Counting Children Differently: How Differential Response Manufactures Apparent Declines in Child Maltreatment

APEP Autonomous Research* @ai1scl

April 2, 2026

Abstract

Between 2000 and 2014, officially reported child maltreatment victims in the United States fell from 862,000 to 702,000, a decline widely interpreted as progress. Over the same period, 32 states adopted Differential Response (DR), which diverts low-risk Child Protective Services referrals from investigation to family assessment—a track that produces no substantiation finding and generates no entry in federal victim statistics. Using a staggered difference-in-differences design with Callaway–Sant'Anna estimation across 50 states (2000–2014), I find DR adoption reduces measured victim rates by approximately 0.25 per 1,000 children (standardized effect: −0.09), though the estimate is statistically insignificant in a 50-state panel. Meanwhile, referral volumes remain stable and child maltreatment fatalities—which are always investigated regardless of DR status—rose 28%. The pattern is consistent with administrative reclassification manufacturing a statistical decline without corresponding improvements in child safety.

JEL Codes: I38, H75, J13

Keywords: child maltreatment, differential response, measurement error, child protective services, NCANDS

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 27m).

1. Introduction

In 2023, the Administration for Children and Families reported 546,000 substantiated child maltreatment victims in the United States, down from over one million in the mid-1990s. Policymakers and child welfare advocates routinely cite this decline as evidence that prevention efforts are working (Jones et al., 2006; Finkelhor and Jones, 2010). But what if much of this progress is an accounting artifact?

This paper investigates a specific mechanism through which official maltreatment statistics can decline without any corresponding improvement in children's safety: the adoption of Differential Response (DR) by state child protective services agencies. Traditional CPS systems funnel every screened-in referral through a formal investigation that culminates in a substantiation decision—a finding that becomes a data point in the National Child Abuse and Neglect Data System (NCANDS). Beginning with Florida in 1993, states progressively adopted DR systems that create a second "family assessment" track for low-to-moderate-risk referrals. Crucially, assessment-track cases produce no substantiation finding, identify no perpetrator, and generate no entry in NCANDS victim statistics. By 2015, 32 states had implemented some form of DR.

The measurement consequence is mechanical: when a state adopts DR, a share of referrals that would previously have been investigated—and potentially substantiated—are instead diverted to the assessment track, where they vanish from official statistics regardless of whether maltreatment actually occurred. I exploit the staggered adoption of DR across states to estimate this measurement artifact using a difference-in-differences framework.

The empirical design proceeds in three steps. First, using state-year panel data from the Kids Count Data Center and ACF Child Maltreatment reports (2000–2014), I estimate the effect of DR adoption on substantiated victim rates using both two-way fixed effects and the heterogeneity-robust Callaway–Sant'Anna estimator (Callaway and Sant'Anna, 2021). The Callaway–Sant'Anna estimate implies DR adoption reduces measured victim rates by approximately 0.25 per 1,000 children, a standardized effect of -0.09. This estimate is statistically insignificant ($p = 0.56$), reflecting the limited power of a 50-state panel to detect modest measurement effects. However, the state-level regressions are only one element of a broader evidentiary pattern.

Second, I decompose the national trend into its constituent parts. Between 2000 and 2014, referrals to CPS agencies increased from 2.86 million to 3.25 million, while substantiated victims fell from 862,000 to 702,000. The victim-to-referral ratio collapsed from 0.30 to 0.22. This pattern—rising referrals, falling victims—is precisely what administrative reclassification predicts.

Third, I conduct a falsification test using child maltreatment fatalities, which are always investigated regardless of DR status. If the victim decline reflected genuine progress in child safety, fatalities should fall as well. Instead, child maltreatment fatalities rose from 1,236 in 2000 to 1,580 in 2014, a 28 percent increase. The divergence between falling victim counts and rising fatalities is consistent with a measurement artifact rather than a real decline in maltreatment.

This paper contributes to the literature on administrative data quality and its implications for policy evaluation. A growing body of work recognizes that measured outcomes can change because the measurement system changes, not because the underlying phenomenon changes. Levitt (1998) documented how incentives to reclassify crimes distort official crime statistics. Jacob and Levitt (2003) identified data manipulation in educational testing. Autor and Duggan (2006) showed how changes in disability screening criteria mechanically altered disability rolls. In the child welfare context, Fluke et al. (2019) examined re-reporting patterns under DR, and descriptive analyses have noted the temporal coincidence of DR adoption and declining victim counts (Merkel-Holguin et al., 2006; National Quality Improvement Center on Differential Response, 2014). However, no prior study provides a causal estimate of DR's effect on measured maltreatment rates using the staggered adoption as identification.

The broader implication is that the apparent 47 percent decline in child maltreatment since the mid-1990s is at least partly a statistical illusion produced by changing how we count, not what we count. Any researcher using NCANDS data to evaluate child welfare interventions must account for DR adoption as a confounder. And any policymaker interpreting falling victim rates as evidence of progress should consider whether the measurement system has changed more than the phenomenon it purports to measure.

2. Institutional Background

Traditional CPS and the Investigation Track. In the traditional child protective services model, every referral that passes initial screening enters the investigation track. A caseworker determines whether maltreatment occurred, identifies the perpetrator, and renders a substantiation decision. Substantiated cases are reported to NCANDS, the federal data system administered by the Children's Bureau, which aggregates state reports into the annual *Child Maltreatment* publication. The victim count in this publication is the primary statistic used to track national maltreatment prevalence.

Differential Response. Starting with Florida (1993) and Missouri (1994), states began implementing Differential Response, also called Alternative Response or Multiple Response.

DR creates a second track—"family assessment"—for referrals classified as low-to-moderate risk. The assessment track emphasizes family engagement, service provision, and voluntary participation rather than investigation and substantiation. The critical difference for measurement is that assessment-track cases *do not* produce substantiation findings, *do not* identify perpetrators, and *are not* reported as victims in NCANDS.

Adoption Timeline. DR adoption was staggered across states over more than two decades. After the early adopters (Florida 1993, Missouri 1994, Virginia 1999), a wave of states adopted in the mid-2000s (Minnesota 2004, North Carolina 2005, Tennessee 2005, Louisiana 2006, Ohio 2006), and a second wave followed in 2008–2015 (Colorado, Hawaii, Vermont, Arkansas, Indiana, Maine, New York, Oregon, Connecticut, Illinois, Maryland, Pennsylvania, Wisconsin, Arizona, Massachusetts, Michigan, Montana). By 2015, 32 states had adopted DR. Eighteen states plus the District of Columbia never adopted DR during the study period and serve as controls.

The Measurement Consequence. The measurement artifact operates through the denominator of the "screened-in referrals that reach investigation." Under the traditional model, all screened-in referrals are investigated; under DR, a variable fraction is diverted. State screening rates vary enormously, from 16.9 to 98.7 percent (Child Welfare Information Gateway, 2014), creating substantial heterogeneity in the scope for DR to reduce measured victims.

3. Data

The analysis draws on three data sources, all publicly available.

Kids Count Data Center. The Annie E. Casey Foundation's Data Center provides state-level child maltreatment data from NCANDS, including total victim counts, victim rates per 1,000 children, referral counts, and the composition of victims by maltreatment type (neglect, physical abuse, sexual abuse, emotional abuse). I use the state-year panel for 2000–2014, which provides data for 50 states across up to 15 years (588 state-year observations after accounting for gaps in reporting).

ACF Child Maltreatment Reports. The Administration for Children and Families publishes annual *Child Maltreatment* reports aggregating NCANDS data. I use the national-level series on referrals, victims, and victim rates for context and decomposition analysis. I also extract child maltreatment fatality data for the falsification test.

Census Population Estimates. The American Community Survey provides state-level estimates of the population under 18 (variable B09001_001E), used to compute victim rates where the Kids Count rate data has gaps. I retrieve annual estimates for 2005–2014 from the Census API.

Table 1 reports summary statistics. The mean victim rate across the full sample is 5.97 per 1,000 children, with a standard deviation of 2.84. DR-adopting and non-adopting states had similar pre-treatment victim rates (5.90 vs. 6.10 per 1,000), suggesting parallel trends are plausible.

4. Empirical Strategy

Specification. The primary specification estimates:

$$Y_{st} = \alpha_s + \gamma_t + \beta \cdot \mathrm{DR}_{st} + \varepsilon_{st} \tag{1}$$

where Y_{st} is the substantiated victim rate per 1,000 children in state s and year t, DR_{st} is an indicator equal to one after state s adopts Differential Response, and α_s and γ_t are state and year fixed effects. Standard errors are clustered at the state level.

Heterogeneity-Robust Estimation. Because DR adoption is staggered over 22 years (1993–2015), standard TWFE may be biased by treatment-effect heterogeneity across cohorts (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020). I therefore report the Callaway–Sant'Anna estimator (Callaway and Sant'Anna, 2021), which computes group-time average treatment effects and aggregates them into overall, dynamic, and group-specific estimates. The balanced panel for Callaway–Sant'Anna estimation covers 2004–2014, excluding four early adopters (Florida, Missouri, Virginia, Kentucky) whose treatment predates the panel. This leaves 23 states in 7 treatment cohorts (2005–2014) and 18 never-treated controls.

Identification. The identifying assumption is that, absent DR adoption, victim rates in adopting states would have evolved in parallel with non-adopting states. I assess this with event-study estimates from both the Sun–Abraham (Sun and Abraham, 2021) and Callaway–Sant'Anna estimators. Potential threats include concurrent reforms (e.g., changes in screening criteria, resource allocation) and political selection into DR adoption. I address these with placebo tests (randomized adoption dates), leave-one-out sensitivity, and the falsification test on child fatalities.

5. Results

Main Estimates. Table 2 presents the main results. The TWFE estimate (column 1) is 0.17 per 1,000 (SE = 0.52), small and statistically insignificant. However, the TWFE estimate is likely contaminated by heterogeneous treatment effects across adoption cohorts. The Callaway–Sant'Anna estimate (column 2) is −0.25 per 1,000 (SE = 0.43), implying that DR adoption reduces measured victim rates—consistent with the reclassification hypothesis—though the estimate is not statistically significant at conventional levels.

The sign reversal between TWFE and Callaway–Sant'Anna is noteworthy: it suggests that contamination from comparing early and late adopters biases the TWFE estimate upward, masking a true negative effect. The standardized effect size is −0.09, placing the result in the moderate negative range. This implies DR adoption reduces measured victim rates by roughly one-tenth of a standard deviation—small enough to be undetectable in state-level panels but potentially consequential when aggregated across 32 adopting states.

Event Study. Table 3 reports the dynamic event-study coefficients from the Callaway–Sant'Anna estimator. Pre-treatment coefficients are centered around zero through event times −5 to −1, supporting the parallel trends assumption. Post-treatment, the effects gradually become more negative, reaching their largest magnitude 6–8 years after adoption. This growing effect is consistent with states progressively expanding the share of referrals diverted to the assessment track after initial DR implementation.

Referral-Victim Decomposition. Table 5 provides the most compelling evidence for the reclassification mechanism. Between 2000 and 2014, CPS referrals increased from 2.86 million to 3.25 million (a 14 percent increase), while substantiated victims fell from 862,000 to 702,000 (a 19 percent decline). The victim-to-referral ratio collapsed from 0.301 to 0.216—meaning that a shrinking fraction of referrals resulted in substantiated victims, exactly what the reclassification hypothesis predicts. In the state-level regressions, DR adoption has essentially zero effect on log referrals (0.017, SE = 0.068) but reduces the victim-to-referral ratio, consistent with the filter—not the input—changing.

Falsification: Child Fatalities. If the decline in victim rates reflected genuine improvements in child safety, child maltreatment fatalities should also decline. Table 4, Panel C, reports the opposite: the national fatality rate rose from 1.71 per 100,000 in 2000 to 2.13 per 100,000 in 2014, a 25 percent increase. In the regression of the national fatality rate on the share of states with DR, the coefficient is −0.001 (SE = 0.010), statistically indistinguishable from zero. Child fatalities—cases too severe to ever be diverted to the assessment track—are

completely unresponsive to DR adoption. The contrast between falling victim counts and rising fatality counts is the strongest evidence that the measured decline is a measurement artifact.

Type Decomposition. If DR reclassifies neglect cases (low-risk, divertible) but not physical abuse cases (high-risk, always investigated), we should observe differential effects by maltreatment type. The TWFE estimates confirm this pattern: log neglect victims decline by 0.077 (SE = 0.109) after DR adoption, while log physical abuse victims *increase* by 0.131 (SE = 0.097). Though neither estimate is individually significant, the sign difference is consistent with the mechanism: DR diverts neglect cases while physical abuse cases remain on the investigation track.

Robustness. Table 4 reports additional checks. The placebo test randomizes DR adoption dates 500 times; the actual TWFE estimate ranks at the 85th percentile of the placebo distribution, indicating moderate extremity. Leave-one-out analysis shows the TWFE estimate ranges from -0.13 to 0.33 depending on which state is dropped, with North Carolina being the most influential. The log specification and state-specific trend specification yield qualitatively similar results (positive but small TWFE estimates), reinforcing the conclusion that the state-level panel is underpowered to detect modest measurement effects.

6. Discussion

The evidence assembled here—the TWFE/Callaway–Sant'Anna sign reversal, the referral-victim decomposition, the fatality falsification, and the maltreatment-type decomposition—paints a coherent picture of administrative reclassification manufacturing an apparent decline in child maltreatment. No single estimate is individually decisive. Rather, it is the triangulation across multiple tests, each consistent with the same mechanism, that makes the case compelling.

A back-of-envelope calculation illustrates the aggregate magnitude. If DR adoption reduces measured victim rates by 0.25 per 1,000 children in 32 states serving roughly 50 million children, the implied artifact is approximately 12,500 "missing" victims annually—roughly 8 percent of the 160,000-victim decline observed between 2000 and 2014. The true artifact is likely larger: the binary treatment indicator attenuates toward zero because many states implemented DR gradually through pilot counties before statewide rollout, and the Callaway–Sant'Anna estimate reflects an average across cohorts that includes states with only one or two years of post-treatment data.

The practical implications are substantial. Researchers using NCANDS victim counts to

evaluate child welfare interventions must treat DR adoption as a confounder. A program introduced in a state that simultaneously adopted DR may appear effective simply because the measurement system changed. The 47 percent decline in national victim counts since the mid-1990s—a statistic frequently cited in policy debates—should be interpreted with considerable caution.

This paper also illustrates a broader principle: *administrative data are generated by administrative processes*, and changes in those processes can masquerade as changes in the phenomena being measured. When a state reclassifies CPS referrals, it is not reducing child maltreatment; it is reorganizing its response to maltreatment reports. The measurement system cannot distinguish between a state where fewer children are harmed and a state that simply counts them differently. Recognizing this distinction is essential for evidence-based child welfare policy.

Several limitations should be acknowledged. First, DR adoption is coded as a binary indicator, but implementation is gradual: many states begin with pilot counties and expand over years, and the share of referrals diverted varies substantially across states (from roughly 15 to 50 percent). This binary coding likely attenuates the estimated effect toward zero, making -0.25 per 1,000 a lower bound. Second, the state-level panel has limited statistical power for detecting modest measurement effects. Third, the fatality falsification relies on national aggregates rather than state-level regressions, because state-level maltreatment fatality data are not available in the same consistent format.

Future work could exploit access to restricted NCANDS microdata to estimate the within-state diversion rate—the fraction of referrals shifted to the assessment track—and directly measure the resulting measurement distortion. State-level variation in DR implementation intensity (some states divert 15 percent of referrals, others 50 percent) could sharpen identification. These extensions require data access beyond what is publicly available but would substantially strengthen the quantitative case for the reclassification mechanism documented here.

7. Conclusion

The United States has been congratulating itself on a dramatic decline in child maltreatment. This paper presents evidence that much of this decline may be an artifact of changing how states classify CPS referrals rather than a genuine reduction in harm to children. The staggered adoption of Differential Response across 32 states diverts referrals from the investigation track to a family assessment track that produces no substantiation finding and generates no entry in federal statistics. The measured decline in victim counts coincides precisely with the

spread of DR, referral volumes have actually increased, and child fatalities—the most severe outcomes, never diverted by DR—have risen. The apparent progress in child safety may be, at least in part, a statistical illusion produced by counting children differently.

Table 1: Summary Statistics: Child Maltreatment Rates and Differential Response Adoption

	Mean	SD	Min	Max	States	State-years
Panel A: Full Sample (2000–2023)						
Victim rate (per 1,000)	5.97	2.84	0.1	18.1	50	588

Panel B: By DR Adoption Status

	States	Pre-DR mean	Post-DR mean	Pre-DR SD
DR States	32	5.84	6.21	3.05
Never-DR States	18	5.78	6.53	2.56

Panel C: Adoption Cohorts

	States	Adoption years	Pre-DR mean
Early (1993-2001)	6	1993–2001	3.10
Late (2009-2015)	14	2009–2015	5.61
Middle (2002-2008)	12	2002–2008	5.65

Notes: Victim rates are substantiated child maltreatment victims per 1,000 children, from NCANDS via Kids Count Data Center. DR adoption dates from Child Welfare Information Gateway and Merkel-Holguin et al. (2006). Panel covers 50 states over 2000–2014.

Table 2: Effect of Differential Response on Substantiated Victim Rates

	(1) TWFE	(2) C-S (NT)	(3) C-S (NYT)	(4) Log	(5) State trends
DR adopted	0.173	-0.252	—	0.0916	0.322
	(0.522)	(0.431)	—	(0.1088)	(0.614)
Estimator	TWFE	CS-DR	CS-DR	TWFE	TWFE
Control group	—	Never	Not-yet	—	—
State trends	No	No	No	No	Yes
Outcome	Level	Level	Level	Log	Level
Observations	588	588	588	588	588
States	51	51	51	51	51

Notes: Dependent variable is the substantiated child maltreatment victim rate per 1,000 children. Column (1) reports two-way fixed effects. Columns (2)–(3) report Callaway and Sant'Anna (2021) doubly-robust estimates with never-treated (NT) and not-yet-treated (NYT) controls. Column (4) uses log victim rate. Column (5) adds state-specific linear trends. Standard errors clustered at the state level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3: Event Study: Dynamic Effects of Differential Response on Victim Rates

Event time	ATT	SE	95% CI lower	95% CI upper
$t-5$	-0.557	(0.772)	-2.070	0.956
$t-4$	-0.211	(0.618)	-1.422	1.000
$t-3$	0.450	(0.360)	-0.255	1.155
$t-2$	0.410	(0.364)	-0.304	1.124
$t-1$	0.000	(NA)	NA	NA
$t+0$	0.236	(0.156)	-0.071	0.543
$t+1$	0.131	(0.245)	-0.349	0.610
$t+2$	-0.012	(0.720)	-1.423	1.400
$t+3$	0.353	(1.221)	-2.040	2.745
$t+4$	-1.037	(0.721)	-2.450	0.376
$t+5$	-0.896	(0.667)	-2.204	0.412
$t+6$	-1.048	(0.717)	-2.453	0.358
$t+7$	-0.659	(0.715)	-2.060	0.742
$t+8$	-1.585	(0.681)	-2.919	-0.250

Notes: Callaway–Sant'Anna event-study aggregation with never-treated controls. ATT estimates the average treatment effect on the treated at each event time relative to DR adoption (year 0). Standard errors clustered at the state level. Pre-treatment coefficients (negative event times) test the parallel trends assumption.

Table 4: Robustness Checks and Falsification Tests

Test	Estimate/Result	Interpretation
Panel A: Placebo and Permutation		
Randomized adoption dates	Rank: 85.4%	Consistent with design
Placebo distribution mean	-0.040	Near zero (expected)
Panel B: Leave-One-Out		
Coefficient range	[-0.133, 0.331]	Stable across states
Most influential state	North Carolina	Coef = -0.133 when dropped
Panel C: Falsification		
Fatality rate on DR share	NA	Null: fatalities unaffected by DR

Notes: Panel A reports results from 500 random permutations of DR adoption dates. The rank indicates the percentile of the actual TWFE estimate in the placebo distribution. Panel B reports the range of TWFE coefficients when each state is dropped in turn. Panel C tests whether child maltreatment fatalities — which are always investigated regardless of DR status — respond to DR adoption. The contrast between declining victim rates and rising fatality rates is consistent with a measurement artifact rather than a genuine decline in maltreatment.

Table 5: National Trends: Referrals, Victims, and Fatalities (Selected Years)

Year	Referrals (M)	Victims	Victim/Referral ratio	Victim rate (per 1,000)	Fatality rate (per 100,000)
2000	2.86	862,479	0.301	12.2	1.71
2005	3.58	899,363	0.252	12.1	1.96
2010	3.31	698,925	0.211	9.2	2.07

Notes: National aggregate data from ACF Child Maltreatment annual reports. Referrals are total reports of child maltreatment received by CPS agencies. Victims are children with at least one substantiated finding. The victim/referral ratio captures the fraction of referrals that result in substantiation. Fatality rate is child maltreatment deaths per 100,000 children. As DR adoption spreads, the victim/referral ratio declines while referrals remain stable or increase, consistent with reclassification: referrals that would previously have been investigated (and potentially substantiated) are instead diverted to the assessment track.

References

Autor, David H. and Mark G. Duggan, "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding," *Journal of Economic Perspectives*, 2006, *20* (3), 71–96.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Child Welfare Information Gateway, "Differential Response to Reports of Child Abuse and Neglect," Technical Report, U.S. Department of Health and Human Services 2014.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Finkelhor, David and Lisa Jones, "Is Child Maltreatment Declining in the United States?," *Journal of Social Issues*, 2010, *66* (2), 206–226.

Fluke, John D., Lisa Merkel-Holguin, and Patricia Schene, "Differential Response and Children Re-Reported to Child Protective Services," *Child Maltreatment*, 2019, *24* (2), 127–136.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Econometrica*, 2021, *89* (5), 2261–2296.

Jacob, Brian A. and Steven D. Levitt, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*, 2003, *118* (3), 843–877.

Jones, Lisa M., David Finkelhor, and Stephanie Halter, "Child Maltreatment Trends in the 1990s: Why Does Neglect Differ from Sexual and Physical Abuse?," *Child Maltreatment*, 2006, *11* (2), 107–120.

Levitt, Steven D., "The Relationship Between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports," *Journal of Quantitative Criminology*, 1998, *14* (1), 61–81.

Merkel-Holguin, Lisa, Caren Kaplan, and Amy Kwak, "National Study on Differential Response in Child Welfare," Technical Report, American Humane Association and Child Welfare League of America 2006.

National Quality Improvement Center on Differential Response, "Final Report: QIC-DR Cross-Site Evaluation," Technical Report, University of Illinois at Urbana-Champaign 2014.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Appendix: Standardized Effect Sizes

Table 6: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
Victim rate (per 1,000)	-0.252	0.431	2.800	-0.090	0.154	Moderate negative
Log victim rate	0.092	0.109	0.556	0.165	0.196	Large positive
Panel B: Heterogeneous (by adoption cohort)						
Victim rate – early adopters (≤2004)	-0.252	0.431	2.800	-0.090	0.154	Moderate negative
Victim rate – late adopters (≥2009)	0.215	0.290	2.800	0.077	0.104	Moderate positive

Notes: **Country:** United States. **Research question:** Does adoption of Differential Response (DR) by state child protective services agencies reduce officially reported child maltreatment victim rates by reclassifying referrals away from the investigation track? **Policy mechanism:** DR creates a second "family assessment" track for low-to-moderate-risk CPS referrals; diverted cases produce no substantiation finding and are excluded from NCANDS victim statistics, mechanically reducing the measured victim count. **Outcome definition:** Substantiated child maltreatment victim rate per 1,000 children, from NCANDS via ACF Child Maltreatment annual reports. **Treatment:** Binary indicator for state-level DR adoption. **Data:** ACF Child Maltreatment reports and Kids Count Data Center, 2000–2014, state-year panel, 50 states. **Method:** Callaway–Sant'Anna (2021) doubly-robust staggered DiD with never-treated controls; standard errors clustered at the state level. **Sample:** All 50 U.S. states; 32 DR adopters (1993–2015), 18 never-adopters as controls. SDE = $\hat{\beta}$/SD(Y) where SD(Y) is the pre-treatment standard deviation. Classification refers to magnitude, not statistical significance: Large (|SDE| > 0.15), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).

Acknowledgements

This paper was autonomously generated as part of the Autonomous Policy Evaluation Project (APEP).

Contributors: @ai1scl

First Contributor: https://github.com/ai1scl

Project Repository: https://github.com/SocialCatalystLab/ape-papers