RESEARCH ARTICLE

Socioeconomic Disadvantage, Not Gun Policy or Mental

Illness, Is Associated With State-Level Mass Shooting Rates: A Negative-Binomial Panel Analysis, 2018–2024

Safaa Dabagh[1,2]

[1] Department of Mathematics, Santa Monica College, Santa Monica, CA, USA

[2] Department of Mathematics, Loyola Marymount University, Los Angeles, CA, USA

Corresponding author: Safaa Dabagh (dabagh_safaa@smc.edu)

Abstract

Background: Mass shootings in the United States rose sharply after 2019, yet debate over their state-level correlates—gun policy, mental illness, or socioeconomic conditions—often rests on single-method, cross-sectional analyses vulnerable to model misspecification. We examined which state characteristics are associated with mass-shooting incidence across a seven-year panel, and how incidence changed over time.

Methods: We assembled a state-year panel of mass-shooting counts from the Gun Violence Archive (GVA; 2018–2024; 50 states; 350 state-years; 3,783 incidents) with state-level socioeconomic, gun-policy, and mental-health predictors. Because counts were overdispersed, we fit negative-binomial regression with a population offset, calendar-year fixed effects, and standard errors clustered by state, with a CR2 small-sample correction given 50 clusters. Predictors were standardized and reported as incidence-rate ratios (IRR) per standard deviation. LASSO provided independent variable selection; tree-ensemble importance was examined as supplementary triangulation.

Results: Incidence rose steeply and persistently relative to 2018, roughly doubling at its 2021 peak (IRR 2.02, 95% CI 1.68–2.43) and remaining elevated through 2024 (IRR 1.55). Poverty was associated with higher incidence (IRR 1.45, 1.18–1.79, $p = .006$ with CR2 SEs) and the more stable socioeconomic predictor across specifications; income inequality carried an equivalent but less robust association (IRR 1.44 alone; the two correlated $r = 0.50$). The poverty association was stable before and after 2020 (interaction $p = .51$). Gun-law strength (IRR 1.11) and household gun ownership (IRR 1.07) were not associated. Serious mental illness—the construct most contested in this literature—showed no association (IRR 1.04, $p = .56$); any mental illness was inversely associated (IRR 0.78), an ecological pattern, not protection. LASSO retained only the socioeconomic-disadvantage measures.

Conclusions: The post-2020 rise in mass shootings was broad and sustained, and its geography was associated with socioeconomic disadvantage—chiefly poverty—rather than gun-law permissiveness or mental-illness prevalence; under the GVA 4+-shot definition, this association did not extend to gun policy or mental illness. The ecological, observational design precludes individual-level and causal inference.

Keywords: mass shootings, socioeconomic disadvantage, poverty, income inequality, gun policy, mental health, negative binomial, state panel, gun violence

Background

Mass shootings are a persistent feature of gun violence in the United States, and their frequency rose markedly after 2019. While public attention often centers on individual-level explanations—particularly mental illness and firearm access—comparatively little population-level research examines which state characteristics are associated with where mass shootings concentrate, or how that burden has changed.

The limited population-level literature points to socioeconomic conditions. Kwon and Cabrera reported that U.S. counties and states with greater income inequality experienced higher mass-shooting rates [2,3], an association often interpreted through Merton's strain theory, a relative-deprivation account in which disparities between observed affluence and blocked opportunity generate conditions conducive to violence [4]. Income inequality and poverty are, however, closely entangled, and several state-level analyses emphasize gun policy or mental health while including only limited socioeconomic covariates; fewer compare socioeconomic, policy, and mental-health domains within a single, appropriately specified count model.

Two methodological issues recur. First, mass-shooting counts are non-negative, skewed, and overdispersed, yet are sometimes modeled with ordinary least squares on a rate, which can distort inference. Second, single-method, single-period, cross-sectional designs are vulnerable to model-specific artifacts. Machine-learning approaches have been applied to firearm violence [5]; we build on this by comparing socioeconomic, policy, and mental-health domains across a seven-year state-year panel using count models appropriate to the outcome, with regularized and ensemble methods as independent checks.

We address three questions: (1) How did state-level mass-shooting incidence change between 2018 and 2024? (2) Which state characteristics are associated with incidence? (3) Are any associations robust across complementary modeling approaches and stable across the post-2020 surge?

Methods

Study design and outcome

This is an observational, ecological analysis of a state-year panel. Mass-shooting counts by state and year (2018–2024) were obtained from the Gun Violence Archive (GVA), which defines a mass shooting as an incident in which four or more people are shot (injured or killed), excluding the shooter [1]. This 4+-shot definition captures the broad burden of community gun violence and differs from fatal-only (4+-killed) definitions, which identify a rarer, qualitatively distinct event class; recent work shows the two definitions select substantially different incidents [6], and the fatal-only subset is too small for stable state-panel modeling. Counts were retrieved on 10 February 2026; because the GVA is continuously updated, counts may differ slightly by retrieval date, and no incidents were excluded.

The unit of analysis was the state-year (50 states $\times$ 7 years = 350 observations). Because all predictors are time-invariant, predictor associations are identified from 50 between-state contrasts (effectively a cross-section with a repeated outcome), while year fixed effects absorb the temporal trend; we therefore treat predictor inference as resting on $n = 50$. The District of Columbia was excluded from all models: it is not a state, does not receive a Giffords gun-law grade, and is an extreme population-density outlier.

Predictor variables

Predictors spanned three focal domains. *Socioeconomic:* Gini coefficient, poverty rate, median household income (Census ACS 2023 [7]). *Mental health:* prevalence of any mental illness (AMI) and serious mental illness (SMI) (SAMHSA NSDUH 2022–2023 [8]; SMI was missing for two states, analysed listwise, $n = 48$ states / 336 state-years); mental-health provider density (KFF 2023 [9]). *Gun policy:* household gun ownership (RAND 2020 [10]) and Giffords gun-law strength score [11]. Additional covariates included violent-crime rate (FBI UCR 2023 [12]), per-capita alcohol consumption [13], population density, percent urban, and percent aged 18–29 (Census [7]). The complete codebook is provided in Additional File 1.

Statistical analysis

Mass-shooting counts were overdispersed: the model-based Pearson dispersion statistic (after the population offset) was 4.2, motivating negative-binomial regression with a log-population offset and calendar-year fixed effects. Population offsets used 2023 Census estimates applied to all years; because population is therefore time-invariant, early-year rates are slightly misstated, a negligible effect at this aggregation. Standard errors were clustered by state; given exactly 50 clusters, we applied the CR2 bias-reduced correction with Satterthwaite degrees of freedom [20,21], which is conservative relative to conventional cluster-robust errors. Continuous predictors were standardized, so coefficients are IRRs per one standard-deviation increase. Because income inequality and poverty are entangled, we examined them individually and jointly and report their correlation. To test stability across the structural break, we estimated the poverty association separately before (2018–2019) and after (2020–2024) 2020 and tested a poverty $\times$ period interaction. LASSO regression (cross-validated penalty) provided independent variable selection [14]; random-forest [15] and gradient-boosting [16] importance were examined as supplementary triangulation but, given $n = 50$, are reported as qualitative corroboration only, not inference. Analyses were conducted by the author in R (`MASS`, `sandwich`, `clubSandwich`, `glmnet`); results are reproducible from the de-identified data and code.

Use of AI-assisted technologies

The statistical analyses were specified and conducted by the author in R; generative AI (Claude, Anthropic) assisted with prose drafting and editing only and did not perform the analysis. All references were verified by the author against primary sources. The author takes full responsibility for the content.

Results

The temporal surge

Across 50 states and seven years, 3,783 mass-shooting incidents were recorded (median 6 per state-year; range 0–86). National counts rose from 329 in 2018 to a peak of 673 in 2021, remaining elevated through 2023 (649) before partially receding to 491 in 2024 (Table 1; Figure 1). In the negative-binomial model, year effects were large and highly significant: relative to 2018, incidence surged from 2020 (IRR 1.89) to a 2021 peak (2.02, 95% CI 1.68–2.43), remaining near double baseline through 2023 (1.98) before declining in 2024 (1.55); all $p \leq .015$. This post-2020 increase was the single largest signal in the data; because predictors are time-invariant, the panel documents but cannot mechanistically explain it.

Table 1: National mass-shooting incidents by year (GVA definition, 50 states)

Year	Incidents	IRR vs. 2018 (95% CI)
2018	329	— (reference)
2019	406	1.25 (1.04–1.51)
2020	602	1.89 (1.60–2.23)
2021	673	2.02 (1.68–2.43)
2022	633	1.91 (1.59–2.30)
2023	649	1.98 (1.64–2.38)
2024	491	1.55 (1.30–1.86)
Total	**3,783**	

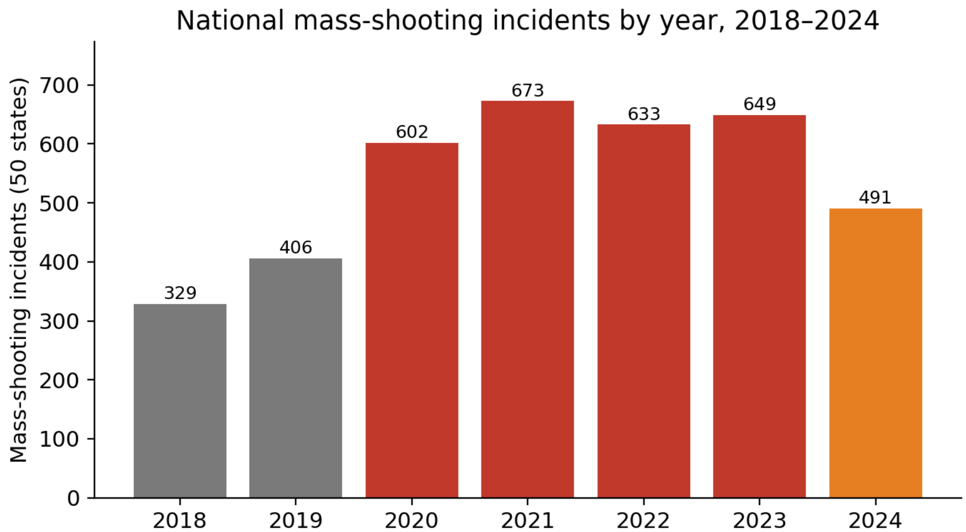

Figure 1: National mass-shooting incidents by year, 2018–2024 (50 states). Source: Gun Violence Archive.

Socioeconomic disadvantage

Table 2 presents the predictor associations across three models; Figure 2 shows each predictor's association taken from the model in which it appears. Poverty was associated with higher incidence (IRR 1.45 per SD, 95% CI 1.18–1.79, p = .006 under CR2 SEs, in the primary model adjusted for gun-law strength and serious mental illness; the poverty-only estimate is 1.40) and was the more stable of the two socioeconomic measures across specifications. Income inequality carried an equivalent single-predictor association (IRR 1.44, 1.17–1.78, p = .001); the two correlated $r = 0.50$ (variance inflation ≈1.3), and entered jointly each attenuated to marginal significance (Gini 1.22, p = .048; poverty 1.26, p = .018)—a pattern more consistent with limited power at $n = 50$ than with true inseparability. LASSO retained poverty (+0.16) far more strongly than Gini (+0.03), so we interpret the association at the level of the socioeconomic-disadvantage construct, with poverty as its most robust single proxy. The poverty association was stable across the structural break: estimated on the same adjusted specification, the poverty IRR was 1.34 (1.07–1.69) before 2020 and 1.49 (1.20–1.85) after—the pooled estimate (1.45) falling between them—with a non-significant poverty × period interaction (p = .51). It also persisted after adjustment for violent-crime rate and gun ownership (IRR 1.28, p = .022; Model 3); violent crime was itself positively associated (IRR 1.27), consistent with mass shootings being embedded in broader community violence.

Table 2: Negative-binomial associations with mass-shooting incidence (IRR per 1 SD; population offset, year fixed effects, CR2 cluster-robust SE by state)

Predictor	Model 1 (primary)	Model 2 (SES alt.)	Model 3 (adjusted)
Poverty rate	1.45 (1.18–1.79)**	—	1.28 (1.04–1.58)*
Income inequality (Gini)	—	1.44 (1.17–1.78)**	—
Gun-law strength	1.11 (0.88–1.41)	1.10 (0.87–1.38)	1.06 (0.84–1.34)
Serious mental illness	1.04 (0.91–1.20)	1.03 (0.90–1.18)	1.05 (0.92–1.20)
Household gun ownership	—	—	1.07 (0.84–1.36)
Violent-crime rate	—	—	1.27 (1.05–1.54)*
Year fixed effects	yes	yes	yes

** $p < .01$, * $p < .05$. Models 1–3 include serious mental illness (listwise): $n = 336$ (48 states).
Single-predictor estimates quoted in text (poverty-only, Gini-alone) use $n = 350$ (50 states). Population offset, CR2 SEs.

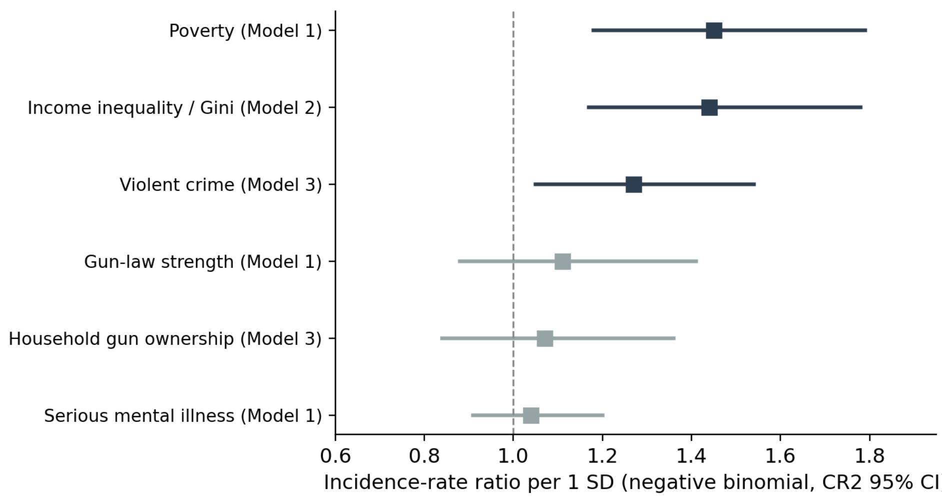

Figure 2: Associations for each predictor of mass-shooting incidence, taken from the model in which it appears (Table 2); negative binomial, IRR per 1 SD, CR2 95% CI. Serious mental illness is shown as the primary mental-illness measure. Intervals right of 1 indicate higher incidence: socioeconomic disadvantage and violent crime are associated with incidence; gun-law strength, household gun ownership, and serious mental illness are not.

Gun policy and mental illness

Neither gun-law strength (IRR 1.11, 0.88–1.41, $p = .39$) nor household gun ownership (IRR 1.07, $p = .60$) was associated with incidence. Serious mental illness—the construct most contested in the mass-shooting literature—showed no association (IRR 1.04, 0.91–1.20, $p = .56$), unchanged after adjustment for household income. Any mental illness was inversely associated (IRR 0.78, $p = .005$); we read this as ecological confounding (states reporting higher prevalence tend to be wealthier, with greater diagnostic infrastructure and less stigma) rather than protection. Testing SMI directly addresses the specific version of the mental-illness claim the literature contests, and it too is null.

Independent variable selection and triangulation

LASSO retained only the two socioeconomic-disadvantage measures (poverty +0.16; Gini +0.03), shrinking all gun-policy and mental-health coefficients to zero. Supplementary tree-ensemble rankings concurred: random-forest importance placed Gini and poverty first and second and gun-law strength last of fourteen predictors, with violent crime also high; gradient boosting gave a similar ordering (Figure 3). Because these ensembles are estimated on 50 observations, their rankings are qualitative corroboration of the negative-binomial findings, not independent evidence.

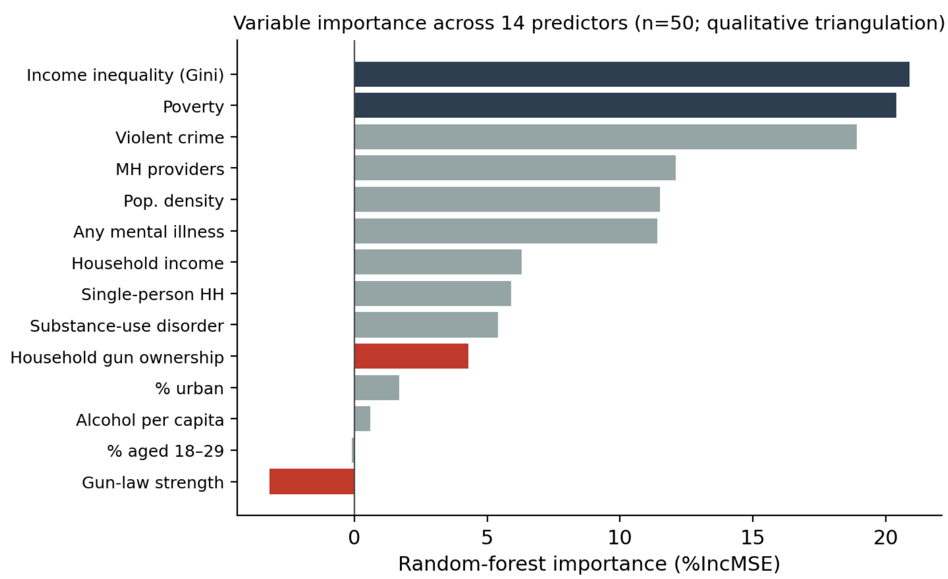

Figure 3: Random-forest variable importance (supplementary triangulation; $n = 50$, qualitative only). Socioeconomic-disadvantage measures rank highest, gun-law strength lowest.

Discussion

Across a seven-year state panel, two findings are robust to model choice. First, mass-shooting incidence rose sharply after 2019 and remained near double its 2018 level through 2023—a broad, sustained increase that is the dominant feature of the data. Second, the geographic concentration of that burden was associated with socioeconomic disadvantage—chiefly poverty, the most stable predictor across negative-binomial, LASSO, and ensemble methods and across the pre/post-2020 break—while gun-law strength, household gun ownership, and mental-illness prevalence were not.

Disadvantage as a construct, and a tension with strain theory

Our results support an association between socioeconomic disadvantage and mass-shooting rates but, unlike some prior work, do not privilege income inequality over poverty: the two are correlated ($r = 0.50$), and poverty was the more stable. This creates a theoretical tension worth naming. Merton's strain account is a *relative*-deprivation theory that fits inequality, yet our dominant predictor is poverty, an *absolute*-deprivation measure; the data are thus more consistent with a general socioeconomic-disadvantage mechanism than with strain theory specifically. That violent crime co-ranked with the socioeconomic measures further suggests mass shootings are one expression of broader community violence rooted in shared structural conditions.

Gun policy and mental illness

The absence of association with gun-law strength and ownership does not establish that firearm policy is ineffective: cross-sectional, state-level data cannot capture enforcement, within-state variation, lethality, or effects on other gun violence. The prior literature is itself mixed and, tellingly, definition-dependent: studies using fatal (4+-killed) definitions have found more permissive state gun laws associated with higher mass-shooting rates [17], whereas analyses using the broader 4+-shot definition—as here—have found no association with state gun-law grades [18]. Our null aligns with the latter, and the divergence from fatal-definition studies reinforces that conclusions about gun policy hinge on which mass-shooting definition is used—indeed, this split appears even within a single study, as Duchesne et al. found a gun-law association with mass-shooting *fatalities* while the *incidence* association was null [18]. What our data indicate is that gun-law permissiveness does not explain the *geographic* pattern of 4+-shot incidence once socioeconomic context is considered. Likewise, neither serious nor any mental illness predicts where mass shootings concentrate. Much of the mental-illness literature concerns rarer public or fatal mass shootings [19]; our null under the broader GVA definition extends rather than contradicts it, while underscoring that conclusions are definition-dependent—which is why we frame the headline claim within the 4+-shot definition rather than as a general statement.

Strengths and limitations

Strengths include a seven-year panel, a count model appropriate to the outcome with CR2 small-sample-corrected inference, convergent variable selection across methods, and a demonstrated stability of the poverty association across the 2020 surge. Limitations are substantial. The design is ecological and observational: state-level associations cannot be attributed to individuals, and no causal inference is warranted. Predictors are time-invariant, so the analysis documents the surge but cannot explain it through changing covariates, and predictor inference rests on 50 between-state contrasts. The data are state incident counts with no perpetrator-level information and cannot speak to perpetrator characteristics such as age. The GVA 4+-shot definition captures predominantly community gun violence and differs from fatal-only definitions. Several predictors reflect periods adjacent to, rather than identical with, the outcome window.

Conclusions

State-level mass-shooting incidence rose sharply and persistently after 2019, and its geographic burden was associated with socioeconomic disadvantage, chiefly poverty. Gun-law strength, household gun ownership, and mental-illness prevalence were not associated with this geography under the GVA definition. These observational findings are consistent with—though they cannot by themselves demonstrate—a role for upstream socioeconomic conditions in shaping where mass-shooting risk concentrates, and they caution against framing mass shootings primarily as a problem of gun-law permissiveness or individual mental illness. Understanding the post-2020 surge, and the role of younger perpetrators that incident-level studies describe, will require individual- and incident-level designs beyond the ecological scope of this analysis.

Declarations

Ethics approval and consent to participate

Not applicable. The study used only publicly available, de-identified aggregate state-level data; no human subjects were involved.

Consent for publication

Not applicable.

Availability of data and materials

The compiled, de-identified analysis dataset and R code are available from the author on request and will be deposited in a public repository (e.g., OSF). Mass-shooting data derive from the Gun Violence Archive (https://www.gunviolencearchive.org/); predictor sources are documented in Methods.

Competing interests

The author declares no competing interests.

Funding

None.

Authors' contributions

SD is the sole author and conducted all analyses in R.

References

1. Gun Violence Archive. Mass shooting data 2013–present. 2025. https://www.gunviolencearchive.org/. Accessed 10 Feb 2026.

2. Kwon R, Cabrera JF. Income inequality and mass shootings in the United States. *BMC Public Health*. 2019;19:1147.

3. Kwon R, Cabrera JF. Income inequality, household income, and mass shooting in the United States. *Front Public Health*. 2018;6:294.

4. Merton RK. Social structure and anomie. *Am Sociol Rev*. 1938;3(5):672–82.

5. Goin DE, Rudolph KE, Ahern J. Predictors of firearm violence in urban communities: a machine-learning approach. *Health Place*. 2018;51:61–67.

6. Knapp KG. Conceptual confusion: an examination of the definitional debates of mass shootings using NIBRS. *Crime Delinq*. 2025. doi:10.1177/00111287251384672.

7. U.S. Census Bureau. American Community Survey 1-year estimates, 2023. 2024. https://data.census.gov/. Accessed 12 Feb 2026.

8. Substance Abuse and Mental Health Services Administration. 2022–2023 NSDUH: model-based estimated prevalence for states. 2025. https://www.samhsa.gov/data/. Accessed 12 Feb 2026.

9. Kaiser Family Foundation. Mental health care health professional shortage areas. KFF State Health Facts. 2023. https://www.kff.org/. Accessed 12 Feb 2026.

10. RAND Corporation. State-level estimates of household firearm ownership. 2020. https://www.rand.org/pubs/tools/TLA243-2-v2.html. Accessed 12 Feb 2026.

11. Giffords Law Center. Annual gun law scorecard. 2023. https://giffords.org/lawcenter/resources/scorecard/. Accessed 12 Feb 2026.

12. Federal Bureau of Investigation. Crime in the United States, 2023. Uniform Crime Reporting Program. 2024. https://ucr.fbi.gov/. Accessed 12 Feb 2026.

13. National Institute on Alcohol Abuse and Alcoholism. Surveillance report #120: apparent per capita alcohol consumption. 2023. https://www.niaaa.nih.gov/. Accessed 12 Feb 2026.

14. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc B*. 1996;58(1):267–88.

15. Breiman L. Random forests. *Mach Learn*. 2001;45:5–32.

16. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29(5):1189–232.

17. Reeping PM, Cerdá M, Kalesan B, Wiebe DJ, Galea S, Branas CC. State gun laws, gun ownership, and mass shootings in the US: cross sectional time series. *BMJ*. 2019;364:l542.

18. Duchesne J, Taghavi S, Toraih E, Simpson JT, Tatum D. State gun law grades and impact on mass shooting event incidence: an 8-year analysis. *J Am Coll Surg*. 2022;234(4):608–16.

19. Peterson J, Densley J. *The Violence Project: How to Stop a Mass Shooting Epidemic*. New York: Abrams Press; 2021.

20. Bell RA, McCaffrey DF. Bias reduction in standard errors for linear regression with multi-stage samples. *Surv Methodol*. 2002;28(2):169–81.

21. Pustejovsky JE, Tipton E. Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *J Bus Econ Stat.* 2018;36(4):672–83.

Additional file

Additional file 1: Variable codebook with sources, definitions, and state-level data. (CSV)