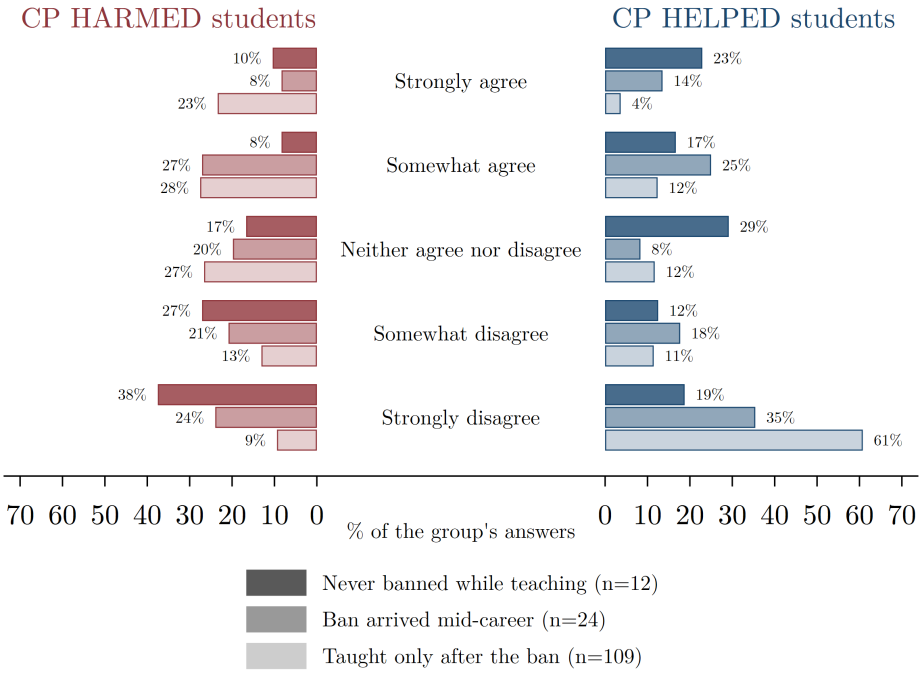
Table C1: Teacher survey: sample characteristics

	(1) All respondents	(2) Single-state respondents, by their state's ban		
		Never banned	1990s	2000s–
Respondents (N)	162	67	17	38
States covered (N)	23	12	7	4
Demographics and career				
Female (%)	75.9	76.1	70.6	71.1
Age in 2026	57.4 (9.0)	55.4 (8.3)	55.5 (5.4)	57.9 (8.5)
Year of birth	1969	1971	1970	1968
First year teaching	2001	2004	2006	2000
Last year teaching	2019	2020	2024	2017
Career length (years)	18.1 (11.5)	16.3 (10.1)	17.7 (12.5)	17.0 (12.4)
Still teaching in 2026 (%)	48.1	53.7	70.6	42.1
School characteristics				
School location				
Urban (%)	22.8	20.9	23.5	31.6
Suburban (%)	52.5	53.7	58.8	39.5
Rural (%)	24.7	25.4	17.6	28.9
Student body				
Predominantly white (%)	41.4	31.3	41.2	50.0
Roughly balanced (%)	40.1	49.3	29.4	36.8
Predominantly students of color (%)	17.9	19.4	23.5	13.2
Free / reduced-price lunch share (%)	55.4 (31.3)	56.3 (30.1)	52.5 (36.2)	55.3 (32.0)
Exposure to corporal punishment				
Experienced or witnessed CP as a pupil (%)	63.6	64.2	47.1	63.2
Used CP themselves (%)	3.7	6.0	0.0	2.6
Colleagues used CP (%)	16.2	21.2	5.9	10.5
Estimated School staff who used CP (%)	7.6 (16.6)	12.1 (22.9)	4.1 (12.0)	4.7 (8.5)
When the ban arrived: self-reported				
Taught only without ban (%)	8.3	12.9	0.0	3.2
Ban arrived mid-career (%)	16.6	17.7	6.2	16.1
Taught only after the ban (%)	75.2	69.4	93.8	80.6
When the ban arrived: state ban record				
State ban after their career (%)	60.7	100.0	0.0	18.4
State ban during their career (%)	17.2	0.0	11.8	50.0
State ban predates their career (%)	22.1	0.0	88.2	31.6
Teaching in the year of the state ban (%)	38.2	—	11.8	50.0
Reports a ban with no state ban in force (%)	66.1	87.1	0.0	58.1

Notes: Characteristics of the teacher survey sample. Column 1 covers all 162 respondents; columns 2–4 cover the 122 who taught in exactly one state, grouped by when that state banned corporal punishment, and exclude the 40 who taught in more than one. Continuous variables report means, with standard deviations in parentheses; year variables report means only. Percentages are shares of the respondents answering the item in that column. "States covered" counts distinct states among single-state respondents. Student-body shares do not sum to 100 because "not sure / other" remains in the denominator. "Staff at the school who used CP" is the respondent's own estimate of the share of teachers at their school who used it, asked of all respondents. The two regime blocks are alternative measures of when the ban arrived: the self-reported block is the respondent's answer for their own school, which can reflect a district- or school-level prohibition; the state ban record block asks only whether a state ban year falls inside their career. "Teaching in the year of the state ban" is over respondents whose state banned while they actively taught (55 in column 1), and is undefined where the state never banned. "Reports a ban with no state ban in force" is over the single-state 109 respondents.

pattern emerges among teachers in states where corporal punishment is permitted, where only 18.8% strongly disagree that it helps students. Teachers who were teaching when a ban arrived fall in between these two views. Perception on whether corporal punishment deters misbehavior and whether it should be allowed follow the same pattern (Figure C2, Panels a and b). Teachers who only taught after the ban see little merit in it: 63% disagree that it deters unruly behavior, and only 8% would permit it if the decision were theirs. Teachers who taught while corporal punishment was still allowed see considerably more value in it: 47% believe it deters unruly behavior, and 31% would permit it.

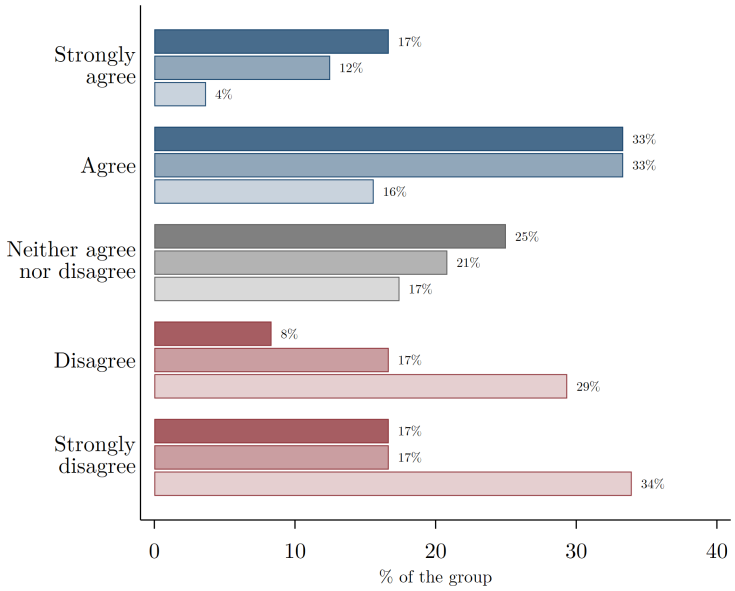
Figure C1: Perceived effects of corporal punishment, by exposure to the ban as teacher



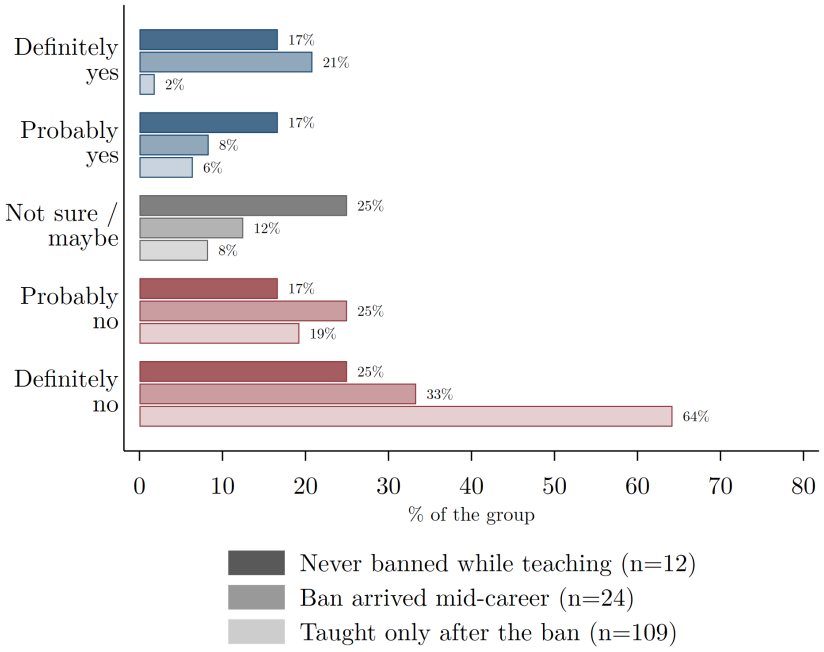
Notes: Beliefs about the effects of corporal punishment, on a five-point agree scale, among all respondents. Left: "I feel that using corporal punishment makes students..." act more aggressively towards their peers / towards authority figures, get suspended more often, act out more outside of school. Right: "I feel that using corporal punishment helps students..." focus better in the classroom, maintain discipline and act out less in class, achieve better test scores, respect authority more. Each side pools its four items, so a group's five bars sum to 100% on each side. Groups follow the self-reported regime question: corporal punishment permitted throughout the teacher's time at the school (darkest), prohibited while they were teaching there (medium), or prohibited before they began (lightest). Respondents answering "I am not sure" are excluded. Group sizes are numbers of teachers.

Figure C2: Perspectives on corporal punishment by exposure to the ban as teacher

(a) Believes that corporal punishment effectively deters unruly behavior



(b) Would permit corporal punishment if it were their decision



Notes: Teachers' attitudes to corporal punishment, split by their exposure to a ban during their own career. Both items are asked of every respondent, not of a regime group. Panel (a): "Do you agree that corporal punishment effectively deters unruly behavior in children and adolescents?" on the five-point agree scale. Panel (b): "If it were your decision, would you permit the use of corporal punishment to discipline students?" on the five-point scale from "definitely no" to "definitely yes". In both panels maroon marks the two answers least favourable to corporal punishment and navy the two most favourable, with the middle category in grey, and a group's five bars sum to 100%. Groups follow the self-reported regime question, as in Figure C1: corporal punishment permitted throughout the teacher's time at the school (darkest), prohibited while they were teaching there (medium), or prohibited before they began (lightest); respondents answering "I am not sure" are excluded. The two panels share the legend shown under panel (b). Group sizes are numbers of teachers.