‘Seven Times More Likely’: Risk Evidence and the Effects of Nonfatal Strangulation Laws

Zubin Jelveh, Xinyi Situ, & Abdul Nasser Rad

October 28, 2026

Abstract

Between 2000 and 2019, 48 U.S. states made nonfatal strangulation a felony. Originally a fix for undercharging, the laws were increasingly justified as homicide prevention, based on evidence that prior strangulation raises the odds of later intimate partner homicide sevenfold. Using the staggered adoption of these laws and heterogeneity-robust difference-in-differences estimators, we estimate their effects on police enforcement and on intimate partner homicide. The laws raised the share of intimate-partner assaults against women that police classify as felonies by about a third, without changing the number of assaults. We find no effect on intimate partner homicide, with point estimates close to zero. Because homicide is rare and the laws vary only across states, a null could reflect low power rather than the absence of an effect. We develop a simulation-based power analysis built on the realized panel, adoption timing, and each estimator’s identifying assumptions, and find that sustained reductions of 20–30% or more would likely have been detected. Two calibration exercises, using published estimates of baseline risk and case attrition, show why plausible effects fall below this threshold. The sevenfold odds ratio implies an annual homicide risk of about 0.1% after a strangulation, and attrition from reporting to conviction means the criminal-legal channel reaches an estimated 2.7–8.2% of future homicides. At plausible effectiveness, the incarceration cost per homicide averted is many times the value of a statistical life. More broadly, a large relative risk for a subgroup need not imply a large policy effect when the base rate is low and the policy reaches few cases.

Jelveh: University of Maryland (zjelveh@umd.edu). Situ: University of Nevada, Las Vegas (xinyi.situ@unlv.edu).

1 Introduction

Between 2000 and 2019, 48 U.S. states enacted statutes reclassifying nonfatal strangulation (NFS) from misdemeanor assault to a felony offense (Training Institute on Strangulation Prevention, 2021). The original impetus was a gap in prosecution. Emerging evidence found that strangulation often leaves no visible external injuries despite potentially serious internal trauma (Hawley et al., 2001). Under existing assault statutes, prosecutors needed visible evidence of serious bodily harm to sustain felony charges, and strangulation cases were routinely charged as misdemeanors or not prosecuted (Strack et al., 2001; Pritchard et al., 2017). These findings informed the first wave of NFS statutes, beginning with Missouri in 2000 and Minnesota in 2005. The justification was corrective: strangulation was serious violence, and existing law undercharged it.

Glass et al. (2008) added a second dimension to the case for reform. Using a case-control design, they reported that prior nonfatal strangulation by an intimate partner was associated with approximately seven-fold higher odds of subsequent homicide, quantifying a risk previously identified through case reviews (Campbell et al., 2003; Snider et al., 2009). The finding increasingly accompanied the public case for adoption: as states adopted, newspaper coverage surrounding passage cited the elevated homicide risk with growing frequency, a shift we document in Section 2 (Figure 1). The Glass finding sharpened the second rationale for the laws: strangulation was increasingly framed not only as serious violence warranting felony treatment but as a quantified precursor to homicide that felony intervention might prevent. This shifted the justification from corrective to predictive, and whether the policy delivers on that rationale—whether felony intervention actually reduces homicide—is what this paper evaluates.

The distinction matters because four separate claims are easily run together: that a harm is serious, that a subgroup faces elevated risk, that a policy changes official behavior, and that a policy reduces the outcome of concern. Only the last two describe what an intervention does, and they can come apart—a policy can change how cases are handled without changing the outcome it was meant to prevent (Weiss, 1979). Even when the evidence is causal—"this program works"—

Rad: University of Oxford (abdul.rad@sociology.ox.ac.uk). We thank Maxwell Davies and Jessica Raskauskas for outstanding research assistance. We also thank the University of Chicago Crime and Education Lab, participants in the Violence, Inequality, and Comparative Economics seminar, and participants at the American Society of Criminology Annual Meeting for helpful comments and feedback. The authors received no external funding for this research and have no financial relationships or other potential conflicts of interest to disclose.

translating it into what a policy will accomplish in a particular context is often underspecified (Sampson et al., 2013). The NFS case compounds this along two dimensions. First, the evidence entering legislation was not a causal estimate but a risk association: the odds ratio speaks to who is at risk, not to how much a law changes the outcome, so even a perfectly transported estimate leaves the treatment effect unidentified. Second, it entered debate on the relative scale—a seven-fold odds ratio—rather than as a risk difference, the quantity that would help bound the effect a law can achieve. Both gaps become consequential under a preventive rationale—part of a broader turn in criminal law toward risk-based reasoning (Feeley and Simon, 1992) and preventive justification (Zedner and Ashworth, 2019)—which implicitly asks the evidence to speak to the *magnitude* of the policy's effect.

Using National Incident-Based Reporting System data on enforcement and Supplementary Homicide Reports data on intimate partner homicide, we estimate the effect of NFS statutes with the heterogeneity-robust staggered difference-in-differences estimators of Callaway and Sant'Anna (2021) and Borusyak et al. (2024). We find that NFS laws changed enforcement: the share of domestic-violence-coded female assaults classified as felonies rose after adoption, a shift consistent with reclassification rather than a change in incidence. Intimate partner homicide, by contrast, shows no detectable decline under either a broad definition or one restricted to incidents coded as intimate partner. A null on a rare outcome immediately raises the question of what the design could have detected: our simulation-based power analysis, tailored to the realized panel and adoption timing, shows that a sustained reduction of 20% or larger would likely have been detected. A reduction of that size would presume that strangulation incidents regularly lead to arrest or conviction—the pathway through which a felony law could prevent a homicide—an assumption we examine in Section 7.

A second question is what that enforcement costs. Felony reclassification can add a year or more of imprisonment to strangulation cases, and because homicide is rare among them, that punishment falls overwhelmingly on cases that would never have ended in a killing. We make that trade-off explicit, setting the homicides a felony upgrade might prevent against the additional incarceration it imposes.

The null we find is not unique to strangulation statutes. Across the domestic violence policy evaluation literature, reforms such as mandatory arrest, no-drop prosecution, and unilateral

divorce have likewise shown little clear evidence of reducing intimate partner homicide (Dugan, 2003; Iyengar, 2009; Chin and Cunningham, 2019; Aizer and Dal Bo, 2009; Stevenson and Wolfers, 2006; Koppa, 2024). Many of these studies rely on two-way fixed effects in staggered-adoption settings, where estimates have proven sensitive to data and specification choices (Goodman-Bacon, 2021; Chin and Cunningham, 2019); our heterogeneity-robust estimators and formal power analysis instead let us interpret the null against a concrete detectable-effect threshold rather than as merely inconclusive.[1] More broadly, the paper connects to work on risk communication and evidence utilization by tracing how an epidemiological risk statistic traveled from publication into legislative adoption and then into retrospective policy evaluation (Gigerenzer and Edwards, 2003; Akl et al., 2011; Weiss, 1979; Bowen et al., 2009).

2 Background

Before the early 2000s, strangulation in domestic violence cases was usually prosecuted under general assault statutes, often as misdemeanor assault. Concern about these charges grew because visible injury is an unreliable indicator of harm: in a review of 300 cases, roughly half of victims showed no visible external injury despite serious internal trauma (Hawley et al., 2001). This likely contributed to undercharging under assault frameworks that relied on visible evidence.

Subsequent research documented the prevalence of nonfatal strangulation and its associated harms in intimate partner violence. A survey of women in domestic violence shelters found that 68% reported experiencing at least one strangulation incident by an intimate partner (Wilbur et al., 2001). National surveys rarely isolate strangulation: the National Intimate Partner and Sexual Violence Survey (NISVS) folds it into a broader category, reporting that about one in four women and one in seven men have experienced severe physical violence—including being choked or suffocated—by an intimate partner in their lifetime (Leemis et al., 2022). A systematic review that isolated strangulation found that between 3% and 10% of women reported ever having been strangled by an intimate partner (Sorenson et al., 2014). Medical research links NFS to neurological injury, stroke, psychological trauma, and in some cases delayed death (Faugno et al., 2013).

[1] We also show in Appendix D that the larger homicide reductions reported by de Assis et al. (2026) are sensitive to missing SHR data and estimator choice.

States began responding to these concerns with statutory changes. Missouri enacted specialized strangulation legislation in 2000, classifying domestic assault by strangulation as a felony offense. Minnesota followed in 2005 with legislation creating a specific felony offense for domestic assault by strangulation. During this period, epidemiological research also examined the relationship between strangulation and homicide risk. Using a case-control design, Campbell et al. (2003) studied 220 intimate partner femicide cases and 343 abused control women across 11 cities and reported that 56% of femicide victims had experienced prior strangulation compared to 10% of abused women who survived.

Glass et al. (2008) expanded this line of work by including attempted homicides alongside completed femicides, increasing the sample to 506 cases and enabling interviews with survivors. The study reported that prior strangulation was associated with substantially higher odds of subsequent intimate partner homicide (an odds ratio of approximately seven). This estimate was frequently cited in policy discussions around NFS legislation. Using newspaper coverage around each state's law adoption, we find that references to the underlying homicide-risk research—an invoked study or a numeric "times-higher" figure—grew more common over time: 0% of states adopting in 2000–2004 had such coverage, rising to 60% of those adopting from 2015 onward (Figure 1). In some cases the statistic entered enacted law directly: Washington's 2021 legislative findings state that domestic violence victims who have been nonfatally strangled are "eight times more likely to become a subsequent victim of homicide at the hands of the same abusive partner" (2021 Wash. Sess. Laws ch. 269, § 1).

By 2019, 48 states had enacted specialized NFS legislation (Training Institute on Strangulation Prevention, 2021). Although the statutes vary, most define strangulation as impeding breathing or blood circulation, remove an explicit visible-injury requirement, and classify NFS as a felony offense. Some are limited to intimate partner or domestic violence contexts.

2.1 Research Translation and Risk Communication

The NFS case illustrates a specific research-to-policy pathway: research informs a change in the *law* rather than a decision to adopt or expand a particular *program*. In this setting, the policy object is a rule change (e.g., creating a new offense, reclassifying conduct, increasing sentencing

exposure), and research functions less as an instrumental input—an estimate of the law's expected effect that feeds the decision—than as a way to frame problem severity and make the case for felony classification (Weiss, 1979).

Legislative translation is also mediated by institutions and products that differ from the typical evidence-to-program pipeline. Evidence is commonly packaged into testimony, summaries, and talking points, while bill analyses and fiscal notes emphasize statutory meaning, feasibility, and budget effects (Mosley and Gibson, 2017; North Carolina General Assembly, Fiscal Research Division, 2026; Colorado General Assembly, Legislative Council Staff, 2026). Because legislatures operate under tight calendars and episodic windows for action, these products tend to privilege short, portable statistics over multi-step causal models with uncertainty and reach assumptions (Aula, 2023).

The separation between these two tracks is visible in the legislative records for NFS bills. Kansas's HB 2034 illustrates the pattern: in committee testimony, a prosecutor cited the Glass et al. odds ratio as a roughly 750% increase in attempted homicide risk; in the fiscal note for the same bill, the quantified content was a prison-bed range from the Kansas Sentencing Commission and a per-day contract bed cost.[2] The risk statistic and the cost projection appeared in the same legislative record but were not connected: no document asked what the cited risk association implies for the expected effect of the statute on intimate partner homicide, and whether the cost would be justified. The pattern extends beyond NFS legislation. New Jersey's 2017 bail reform—supported by a judiciary-led commission and national foundation research—produced risk-tool validation and cost projections but no prospective estimate of the reform's expected effects on pretrial public safety outcomes.[3] In a different policy domain, an Oxfam insider describes how a valid estimate of the corporate tax gap was treated as a forecast of the revenue that enforcement policy would recover, sidelining alternative policies in favor of the intervention the statistic seemed to support (Lewis, 2024).

These institutional constraints interact with a well-documented risk communication issue: relative measures (risk ratios, odds ratios, "X times more likely") are more persuasive than absolute

[2] Testimony of Jessica G. Domme, Assistant Attorney General, in Support of HB 2034, House Committee on Judiciary, January 18, 2017; Fiscal Note for HB 2034, Kansas Legislature, January 18, 2017.

[3] Report of the Joint Committee on Criminal Justice, New Jersey Courts (March 2014); Legislative Fiscal Estimate, S946 (Third Reprint), New Jersey Office of Legislative Services (August 2014).

risk differences when baseline risk is small, and they are frequently presented without the baseline rates needed for impact assessment (Akl et al., 2011; Gigerenzer and Edwards, 2003; Zipkin et al., 2014). The seven-fold odds ratio is a case in point: informative as a risk marker among IPV-exposed women, but silent on how much a felony statute would change homicide incidence—a translation from relative to absolute risk we take up in Section 8.

The most directly comparable evidence concerns domestic-violence firearm laws. Firearms are the most common means of intimate partner homicide—a firearm is used in a majority of cases (56% of single-victim female intimate partner homicides in the SHR, 1990–2019)—and these laws, like NFS statutes, target abusers and are evaluated by how much they prevent homicide. That record is mixed, and the effects track a specific feature: whether the law couples a firearm prohibition with an enforceable requirement that prohibited abusers surrender their guns. Díez et al. (2017) find that restraining-order firearm laws with a relinquishment provision are associated with intimate partner homicide rates roughly a tenth lower, while otherwise-similar laws without one are not, and Zeoli et al. (2018) and Raissian (2016) report the same feature-dependence. Prohibitions triggered by a misdemeanor conviction alone have generally not moved the outcome (Vigdor and Mercy, 2006), and a synthesis of this literature grades even the best-supported of these policies as resting on moderate evidence, with no clear effect on total homicide (RAND Corporation, 2024).

A second, larger literature evaluates criminal-legal responses to domestic violence and documents a recurring gap between enforcement and outcomes. The Minneapolis arrest experiment (Sherman and Berk, 1984) reported that arrest reduced repeat abuse and prompted mandatory-arrest adoption across the country, but the federally funded replications that followed did not reproduce it cleanly: several found no deterrent effect, and in two sites arrest reduced repeat violence among employed suspects while increasing it among unemployed ones (Maxwell et al., 2002). No-drop prosecution and specialized domestic-violence courts show the same divergence, raising conviction rates and shortening case processing without a demonstrated reduction in re-offending (Davis et al., 2001; Cissner et al., 2013). A natural experiment in Milwaukee is more cautionary still: when prosecutors expanded the share of arrests carried forward regardless of victim preferences, case processing slowed and pretrial re-offending rose (Davis et al., 2003). Across these evaluations, mandated legal responses move the enforcement metric more reliably than the victimization they are meant to reduce.

Criminological theory also sharpens expectations about what felony reclassification should accomplish. Deterrence research consistently finds that the certainty of apprehension deters more reliably than the severity of punishment (Nagin, 2013; Durlauf and Nagin, 2011; Chalfin and McCrary, 2017). The empirical record on severity-margin statutes bears this out: a widely cited finding that California's Proposition 8 sentence enhancements sharply reduced serious crime (Kessler and Levitt, 1999) did not replicate under re-analysis with the full annual series, in which the decline proved to predate the law (Webster et al., 2006). NFS statutes operate primarily on the severity margin: they raise the sanction attached to conduct that was already criminal and make felony charges easier to sustain, but they leave detection and reporting unchanged, and fewer than half of intimate partner violence victimizations are reported to police (Bureau of Justice Statistics, 2024). Intimate partner violence also occurs in private settings and within relationship dynamics that are not obviously responsive to marginal changes in expected sentence length.

3 Data

3.1 Supplementary Homicide Reports

We use the FBI's Supplementary Homicide Reports (SHR), which provide incident-level information on homicides reported to law enforcement. We restrict to single-victim incidents involving women aged 18–49 and aggregate to a state-year panel for 1990–2019. Our primary outcome is the intimate partner homicide (IPH) rate per 100,000 women.

SHR reporting is incomplete for some state-years. For the SHR analysis, rather than excluding entire states with incomplete reporting, we drop only those state-years flagged as unreliable under the reporting-coverage screen used in the Office of Juvenile Justice and Delinquency Prevention's (OJJDP) Easy Access to the FBI's Supplementary Homicide Reports (EZASHR) tool (Office of Juvenile Justice and Delinquency Prevention, 2026), yielding an unbalanced panel of 1350 state-year observations across 47 states. Appendix Table 9 lists the 73 excluded state-years across 14 states; notably, Alabama loses all post-treatment observations and Nebraska loses most of its pre-treatment period.

The SHR victim–offender relationship field is recorded for 71% of single-victim female homicides aged 18–49 in our window; the remaining cases are coded unknown. We therefore construct

two outcome definitions intended to bracket the true IPH count. The *narrow* definition includes only cases where the relationship is explicitly coded as an intimate partner (spouse, ex-spouse, boyfriend/girlfriend, or common-law partner).[4] This definition is conceptually closest to IPH but will undercount when the relationship field is missing. The *broad* definition includes all single-victim female homicides aged 18–49 regardless of relationship coding. This definition is robust to missing relationship data but will overcount IPH by including stranger and acquaintance homicides. A sufficiently large reduction in IPH would ordinarily be expected to appear in both measures, although the broad measure may be attenuated by the inclusion of non-intimate-partner homicides.

We construct analogous male homicide outcomes as *negative-control outcomes*—groups the targeted enforcement should not move if the main effect is real. A null estimate there is a *specificity* check: evidence that the effect is specific to the targeted group rather than a coincident trend. Because NFS enforcement operates in intimate partner violence contexts in which victims are predominantly women, male victims serve this role. We add two further negative controls that hold sex fixed while shifting the other targeted dimensions: intimate-partner homicides of women aged 50–70, above the 18–49 range in which the enforcement response concentrates, and stranger-perpetrated homicides of women aged 18–49, outside the intimate-partner relationship the laws address. We report all three as specificity tests (Section 5, Table 3). Population denominators come from the Surveillance, Epidemiology, and End Results (SEER) program population estimates by sex, age, and state-year.

3.2 National Incident-Based Reporting System

We measure enforcement responses using the FBI's National Incident-Based Reporting System (NIBRS), which records incident-level data including offense type and, when available, victim–offender relationship. Unlike the Uniform Crime Reports' summary counts, NIBRS records offenses within incidents, allowing us to distinguish aggravated from simple assault classifications.

We construct a state-year panel from continuously reporting agencies. An agency-year enters

[4] The intimate-partner category comprises current and former spouses, dating partners, and common-law partners, following the relationship coding used in the SHR. Incidents recorded under a same-sex partnership code, a small share of the total, are not included in the intimate-partner count; we apply this convention in both datasets so the definition is consistent across sources.

only if the agency reports in all twelve months of that year, and an agency with any incomplete or missing year inside its reporting span is excluded entirely, so every included agency reports completely and without interruption over its span. Spans may begin after 1991 or end before 2019, but for agencies in adopting states the span must cover the three years before the state's adoption and the three years beginning with it; agencies in never-adopting states must report for at least six consecutive years. State-year outcomes aggregate over these agencies. We restrict the analysis window to 1991–2019. Appendix Table 6 shows that the enforcement results are consistent regardless of these sample-construction choices. We restrict to single-victim incidents involving adults aged 18–49. Our primary enforcement outcome is the *aggravated assault share*, defined as the fraction of assault incidents classified as aggravated rather than simple. In NIBRS, aggravated assault is the felony-level classification and simple assault the misdemeanor-level classification—both recorded by the reporting police agency, not a prosecutor's charge or a court disposition—so this share captures the felony-reclassification margin that NFS statutes target: the classification of conduct that would otherwise be recorded as simple assault. Table 1 shows that total assault rates (aggravated plus simple, per 100,000) are largely unchanged following NFS law adoption.

As in the SHR, the victim–offender relationship in NIBRS is not always identified—the field carries a substantive relationship code, rather than an explicit unknown or a blank, for 92% of female assault incidents, versus 71% of female homicides in the SHR—so we construct two versions of the aggravated share. The *IP-coded* measure restricts to incidents where the relationship is recorded as an intimate partner. This is precise but will undercount DV assaults when the relationship field is missing. The *proxy* measure includes all single-victim adult female assault incidents regardless of relationship coding. As with the SHR outcome, male victims serve as a negative control, as do the same age and relationship specificity tests applied to the aggravated share.

NIBRS also records how each incident is resolved, which lets us follow the enforcement response past classification. An incident is *cleared by arrest* when at least one arrest is recorded for it. It is *cleared exceptionally* when the agency has identified the offender and established grounds for an arrest but circumstances outside its control prevent one—the offender has died, prosecution is declined, extradition is refused, or the victim stops cooperating. Our *clearance rate* is the share of incidents cleared by either route; the *arrest rate* counts only clearances by arrest.

NIBRS coverage expands over time and is not nationally representative in early years. Our state-

year panel is therefore unbalanced; the estimators described in Section 4 permit an unbalanced panel. Appendix Figure 2 shows the rollout by state-year: when each state adopts, which state-years fall outside NIBRS coverage, and which enter the balanced event-study subsample.

4 Empirical Strategy

We use staggered adoption of felony nonfatal strangulation statutes across U.S. states to estimate their effects on assault classifications and intimate partner homicide. We obtain adoption dates from the Training Institute on Strangulation Prevention's legislative database (Training Institute on Strangulation Prevention, 2024), updated through manual review of state criminal statutes; Appendix Table 7 lists each state's legislative passage year and flags exceptions. We code each state's treatment year as this legislative passage year. Adoption is staggered—most states enacted felony NFS statutes between 2000 and 2019—which allows comparisons of outcomes in newly treated states to those in not-yet-treated states. We exclude Oregon, which enacted a strangulation statute in 2003 but did not classify NFS as a felony until 2019.[5] For the SHR analysis, we apply the state-year missingness filter described in Subsection 3.1. Our identifying assumption is that, absent adoption, outcomes in treated states would have followed the same trends as outcomes in states that had not yet adopted NFS laws. Appendix Figure 1 shows the staggered adoption across the SHR panel, marks the state-years the OJJDP screen excludes, and outlines the balanced subsample used in the event study.

We study enforcement outcomes (the NIBRS aggravated assault share) and lethal-violence outcomes (the SHR homicide rate), each constructed in the two variants defined in Section 3 (Subsections 3.2 and 3.1).

We estimate treatment effects using two approaches designed for staggered adoption settings; they differ in how they construct the counterfactual—what a treated state's outcome would have been absent the law. First, we use the doubly robust estimator of Callaway and Sant'Anna (2021).[6] It estimates a separate effect for each adoption cohort in each post-adoption year—a group-time

[5] Oregon's 2003 law addressed strangulation but retained misdemeanor classification; its felony statute passed in 2018 (the year in Appendix Table 7) and took effect in 2019.

[6] The term "doubly robust" refers to how the estimator is built: it combines a model of the outcome with a model of which states adopt in which year, and returns consistent estimates as long as at least one of the two models is correctly specified.

effect, the change for states adopting in a given year benchmarked against states not yet treated in that year—and averages these into an overall average treatment effect on the treated (ATT). Second, we use the imputation estimator of Borusyak et al. (2024). It fits a model with state and year fixed effects to the untreated state-years only, uses that model to predict what each treated state would have experienced without the law, and defines the ATT as the average gap between observed and predicted outcomes. Because the two estimators construct the counterfactual under different assumptions, agreement between them makes a result robust to that choice rather than an artifact of one approach.[7]

We cluster standard errors at the state level—the unit of treatment assignment—to accommodate serial correlation within states over time (Bertrand et al., 2004); both estimators absorb fixed state differences and common year shocks. We estimate without time-varying controls. With few treated states spread across many adoption cohorts, adding covariates inflates the pre-treatment coefficients' standard errors several-fold, leaving the controlled specifications too unstable and imprecise to interpret and precluding a parallel-trends test. We report specifications that add the state unemployment rate (Bureau of Labor Statistics, Local Area Unemployment Statistics) and overall crime rate (FBI Uniform Crime Reports, Crimes Known to Law Enforcement) in Appendix Table 11 but do not rely on them. Our main analyses use the unbalanced panel described in Section 3. The event-study estimates reported below are instead computed on a presence-filtered sample requiring each state to be present at every event time in the estimation window, so the plotted dynamics reflect a fixed set of units rather than changing panel composition. The full unbalanced panel comprises 47 states (45 treated) and 1350 state-years for the SHR analysis and 29 states (27 treated) and 597 state-years for NIBRS; the balanced sample retains the 38 and 21 states (36 and 19 treated, respectively) observed at every event time, dropping late adopters that lack a full pre/post window. Re-estimating the main results on this balanced sample leaves the headline conclusions intact—the enforcement effect on the aggravated-assault share persists and the female-victim homicide estimates remain null (Appendix Tables 4 and 5).

We present results in several steps. First, we estimate effects on enforcement outcomes in NIBRS to assess classification responses to NFS statutes. Second, we estimate effects on homicide

[7] Both estimators also avoid a bias that can affect conventional two-way fixed-effects regressions, which can implicitly place negative weights on some cohorts' effects when treatment effects differ across adoption cohorts or over time (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020).

outcomes in SHR. We then conduct a simulation-based power analysis to characterize detectable effect sizes, translate the odds ratio into absolute risk levels and plausible reach through the criminal-legal process, and evaluate cost-effectiveness.

5 Results

5.1 Direct Effect: Law Enforcement Response

Table 1 reports the effects of NFS laws on five enforcement outcomes in NIBRS: the aggravated assault share, the total and aggravated assault rates (which separate a change in how assaults are classified from a change in how many occur), and the clearance and arrest rates. Together they trace the enforcement response from how assaults are classified through whether they are cleared and result in arrest. Each is estimated for female victims: Panel A uses the proxy definition (all single-victim adult female assaults) and Panel B the IP-coded definition restricted to recorded intimate-partner incidents. Male victims serve as a sex negative control, reported alongside the age and relationship specificity tests in Table 3.

The estimates indicate a shift toward felony-level classifications for female-victim assaults under both measures. The proxy measure (Panel A) yields a 0.021 share-point increase (SE = 0.012), a 17.5% rise over the baseline mean; the IP-coded measure (Panel B) yields a 0.037 share-point increase (SE = 0.015), or 33.6% over baseline. Estimates are similar across estimators; the IP-coded estimate is statistically significant, while the broad-proxy estimate is significant only at the 10% level.

Figure 2 plots event-study estimates for assault classifications (see Section 4) for both estimators, with female victims in the top panel and the male negative control—discussed with the other specificity tests below—in the bottom panel. For female victims, pre-treatment coefficients are stable and near zero for both estimators, and post-treatment estimates are positive.

The other enforcement margins move little. For IP-coded incidents, the overall volume of assaults is essentially unchanged—the total assault rate changes by -22.531 per 100,000 (p = 0.687), not distinguishable from zero. That the aggravated share rises while the total volume holds flat is consistent with existing assaults being reclassified across the felony threshold rather than with a change in how many assaults occur; the aggravated rate points the same way (13.284 per 100,000,

p = 0.247), as reclassification of simple into aggravated assaults implies. Clearance and arrest rates are similar: the IP-coded clearance rate changes by 0.018 (p = 0.168) and the arrest rate by 0.008 (p = 0.596), with wide standard errors on both. The response we can detect is concentrated in how incidents are classified rather than in whether they are cleared or result in arrest.

Table 3 collects the specificity tests for both outcomes, each varying one targeted dimension from the intimate-partner female victims aged 18–49 who anchor the main estimates; we discuss the enforcement tests here and the homicide tests alongside the SHR results below. The first dimension is sex: because victims in the targeted enforcement context are predominantly women (Section 3), large effects on male outcomes would be unexpected. For assault classifications, the Callaway–Sant'Anna estimator yields null estimates in both panels: the broad proxy shows a 0.010 share-point change (SE = 0.016, p = 0.557), and the IP-coded definition shows a 0.005 share-point change (SE = 0.016, p = 0.748). The imputation estimator yields a null estimate under the broad proxy (a 0.019 share-point change, p = 0.300) but a larger, statistically insignificant change under the IP-coded definition (0.021, p = 0.321). The event study attributes this larger estimate to a pre-existing trend rather than the law: in the male panels of Figure 2, the imputation estimator's pre-treatment coefficients trend negative and its post-treatment estimates rise in step, while the Callaway–Sant'Anna coefficients stay near zero throughout.

The enforcement effect also disappears when the sample shifts along age or victim–offender relationship: for intimate-partner female victims aged 50–70 the aggravated-share estimate is -0.032 share points (p = 0.583), and for stranger-perpetrated assaults among 18–49-year-olds it is -0.061 (p = 0.189). The reclassification response concentrates among intimate-partner victims aged 18–49 and is absent in the negative-control groups—the pattern expected if the response is specific to the targeted enforcement context rather than a general change in how assaults are recorded.

5.2 Effect on Intimate Partner Homicides

Table 2 reports the effects of NFS laws on homicide outcomes in the SHR. Panel A reports the broad outcome and Panel B the narrow IP-coded outcome, both for female victims; male victims serve as a sex negative control (Table 3).

Under both measures, the two estimators agree, and neither is distinguishable from zero.

Under the broad measure (Panel A), Callaway–Sant'Anna estimates an ATT of -0.155 (SE = 0.165, p = 0.348) and the imputation model -0.056 (SE = 0.179, p = 0.755). Under the narrow, IP-coded measure (Panel B), both again yield null estimates (Callaway–Sant'Anna: 0.075, SE = 0.113; the imputation estimator is likewise statistically insignificant).

Figure 3 plots event-study estimates for homicide outcomes, with female victims in the top panel and the male negative control in the bottom panel. Pre-treatment coefficients for female outcomes are noisy but centered near zero under both estimators, and post-treatment estimates are null.

For male homicides, the two estimators diverge: Callaway–Sant'Anna returns a near-zero estimate (0.063, SE = 0.600, p = 0.916), while the imputation estimator returns a larger, imprecise decline that reaches significance only at the 10% level (-1.089, SE = 0.637, p = 0.087). The male panels of Figure 3 tie the imputation estimate to unstable pre-treatment trends on these sparse male-homicide counts—its pre-treatment coefficients are visibly less stable than Callaway–Sant'Anna's—and the disagreement between estimators points the same way. Neither provides evidence of a male-specific homicide effect, as expected for a policy targeted at intimate partner violence. The homicide estimates are also null on the age and relationship cuts (0.094 per 100,000 for victims aged 50–70, p = 0.278; -0.026 for stranger homicides, p = 0.496; Table 3).

Across outcomes, we find clear evidence of enforcement uptake but no detectable reduction in intimate partner homicides. The two estimators agree on the female homicide null under both the broad and narrow definitions, and neither male negative control shows a robust effect. Because Callaway–Sant'Anna and imputation construct the counterfactual in different ways—the first comparing each adopting cohort to states not yet treated, the second imputing outcomes from a model fit on untreated state-years—their agreement makes the homicide null robust to that modeling choice.

Interpreting this null requires three further steps. First, Section 6 asks what size homicide reduction this research design could have detected, so the null can be read against a concrete sensitivity benchmark rather than as a bare failure to reject. Second, Section 7 asks whether an effect of that size is plausible through the law's most direct mechanism, the criminal-legal processing of the prior strangulation incident. Third, Section 8 translates the homicide-prevention margin into a break-even framework, comparing the value of lives potentially saved with the prison

time added by felony treatment and making explicit where other benefits outside our data would enter.

6 Power Analysis

We complement the staggered difference-in-differences estimates with a simulation-based power analysis designed to answer a practical question: if NFS laws reduced intimate partner homicides by a given percentage, how likely is our study design to detect it? Because power in complex observational designs—difference-in-differences on rare outcomes among them—cannot be read from a standard formula, Black et al. (2022) estimate it by simulation: they randomly assign a pseudo-treatment to real untreated data, impose a known effect, and re-run the estimator to measure how often it is detected.[8] We build on that approach with two changes. Rather than randomly assigning treatment, we preserve this paper's actual staggered adoption timing and construct the no-law counterfactual to satisfy each estimator's parallel-trends assumption. We then impose a grid of hypothetical effect sizes, re-run each of the main-analysis estimators, and record how often each imposed effect is detected — under both the main analysis's off-the-shelf inference and a wild-bootstrap correction for the few-cluster setting. Because the estimators differ in the parallel-trends assumption each requires, we run the analysis under each, so the reported sensitivity does not rest on a single identifying assumption.

The simulation constructs a "no-law" counterfactual path in the observed panel, imposes a known post-adoption effect, and re-estimates the models used in the main analysis. We build that counterfactual both under each estimator's own identifying assumptions (a matched configuration) and, as a stress test, under the other's (a mismatched configuration), so the diagnostics show both whether the pipeline recovers the null when assumptions hold and how much bias a violation introduces. We evaluate a grid of hypothetical effects defined as multiplicative changes in post-adoption outcomes, and summarize detection probabilities as the fraction of simulation draws in

[8] This is distinct from post-hoc (or "observed") power, which plugs a study's realized effect estimate and standard error into a power formula after the fact. Observed power is a monotone function of the p-value and so conveys nothing beyond it: a non-significant result mechanically yields low observed power, which therefore cannot be read as evidence that the design was underpowered (Hoenig and Heisey, 2001). Our analysis instead imposes a grid of hypothetical effect sizes and asks how often each would be detected, characterizing the design's sensitivity independently of the realized estimate.

which the null of no effect is rejected at the 5% level. Appendix A provides the formal setup, and Appendix B reports diagnostics that verify the simulation pipeline is well-behaved under matched data-generating-process–estimator configurations. The simulation pipeline is implemented in the open-source R package `staggeredpower`.

6.1 Results

Figure 4 reports simulation-based power curves for both female SHR homicide outcomes—the broad definition and the narrower IP-coded definition—under off-the-shelf cluster-robust inference and the restricted wild cluster bootstrap. For homicide the Callaway–Sant'Anna estimator is already correctly sized off the shelf, so it needs no bootstrap correction; it reaches 80% power against reductions of about 20–30%. The imputation estimator's off-the-shelf test over-rejects in this few-cluster setting, so we read its power under the wild bootstrap, where it reaches 80% power against reductions of about 20%.

These detectability results help interpret the homicide estimates in Section 5. The female-victim estimates are small and statistically insignificant under both estimators—the imputation point estimate is -1.7%—and fall well below the 20% reduction the design can reliably detect. The design therefore makes large homicide reductions unlikely to have gone undetected but cannot resolve smaller ones. How much of that homicide risk the criminal-legal process can actually reach—and thus how large an effect the direct channel could plausibly produce—is the subject of Section 7.[9]

7 The Reach of the Direct Criminal-Legal Channel

The power analysis showed that our design can reliably detect sustained reductions in intimate partner homicide of 20% or larger. Whether a reduction of that size is realistic depends on how the law would actually prevent a homicide. A felony statute can work directly, through the criminal-

[9] The same simulation framework calibrates detectability for the enforcement outcomes. Read at a hypothesized effect equal to each observed Callaway–Sant'Anna estimate, the design's independently estimated power curve gives about 89% power for the IP-coded margin (33.6%) and about 65% for the broad proxy (17.5%); the Callaway–Sant'Anna test is well sized off the shelf on these shares, so no bootstrap correction is needed. The IP-coded enforcement response is statistically significant, and the design is well powered to detect a change of this size; the broad-proxy response is significant only at the 10% level.

legal processing of strangulation cases, or indirectly, by deterring abuse or changing victim and system behavior. We focus on the direct pathway and return to the indirect channels at the end of the section, since the difference-in-differences estimate already includes them. The remainder of this section bounds the direct pathway's reach: the share of intimate partner homicides it could plausibly avert.

A felony NFS statute can affect future homicides through the direct criminal-legal channel only when the prior NFS incident enters the justice system. A prior NFS incident must be reported, result in arrest, lead to felony charges, and produce a conviction that triggers incapacitation or firearm prohibition. Each stage reduces the share of potential cases the law can reach. A benchmark for the fraction of intimate partner homicides the direct channel could reach, under the stage probabilities below, is

$$\tau_{\max} \;=\; p_{S|H}\; p_R\; p_A\; p_C\; p_V. \tag{1}$$

where $p_{S|H}$ is the share of homicide victims with a prior strangulation, and p_R, p_A, p_C, and p_V are the reporting, arrest, charging, and conviction rates, each conditional on the previous stage. The bound is deliberately generous: it counts every case that clears all four stages and assumes each one averts a homicide that would otherwise occur.

To benchmark this quantity, we use literature-based inputs for the four stages. The NCVS estimates that 47% of intimate partner violence victimizations are reported to police (Bureau of Justice Statistics, 2024), and of those reports, 39% result in arrest (Reaves, 2017). A meta-analysis of 135 studies reports that approximately 60% of domestic violence arrests lead to charges and roughly 55% of charged cases result in conviction (Garner and Maxwell, 2009). Substituting,

$$\tau_{\max} \;=\; 0.448 \times 0.47 \times 0.39 \times 0.60 \times 0.55 \;\approx\; 0.027. \tag{2}$$

Even if a felony conviction were perfectly protective against subsequent homicide, the reduction achievable through the conviction channel would be about 2.7% under these probabilities. Treating arrest rather than conviction as the operative mechanism—for example, via short-term incapacitation—gives $0.448 \times 0.47 \times 0.39 \approx 0.082$, or roughly 8.2%. These stage probabilities come

from general domestic-violence populations and may differ for strangulation cases or for cases that later end in homicide. The figure is thus a literature-calibrated benchmark under stated assumptions rather than a hard upper bound.

Whether the operative stage is conviction or arrest, this benchmark sits well below the 20% minimum detectable effect established in Section 6, so the direct channel alone is unlikely to have produced a detectable reduction. The difference-in-differences design, however, estimates the law's total effect—the direct channel together with any indirect ones, such as deterrence and changes in victim or system behavior. The null therefore covers both: whatever the indirect channels added to the direct pathway, the combined effect remained too small for the design to detect.

8 Homicide-Prevention Benefits and Incarceration Costs

Because the direct criminal-legal pathway can reach only a share of future intimate partner homicides, we next compare the homicides a felony upgrade might prevent with the prison time it adds. A felony law may also generate benefits our design cannot directly measure, including fewer nonfatal injuries, more victims connected to services, deterrence, and a clearer legal signal that strangulation is serious. We therefore treat the cost-benefit analysis as a structured accounting exercise: first valuing the homicide-prevention benefit we can estimate, and then showing how the conclusions change once these additional, unmeasured benefits are allowed to enter.

8.1 The homicide risk behind a strangulation

The key benefit-side input is the probability that a person strangled by an intimate partner is later killed by that partner. This is an absolute, forward-looking risk. It is therefore different from the seven-fold odds ratio discussed in Section 2.

The odds ratio comes from a case-control study. In that design, researchers separately sample women who were killed and abused women who survived; the sizes of those groups are set by the study design rather than by their frequency in the population. As a result, the study can show that prior strangulation was much more common among homicide victims than among comparison women, but it cannot by itself tell us the probability that a strangled woman will later be killed. For the cost-benefit analysis, we need a prospective quantity: the risk that follows an IPV-exposed

population forward from an index event, analogous to actuarial tools such as the Public Safety Assessment, which estimate the probability of later observed outcomes such as failure to appear or rearrest (VanNostrand and Lowenkamp, 2013).

Converting the case-control odds ratio into an absolute risk requires a baseline risk that the case-control design does not supply (Grant, 2014): the annual rate of intimate partner homicide among IPV-exposed women. We estimate that baseline by dividing the annual number of intimate partner homicides of women aged 18–49 in the SHR by the estimated number of IPV-exposed women in the same age range. Because that exposed population is not directly observed, we do not rely on a single point estimate. Instead, we bracket the baseline using two prevalence assumptions: the two-year prevalence from the survey Glass et al. (2008) used to recruit their comparison group, which implies a larger IPV-exposed population and therefore a lower homicide risk, and the 12-month prevalence from the National Intimate Partner and Sexual Violence Survey (Smith et al., 2018), which implies a smaller exposed population and therefore a higher risk.

We then combine this baseline risk with the two strangulation shares reported by Glass et al. (2008): the share of homicide victims with a prior strangulation and the share of comparison women with a prior strangulation. Bayes' theorem gives the resulting probability of intimate partner homicide conditional on prior strangulation. Appendix C provides the derivation and inputs.

We estimate this conditional risk at 0.079% (79 per 100,000; 95% CI 55–124) under the higher-prevalence assumption and 0.137% (137 per 100,000; 95% CI 95–218) under the more conservative (higher-risk) one, with intervals obtained by Monte Carlo simulation propagating sampling and prevalence uncertainty. In both cases the absolute risk remains below 0.2%, on the order of one in a thousand (Table 4). Expressed as natural frequencies, among 100,000 IPV-exposed women followed for one year, approximately 79 to 137 would be killed by an intimate partner if all had experienced a prior strangulation, compared with 10 to 18 if none had (Gigerenzer and Edwards, 2003).

8.2 Break-even framework

We translate the estimated homicide risk into a per-case cost–benefit comparison. Let p denote the per-case homicide risk derived above. If the felony upgrade costs C dollars per NFS case and prevents homicide with probability e among those who would otherwise be killed, then expected lives saved per NFS case are $e \cdot p$, and the implied *cost per life saved* is $C/(e\,p)$. This links the estimated risk to the per-case cost required to break even for a given value of a statistical life. We report this as a *cost–benefit ratio*, the per-case cost relative to the expected homicide-prevention benefit ($C/(e\,p\,\text{VSL})$, equivalently the cost per life saved relative to the value of a statistical life); a ratio above one means costs exceed benefits, and the break-even cost reported below is the value of C at which the ratio equals one.

8.3 Parameter choices

We define the felony-upgrade cost C as the *incremental* incarceration cost of a felony NFS conviction relative to the misdemeanor charge it replaces: $C = \Delta T \times \kappa$, where ΔT is additional time served and κ is cost per inmate-year. Using sentencing data for domestic violence felonies (Smith et al., 2008), we put the imposed sentence for a felony domestic-violence conviction at roughly 27 months, against one to two months served under the misdemeanor charge it replaces. What separates our two scenarios is how much of that imposed sentence is actually served. Under parole-era release rules, violent offenders typically served about half of the imposed term (Ditton and Wilson, 1999); truth-in-sentencing laws, which most states adopted after 1994, instead require violent offenders to serve 85% of the sentence before release. Applying the two release regimes to the same imposed sentence brackets the incremental time served at 11 months (parole-era release) to 21 months (85% truth-in-sentencing).

For the per-inmate cost κ, our primary figure is the long-run *marginal* cost of \$17,922 per year (Wilson and Lemoine, 2022)—since felony reclassification adds inmate-years at the margin rather than changing average operating costs. We also carry the median state expenditure of \$60,989 per year (USAFacts, 2023) as an upper bound; this *average* cost likely overstates the true incremental cost, because it embeds fixed facility and administrative expenses that do not scale

with an additional inmate-year.[10] Crossing these inputs yields four cost scenarios ranging from approximately $16,428 to $106,731 per NFS case. These scenarios omit prosecution, court, and community-supervision costs, so if anything they understate the cost of felony processing, biasing the comparison in the law's favor.

For effectiveness, we consider $e = 1.0$ (a mechanical upper bound in which the felony upgrade prevents every homicide among those who would otherwise be killed) and $e = 0.2$ (effectiveness at the scale of the 20% minimum detectable effect from Section 6). We use values of a statistical life of $10M and $20M. The $10M figure sits near the values used in U.S. federal regulatory analysis—the central Department of Transportation value is approximately $13M in 2023 dollars (U.S. Department of Transportation, 2023)—while the $20M figure, above the federal range, stress-tests the result by roughly doubling the homicide-prevention benefit. Because the incarceration-cost inputs are likewise in 2022–2023 dollars, the cost–benefit ratios require no cross-year inflation adjustment.

8.4 Results

Table 5 reports cost–benefit ratios across the parameter grid. At a $10M VSL, costs exceed the homicide-prevention benefit in every scenario: even under the most favorable assumptions—marginal incarceration cost, short incremental sentence, perfect effectiveness, and the higher absolute-risk estimate—the ratio is 1.20. The $20M stress test pulls that single most-favorable scenario below break-even (0.60), but reaching it requires doubling the standard value of a statistical life and assuming the felony upgrade prevents every homicide among those who would otherwise be killed. Under more plausible assumptions ($e = 0.2$), ratios range from 3.0 (at a $20M VSL) to 67.7 (at $10M) across the cost and value-of-life grid. The bottom row reports the break-even per-case cost $C^* = \text{VSL} \times p \times e$, the maximum expenditure at which homicide-prevention benefits equal costs. At $e = 0.2$ and a $10M VSL, break-even costs range from $1,576 to $2,737; even doubled at the $20M VSL, they remain well below the incarceration costs implied by the sentencing scenarios.

[10] The HB 2034 fiscal note (Section 2) priced the additional prison beds at a contract rate of $40 per day, or about $14,600 per bed-year—close to the marginal per-inmate cost used here ($17,922) and below the median-state figure. Substituting the state's rate for ours leaves the per-case cost above the break-even benchmarks reported in Table 5, so the cost–benefit comparison does not depend on the higher per-inmate figures.

These calculations also clarify why "detectable" effects can still be economically implausible. The MDE exceeds the plausible reach calculations derived in Section 7, so the policy-relevant region is one in which effects may be too small to detect and too small to generate favorable cost–benefit ratios unless per-case costs are low or benefits extend well beyond homicide prevention. Concretely, at $e = 0.2$, the non-homicide benefits of felony processing would need to be 2 to 67 times as large as the homicide-prevention benefit for the upgrade to break even.

9 Discussion and Implications for Policy and Research

Felony NFS statutes produced a large enforcement response without a detectable reduction in intimate partner homicide. The power analysis clarifies the scope of that null: sustained reductions of 20% or larger would likely have been detected, but smaller effects remain below the design's reliable reach. Sections 7 and 8 explain why that is the policy-relevant region. The direct criminal-legal pathway reaches only a small share of future homicides, and the absolute homicide risk following prior strangulation is low enough that incarceration costs can exceed expected homicide-prevention benefits unless effects are very large, per-case costs are low, or benefits extend beyond homicide prevention.

These findings speak to the predictive justification for NFS statutes—that felony intervention would reduce intimate partner homicide. The corrective justification, that strangulation is serious violence and prior law undercharged it, is retributive and expressive—an assertion of society's condemnation of the conduct, independent of any deterrent effect (Feinberg, 1965): it does not depend on demonstrating homicide reduction. The enforcement results confirm that the laws changed classification behavior, which is what the corrective case calls for.

The shift from corrective to predictive justification is not incidental to this case. Under broader shifts in criminal law toward risk-based reasoning (Feeley and Simon, 1992) and preventive justification (Zedner and Ashworth, 2019), large risk estimates can be taken up as if they answered a forward-looking policy question, even when they do not directly estimate the expected effect of a statute. That may help explain why the Glass odds ratio traveled as effectively as it did: it connected a present indicator to a future outcome in a form that could be read as predictive, even though it came from a case-control study rather than a policy evaluation.

This distinction has a measurement corollary. The power analysis implies that, for outcomes as rare as intimate partner homicide, plausible policy effects can lie below what any feasible state-level design can detect: the question "did the statute reduce homicide?" may be unanswerable at policy-relevant scales, not merely unanswered. The same limit applies to effect heterogeneity: because any split by state, region, or adoption cohort leaves only a handful of treated clusters, the smallest detectable effect at that resolution sits well above even the 20% reduction the pooled homicide design can detect (Section 6).

Policies will sometimes be adopted anyway—because the conduct is grave, because victims' claims to recognition are strong, and because legislatures must act under uncertainty. Our results counsel transparency about which justification is doing the work in such cases. A legislature that upgrades strangulation because the conduct deserves felony treatment makes a normative claim our findings do not bear on; a legislature that promises homicide reduction makes an empirical one, staking the policy on an effect that available evidence cannot verify and our bounds suggest is small.

The legislative process for NFS bills illustrates why this translation rarely occurs. As documented in Section 2, risk evidence and cost projections typically move on separate institutional tracks: advocates and expert witnesses present the epidemiological claims in testimony, while legislative staff estimate prison-bed or budget effects in fiscal notes. Our calculations show what an added translation step would require. For criminal statute changes, a fiscal note or companion analysis would need not only a cost estimate, but also a statement of the mechanism linking the cited risk evidence to the claimed policy effect and a prospective estimate of that effect given the outcome's base rate, the intervention's reach, and the costs triggered by felony processing. Existing state-level tools for prospective effect estimation—such as the Washington State Institute for Public Policy's benefit-cost model (Aos et al., 2011) and the Pew–MacArthur Results First model (Pew-MacArthur Results First Initiative, 2013)—were designed mainly for bounded service interventions; adapting that function to criminal statute changes would likely require housing it in a sentencing commission, legislative services office, or similar institution.

A different use of strangulation risk evidence is already common in practice: guiding protective responses rather than justifying a criminal penalty. In the Lethality Assessment Program (LAP), officers responding to domestic violence incidents administer an 11-item screen, and survivors

flagged as high danger are immediately connected to an advocate for safety planning and service referral (Messing et al., 2024). The revised Lethality Screen treats strangulation as an automatic high-danger classification within that protocol. The Danger Assessment uses strangulation as one item in triaging intervention intensity (Campbell et al., 2009). Glass et al. (2008) themselves concluded that their results underscore the need to screen for nonfatal strangulation when assessing abused women in emergency settings—a use of the evidence that matches what these tools do.

Whether screening tools by themselves reduce homicide remains uncertain: one quasi-experimental evaluation of the LAP reports large reductions in female homicide victimization, but the estimates do not survive heterogeneity-robust specifications and the design is limited to 39 agencies in a single state (Koppa, 2024). The tools' clearer contribution is connecting survivors to safety planning and services that address the far more common outcomes of ongoing violence and serious reassault.

The base-rate logic also clarifies where a detectable homicide effect is most plausible: among victims identified as high risk—whether by screening instruments or statistical risk prediction (Jelveh et al., 2026; Bhatt et al., 2024)—annual homicide risk can run well above the population base rate, so a response concentrated there operates on a margin a population-wide statute cannot target. Consistent with this, a quasi-experimental evaluation of an intensive prosecutorial intervention targeting such cases finds large, detectable reductions in victim homicide (Arora et al., 2026). Targeting has its own limit, the prevention paradox familiar from public health: most intimate partner homicide victims had no prior strangulation (a majority in the Glass et al. (2008) case sample), so high-risk strategies, however effective for those they reach, cannot address the larger share of deaths arising outside the flagged group (Rose, 1985).

Several limitations qualify these conclusions. First, the homicide analysis depends on SHR data, which have known missingness problems. States vary in the completeness of their reporting to the SHR program, and missing state-years are not random—they are more common among states with incomplete agency-level coverage or administrative transitions. These gaps are consequential: in Appendix D, we show that the larger effects reported by de Assis et al. (2026) are sensitive to how missing state-years are handled and to the choice of estimator. Our own conclusion is not sensitive to that handling: including the state-years the OJJDP screen flags shifts the point estimates somewhat but leaves the female homicide estimates statistically indistinguishable from zero under

both estimators and definitions (Appendix Table 10).

Second, the enforcement analysis uses NIBRS, whose coverage is incomplete in early years and expands over time. Because NIBRS adoption is itself staggered and not nationally representative, our enforcement estimates reflect the experience of reporting agencies rather than the full population of agencies in each state. To the extent that early NIBRS adopters differ systematically from later adopters—for instance, in urbanicity or agency size—external validity is limited. A related concern is the thinness of the comparison group: NIBRS contains few untreated states, and by the end of the window the pool consists only of South Carolina and Ohio, both untreated because their 2023 adoptions fall outside it. The enforcement results do not hinge on this margin: excluding Ohio leaves all enforcement estimates somewhat larger, with the IP-coded estimate remaining statistically significant under both estimators (Appendix Table 6).

Third, the cost-benefit exercise values only homicide prevention and therefore omits non-lethal benefits of felony classification, including reduced nonfatal injury, improved psychological well-being, increased access to victim services triggered by felony case processing, and the expressive or norm-setting effects of treating strangulation as a serious offense. Section 8 quantifies how large these omitted benefits would need to be for felony upgrading to break even. The framework should therefore be read as a set of transparent thresholds on the homicide-prevention margin, not as a comprehensive welfare assessment. Relatedly, the design cannot distinguish whether the homicide null reflects an effective direct channel with limited reach or the absence of substantial effects through any channel. If reach is the binding constraint, expanding that reach would change both sides of the ledger: it could expose more high-risk cases to intervention, but it would also expand the number of cases subject to felony processing and incarceration.

Finally, as in other staggered difference-in-differences applications, identification rests on a standard but not directly testable parallel-trends assumption. Event-study estimates show no evidence of differential pre-trends for female enforcement outcomes, and the pre-treatment homicide coefficients are noisy but centered near zero (Figure 3); both support but cannot confirm that assumption.

Several directions for future research follow from the analysis. Screening- and service-delivery interventions that use strangulation as a risk marker have limited causal evidence on their effects on intimate partner homicide or serious reassault. Designs that exploit variation in the adoption

or geographic coverage of such interventions could test whether they produce downstream effects that felony reclassification does not. The enforcement response documented here also raises questions about what happens after felony classification: how charges are resolved through plea bargaining, what sentences result, and whether felony processing changes offender trajectories through incapacitation, probation conditions, or firearm prohibitions. Linked administrative data connecting incident reports to court records and corrections data would support that analysis, and would sharpen the cost-benefit calculation of Section 8, whose incarceration cost rests on literature-based sentencing assumptions rather than observed dispositions. Finally, the simulation-based power-analysis framework developed here could be used more broadly in staggered-adoption settings to audit published research, clarify what existing designs were capable of detecting, and improve prospective study design.

References

Aizer, Anna and Pedro Dal Bo (2009) "Love, hate and murder: Commitment devices in violent relationships," *Journal of public Economics*, 93 (3-4), 412–428.

Akl, Elie A, Andrew D Oxman, Jeph Herrin et al. (2011) "Using alternative statistical formats for presenting risks and risk reductions," *The Cochrane Database of Systematic Reviews*, 2011 (3), CD006776, 10.1002/14651858.CD006776.pub2.

Aos, Steve, Stephanie Lee, Elizabeth Drake et al. (2011) "Return on Investment: Evidence-Based Options to Improve Statewide Outcomes,"Technical Report 11-07-1201, Washington State Institute for Public Policy, Olympia, WA.

Arora, Ashna, Xander Beberman, Zubin Jelveh, and Ashley Motta (2026) "Targeted Interventions for High-Risk Domestic Violence Victims," Working paper.

de Assis, Dércio, Arpita Ghosh, Sonia Oreffice, and Climent Quintana-Domeque (2026) "Disrupting Violence, Protecting Lives: Strangulation Laws and Intimate Partner Homicides," NICEP Working Paper 2026-02, Nottingham Interdisciplinary Centre for Economic and Political Research (NICEP), Nottingham, UK.

Aula, Ville (2023) "Evidence-based policymaking in the legislatures: timeliness and politics of evidence in Finland," *Policy & Politics*, 51 (4), 673–694, 10.1332/030557321X16873376680333.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119 (1), 249–275, 10.1162/003355304772839588.

Bhatt, Monica P., Sara B. Heller, Max Kapustin, Marianne Bertrand, and Christopher Blattman (2024) "Predicting and Preventing Gun Violence: An Experimental Evaluation of READI Chicago," *The Quarterly Journal of Economics*, 139 (1), 1–56.

Black, Bernard, Alex Hollingsworth, Letícia Nunes, and Kosali Simon (2022) "Simulated power analyses for observational studies: An application to the Affordable Care Act Medicaid expansion," *Journal of Public Economics*, 213, 104713, 10.1016/j.jpubeco.2022.104713.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024) "Revisiting Event-Study Designs: Robust and Efficient Estimation," *The Review of Economic Studies*, rdae007, 10.1093/restud/rdae007.

Bowen, S., A. B. Zwi, P. Sainsbury, and M. Whitehead (2009) "Killer facts, politics and other in-

fluences: What evidence triggered early childhood intervention policies in Australia?" *Evidence and Policy*, 5 (1), 5–32.

Bureau of Justice Statistics (2024) *Key Findings from the 2023 NCVS: Domestic Violence*: U.S. Department of Justice, Office of Justice Programs.

Callaway, Brantly and Pedro H. C. Sant'Anna (2021) "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 225 (2), 200–230, 10.1016/j.jeconom.2020.12.001.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008) "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 90 (3), 414–427, 10.1162/rest.90.3.414.

Campbell, Jacquelyn C., Daniel Webster, Jane Koziol-McLain et al. (2003) "Risk Factors for Femicide in Abusive Relationships: Results From a Multisite Case Control Study," *American Journal of Public Health*, 93 (7), 1089–1097.

Campbell, Jacquelyn C., Daniel W. Webster, and Nancy Glass (2009) "The Danger Assessment: Validation of a Lethality Risk Assessment Instrument for Intimate Partner Femicide," *Journal of Interpersonal Violence*, 24 (4), 653–674, 10.1177/0886260508317180.

de Chaisemartin, Clément and Xavier D'Haultfœuille (2020) "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110 (9), 2964–2996, 10.1257/aer.20181169.

Chalfin, Aaron and Justin McCrary (2017) "Criminal Deterrence: A Review of the Literature," *Journal of Economic Literature*, 55 (1), 5–48, 10.1257/jel.20141147.

Chin, Yoo-Mi and Scott Cunningham (2019) "Revisiting the effect of warrantless domestic violence arrest laws on intimate partner homicides," *Journal of Public Economics*, 179, 104072, 10.1016/j.jpubeco.2019.104072.

Cissner, Amanda B., Melissa Labriola, and Michael Rempel (2013) *Testing the Effects of New York's Domestic Violence Courts: A Statewide Impact Evaluation*, New York, NY: Center for Court Innovation.

Colorado General Assembly, Legislative Council Staff (2026) "Fiscal Notes."

Davis, Robert C., Barbara E. Smith, and Heather J. Davies (2001) "Effects of No-Drop Prosecution of Domestic Violence upon Conviction Rates," *Justice Research and Policy*, 3, 1–13.

Davis, Robert C., Barbara E. Smith, and Bruce Taylor (2003) "Increasing the Proportion of Domestic

Violence Arrests That Are Prosecuted: A Natural Experiment in Milwaukee," *Criminology & Public Policy*, 2 (2), 263–282.

Díez, Carolina, Rachel P. Kurland, Emily F. Rothman et al. (2017) "State Intimate Partner Violence–Related Firearm Laws and Intimate Partner Homicide Rates in the United States, 1991 to 2015," *Annals of Internal Medicine*, 167 (8), 536–543, 10.7326/M16-2849.

Ditton, Paula M. and Doris James Wilson (1999) "Truth in Sentencing in State Prisons," Special Report NCJ 170032, Bureau of Justice Statistics, U.S. Department of Justice, Washington, DC.

Dugan, Laura (2003) "Domestic violence legislation: Exploring its impact on the likelihood of domestic violence, police involvement, and arrest," *Criminology & Public Policy*, 2 (2), 283–312.

Durlauf, Steven N. and Daniel S. Nagin (2011) "Imprisonment and Crime: Can Both Be Reduced?" *Criminology & Public Policy*, 10 (1), 13–54, 10.1111/j.1745-9133.2010.00680.x.

Faugno, Diana, Daria Waszak, Gael B. Strack, Melodie Ann Brooks, and Casey G. Gwinn (2013) "Strangulation forensic examination: best practice for health care providers," *Advanced Emergency Nursing Journal*, 35 (4), 314–327, 10.1097/TME.0b013e3182aa05d3.

Feeley, Malcolm M. and Jonathan Simon (1992) "The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications," *Criminology*, 30 (4), 449–474.

Feinberg, Joel (1965) "The Expressive Function of Punishment," *The Monist*, 49 (3), 397–423.

Garner, Joel H. and Christopher D. Maxwell (2009) "Prosecution and Conviction Rates for Intimate Partner Violence," *Criminal Justice Review*, 34 (1), 44–79, 10.1177/0734016808324231.

Gigerenzer, Gerd and Adrian Edwards (2003) "Simple tools for understanding risks: from innumeracy to insight," *BMJ*, 327 (7417), 741–744, 10.1136/bmj.327.7417.741.

Glass, Nancy, Kathryn Laughon, Jacquelyn Campbell, Carolyn Rebecca Block, Ginger Hanson, Phyllis W. Sharps, and Ellen Taliaferro (2008) "Non-fatal strangulation is an important risk factor for homicide of women," *The Journal of Emergency Medicine*, 35 (3), 329–335, 10.1016/j.jemermed.2007.02.065.

Goodman-Bacon, Andrew (2021) "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 225 (2), 254–277, 10.1016/j.jeconom.2021.03.014.

Grant, Robert L. (2014) "Converting an odds ratio to a range of plausible relative risks for better communication of research findings," *BMJ*, 348, f7450, 10.1136/bmj.f7450.

Hawley, Dean A, George E McClane, and Gael B Strack (2001) "A review of 300 attempted stran-

gulation cases part III: injuries in fatal cases," *The Journal of Emergency Medicine*, 21 (3), 317–322, 10.1016/S0736-4679(01)00401-2.

Hoenig, John M. and Dennis M. Heisey (2001) "The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis," *The American Statistician*, 55 (1), 19–24.

Iyengar, Radha (2009) "Does the Certainty of Arrest Reduce Domestic Violence? Evidence from Mandatory and Recommended Arrest Laws," *Journal of Public Economics*, 93 (1–2), 85–98, 10.1016/j.jpubeco.2008.09.006.

Jelveh, Zubin, Andrea Tentner, Xander Beberman, and Harold Pollack (2026) "Predicting Behavioral Health-Involved Police Encounters: A Machine Learning Approach," *Journal of Quantitative Criminology*, 10.1007/s10940-026-09671-9.

Kessler, Daniel and Steven D. Levitt (1999) "Using Sentence Enhancements to Distinguish between Deterrence and Incapacitation," *Journal of Law and Economics*, 42 (1), 343–363, 10.1086/467428.

Koppa, Vijetha (2024) "Can information save lives? Effect of a victim-focused police intervention on intimate partner homicides," *Journal of Economic Behavior & Organization*, 217, 756–782.

Leemis, Rachael W., Nickolas Friar, Sambista Khatiwada, Megan S. Chen, Marcie-jo Kresnow, Sharon G. Smith, Sondra Caslin, and Kathleen C. Basile (2022) *The National Intimate Partner and Sexual Violence Survey: 2016/2017 Report on Intimate Partner Violence*, Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Lewis, Mike (2024) "Are killer facts 'a strategy that can come back to bite'?," September.

Maxwell, Christopher D., Joel H. Garner, and Jeffrey A. Fagan (2002) "The Preventive Effects of Arrest on Intimate Partner Violence: Research, Policy and Theory," *Criminology & Public Policy*, 2 (1), 51–80.

Messing, Jill Theresa, Jacquelyn Campbell, Darrell Holly, and Dave Corwin (2024) "The Lethality Assessment Program 2.0: Adjusting Intimate Partner Violence Risk Assessment to Account for Strangulation Risk," *Policing: A Journal of Policy and Practice*, 18, paae060, 10.1093/police/paae060.

Mosley, Jennifer E. and Katherine Gibson (2017) "Strategic use of evidence in state-level policymaking: matching evidence type to legislative stage," *Policy Sciences*, 50 (4), 697–719, 10.1007/s11077-017-9289-x.

Nagin, Daniel S. (2013) "Deterrence in the Twenty-First Century," in Tonry, Michael ed. *Crime and*

Justice, 42, 199–264: University of Chicago Press, 10.1086/670398.

North Carolina General Assembly, Fiscal Research Division (2026) "Fiscal Notes."

Office of Juvenile Justice and Delinquency Prevention (2026) "EZASHR Methods."

Pew-MacArthur Results First Initiative (2013) *States' Use of Cost-Benefit Analysis: Improving Results for Taxpayers*, Washington, DC: The Pew Charitable Trusts.

Pritchard, Adam J., Amy Reckdenwald, and Chelsea Nordham (2017) "Nonfatal Strangulation as Part of Domestic Violence: A Review of Research," *Trauma, Violence, & Abuse*, 18 (4), 407–424, 10.1177/1524838015622439.

Raissian, Kerri M. (2016) "Hold Your Fire: Did the 1996 Federal Gun Control Act Expansion Reduce Domestic Homicides?" *Journal of Policy Analysis and Management*, 35 (1), 67–93, 10.1002/pam.21856.

RAND Corporation (2024) *The Science of Gun Policy: A Critical Synthesis of Research Evidence on the Effects of Gun Policies in the United States*, Santa Monica, CA: RAND Corporation, 4th edition.

Reaves, Brian A. (2017) *Police Response to Domestic Violence, 2006–2015*: Bureau of Justice Statistics, U.S. Department of Justice.

Rose, Geoffrey (1985) "Sick Individuals and Sick Populations," *International Journal of Epidemiology*, 14 (1), 32–38.

Sampson, Robert J, Christopher Winship, and Carly Knight (2013) "Translating causal claims: Principles and strategies for policy-relevant criminology," *Criminology & Pub. Pol'y*, 12, 587.

Sherman, Lawrence W. and Richard A. Berk (1984) "The Specific Deterrent Effects of Arrest for Domestic Assault," *American Sociological Review*, 49 (2), 261–272.

Smith, Erica L., Matthew R. Durose, and Patrick A. Langan (2008) "State Court Processing of Domestic Violence Cases,"Technical Report NCJ 214993, Bureau of Justice Statistics, Washington, DC.

Smith, Sharon G., Xinjian Zhang, Kathleen C. Basile, Melissa T. Merrick, Jing Wang, Marcie-jo Kresnow, and Jieru Chen (2018) *The National Intimate Partner and Sexual Violence Survey (NISVS): 2016/2017 Report on Intimate Partner Violence*, Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Snider, Cynthia, Daniel Webster, Chris S. O'Sullivan, and Jacquelyn Campbell (2009) "Intimate Partner Violence: Development of a Brief Risk Assessment for the Emergency Department,"

Academic Emergency Medicine, 16 (11), 1208–1216.

Sorenson, Susan B., Manisha Joshi, and Elizabeth Sivitz (2014) "A Systematic Review of the Epidemiology of Nonfatal Strangulation, a Human Rights and Health Concern," *American Journal of Public Health*, 104 (11), e54–e61, 10.2105/AJPH.2014.302191.

Stevenson, Betsey and Justin Wolfers (2006) "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress," *The Quarterly Journal of Economics*, 121 (1), 267–288, 10.1162/qjec.2006.121.1.267.

Strack, Gael B, George E McClane, and Dean Hawley (2001) "A review of 300 attempted strangulation cases part i: criminal legal issues," *The Journal of Emergency Medicine*, 21 (3), 303–309, 10.1016/S0736-4679(01)00399-7.

Training Institute on Strangulation Prevention (2021) "All 50 States: Summary of Strangulation Laws," Training Institute on Strangulation Prevention, Alliance for HOPE International.

——— (2024) "Strangulation Legislation."

U.S. Department of Transportation (2023) *Departmental Guidance on Valuation of a Statistical Life in Economic Analysis*: Office of the Secretary, U.S. Department of Transportation.

USAFacts (2023) "How Much Do States Spend on Housing Prisoners?."

VanNostrand, Marie and Christopher T. Lowenkamp (2013) *Assessing Pretrial Risk without a Defendant Interview*, Houston, TX: Laura and John Arnold Foundation.

Vigdor, Elizabeth Richardson and James A. Mercy (2006) "Do Laws Restricting Access to Firearms by Domestic Violence Offenders Prevent Intimate Partner Homicide?" *Evaluation Review*, 30 (3), 313–346, 10.1177/0193841X06287307.

Webster, Cheryl Marie, Anthony N. Doob, and Franklin E. Zimring (2006) "Proposition 8 and Crime Rates in California: The Case of the Disappearing Deterrent," *Criminology & Public Policy*, 5 (3), 417–448, 10.1111/j.1745-9133.2006.00403.x.

Weiss, Carol H. (1979) "The Many Meanings of Research Utilization," *Public Administration Review*, 39 (5), 426–431, 10.2307/3109916.

Wilbur, Lee, Michelle Higley, Jason Hatfield, Zita Surprenant, Ellen Taliaferro, Donald J Smith, and Anthony Paolo (2001) "Survey results of women who have been strangled while in an abusive relationship," *The Journal of Emergency Medicine*, 21 (3), 297–302, 10.1016/S0736-4679(01)00398-5.

Wilson, Stuart John and Jocelyne Lemoine (2022) "Methods of Calculating the Marginal Cost

of Incarceration: A Scoping Review," *Criminal Justice Policy Review*, 33 (6), 639–663, 10.1177/08874034211060336.

Zedner, Lucia and Andrew Ashworth (2019) "The Rise and Restraint of the Preventive State," *Annual Review of Criminology*, 2, 429–450, 10.1146/annurev-criminol-011518-024526.

Zeoli, April M., Alexander McCourt, Shani Buggs, Shannon Frattaroli, David Lilley, and Daniel W. Webster (2018) "Analysis of the Strength of Legal Firearms Restrictions for Perpetrators of Domestic Violence and Their Associations with Intimate Partner Homicide," *American Journal of Epidemiology*, 187 (11), 2365–2371, 10.1093/aje/kwy174.

Zipkin, Daniella A, Craig A Umscheid, Nancy L Keating et al. (2014) "Evidence-based risk communication: a systematic review," *Annals of Internal Medicine*, 161 (4), 270–280, 10.7326/M14-0295.

Tables

Table 1: Enforcement Effects of NFS Laws: Classification, Incidence, and Clearance

	Female victims	
	CS	Imputation
Panel A: Proxy Definition (all female-victim assaults)		
Aggravated assault share	0.021†	0.022†
	(0.012)	(0.013)
Total assault rate (per 100k)	-41.0	-48.4
	(88.9)	(132.7)
Aggravated assault rate (per 100k)	12.8	16.3
	(19.9)	(27.8)
Clearance rate	0.025†	0.034*
	(0.015)	(0.016)
Arrest rate	0.010	0.015
	(0.013)	(0.012)
Panel B: IP-Coded Definition (relationship-verified)		
Aggravated assault share	0.037*	0.034*
	(0.015)	(0.014)
Total assault rate (per 100k)	-22.5	-24.3
	(55.9)	(72.5)
Aggravated assault rate (per 100k)	13.3	13.8
	(11.5)	(13.6)
Clearance rate	0.018	0.014
	(0.013)	(0.013)
Arrest rate	0.008	0.002
	(0.015)	(0.011)

Notes: All estimates are for female victims. Standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The aggravated assault share, clearance rate, and arrest rate are proportions (0–1); total and aggravated assault rates are counts per 100,000. Clearance counts cases closed by arrest or exceptional means; the arrest rate counts arrest closures only. Sex, age, and relationship specificity tests are reported in Table 3. No time-varying controls included.

Table 2: Effects of NFS Laws on Intimate Partner Homicide

	Female victims	
	CS	Imputation
Panel A: Broad Definition		
ATT	-0.155	-0.056
	(0.165)	(0.179)
% of baseline	-4.6%	-1.7%
Panel B: IP-Coded Definition		
ATT	0.075	0.099
	(0.113)	(0.106)
% of baseline	5.2%	6.8%

Notes: All estimates are for female victims. Standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ATT is in intimate partner homicides per 100,000. Percent of baseline is ATT divided by the control-group mean. Sex, age, and relationship specificity tests are reported in Table 3. No time-varying controls included.

Table 3: Specificity Tests by Sex, Age, and Victim–Offender Relationship

	NIBRS agg. share		SHR homicide rate	
	CS	Imputation	CS	Imputation
Panel A: Broad Definition				
Female, 18–49 (treatment)	0.021†	0.022†	-0.155	-0.056
	(0.012)	(0.013)	(0.165)	(0.179)
Male, 18–49	0.010	0.019	0.063	-1.089†
	(0.016)	(0.018)	(0.600)	(0.637)
Female, 50–70	0.010	0.015	0.247	0.103
	(0.017)	(0.014)	(0.156)	(0.080)
Panel B: IP-Coded Definition				
Intimate partner, 18–49 (treatment)	0.037*	0.034*	0.075	0.099
	(0.015)	(0.014)	(0.113)	(0.106)
Male, 18–49	0.005	0.021	-0.045	-0.031
	(0.016)	(0.021)	(0.069)	(0.040)
Intimate partner, 50–70	-0.032	0.021	0.094	0.026
	(0.058)	(0.018)	(0.087)	(0.055)
Stranger, 18–49	-0.061	-0.030	-0.026	-0.007
	(0.046)	(0.024)	(0.038)	(0.029)

Notes: Standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The NIBRS outcome is the aggravated assault share (proportion); the SHR outcome is homicides per 100,000. Panel A uses the broad definition (all female victims; relationship not required), Panel B the IP-coded definition (relationship verified as intimate partner). Within each panel the treatment row is the targeted group, and each specificity row varies one dimension—sex, age, or victim–offender relationship—from it. The stranger cut appears only under the IP-coded definition, since isolating strangers requires the recorded relationship. No time-varying controls.

Table 4: Estimated Annual Intimate Partner Homicide Risk Conditional on Prior Strangulation

IPV Prevalence Estimate	Precision (Prior NFS)	Precision (No NFS)	Risk Difference
Glass survey (11.7% IPV prevalence)			
Median	0.079%	0.010%	0.069%
	(78.8 per 100k)	(10.2 per 100k)	(68.6 per 100k)
95% CI	[0.055-0.124%]	[0.008-0.014%]	[0.045-0.112%]
	[55-124]	[8-14]	[45-112]
NISVS (4.5% IPV prevalence)			
Median	0.137%	0.018%	0.119%
	(136.8 per 100k)	(17.7 per 100k)	(119.1 per 100k)
95% CI	[0.095-0.218%]	[0.014-0.025%]	[0.078-0.197%]
	[95-218]	[14-25]	[78-197]

Notes: Conditional risk calculated using Bayes' theorem with Glass et al. (2008) conditional probabilities. Estimates based on Monte Carlo simulation (n=10,000) incorporating uncertainty from: (1) Glass case-control sample (binomial), (2) IPV prevalence estimates (Glass survey: 427/3,637; NISVS: Smith et al. 2018), and (3) homicide count variation across years. Glass survey estimate uses population adjusted for relationship status (67%). Appendix Figure 6 plots the full distributions.

Table 5: Homicide-Prevention Break-Even Ratios

	$e = 1.0$		$e = 0.2$	
Incremental cost (C)	$p = 0.079\%$	$p = 0.137\%$	$p = 0.079\%$	$p = 0.137\%$
\$ 16,428	2.08	1.20	10.42	6.00
\$ 31,364	3.98	2.29	19.90	11.46
\$ 55,907	7.09	4.09	35.47	20.43
\$106,731	13.54	7.80	67.71	39.00
Break-even C^*	\$7,882	\$13,685	\$1,576	\$2,737

Notes: Each cell reports the cost–benefit ratio CPLS/VSL at VSL = \$10M; halve each entry for VSL = \$20M. p is the median conditional risk $P(H = 1 \mid S = 1)$ from Table 4: the Glass survey estimate uses 11.7% IPV prevalence (higher prevalence $\Rightarrow$ lower conditional risk) and the NISVS estimate uses 4.5% (lower prevalence $\Rightarrow$ higher conditional risk). e is intervention effectiveness—the probability that a homicide is prevented among targeted individuals who would otherwise be killed. C is the incremental incarceration cost of a felony vs. misdemeanor charge, computed as $\Delta T \times \kappa$ where ΔT is incremental time served (11 months under parole-era release, 21 months under 85% truth-in-sentencing) and κ is cost per inmate-year (marginal \$17,922 per Wilson & Lemoine 2022; median state average \$60,989 per USAFacts). Break-even $C^* = \text{VSL} \times p \times e$ is the maximum per-case cost at which benefits equal costs.

Figures

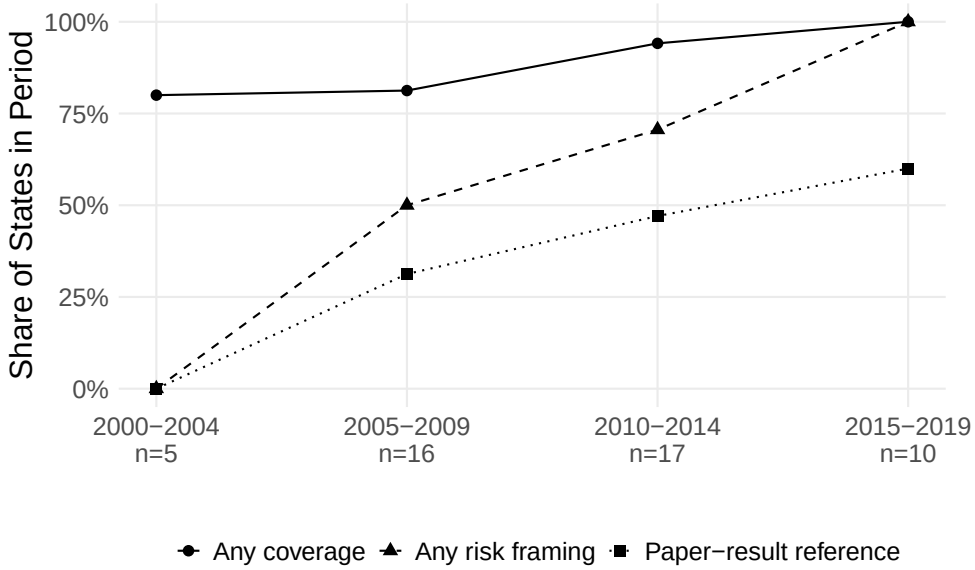


Figure 1: Newspaper coverage around NFS law adoption, by adoption-year bin. Lines show the share of adopting states with any located coverage of the law; the share whose coverage framed nonfatal strangulation as a precursor to future homicide (any risk framing); and the share whose coverage referenced the underlying homicide-risk research, by invoking a study or citing a numeric multiplier such as "seven times more likely" (a paper-result reference). Paper-result references rose from 0% of states adopting in 2000–2004 to 60% of those adopting in 2015–2019. States and supporting evidence are listed in Appendix Table 8.

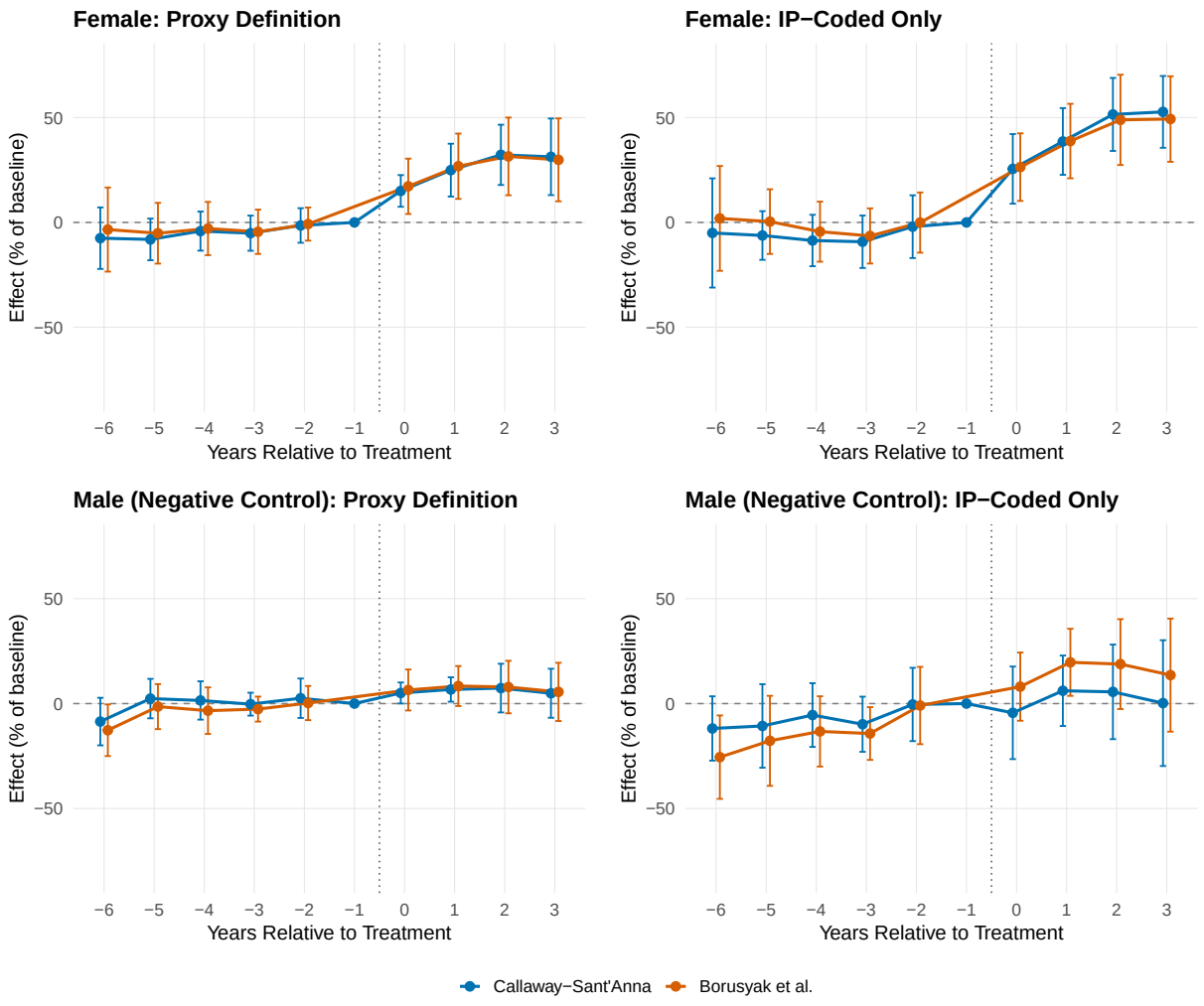


Figure 2: Event-study estimates of NFS law effects on assault classifications (NIBRS). The top panel shows estimates for female victims; the bottom panel shows male victims as a negative control. Columns compare the Callaway–Sant'Anna and imputation estimators. Pre-treatment coefficients for female victims are stable near zero; post-treatment estimates are positive. For the male negative control, the imputation estimator's pre-treatment coefficients trend negative and its post-treatment estimates rise correspondingly, so its larger point estimate reflects a pre-existing differential trend rather than a treatment effect; the Callaway–Sant'Anna estimates stay near zero throughout.

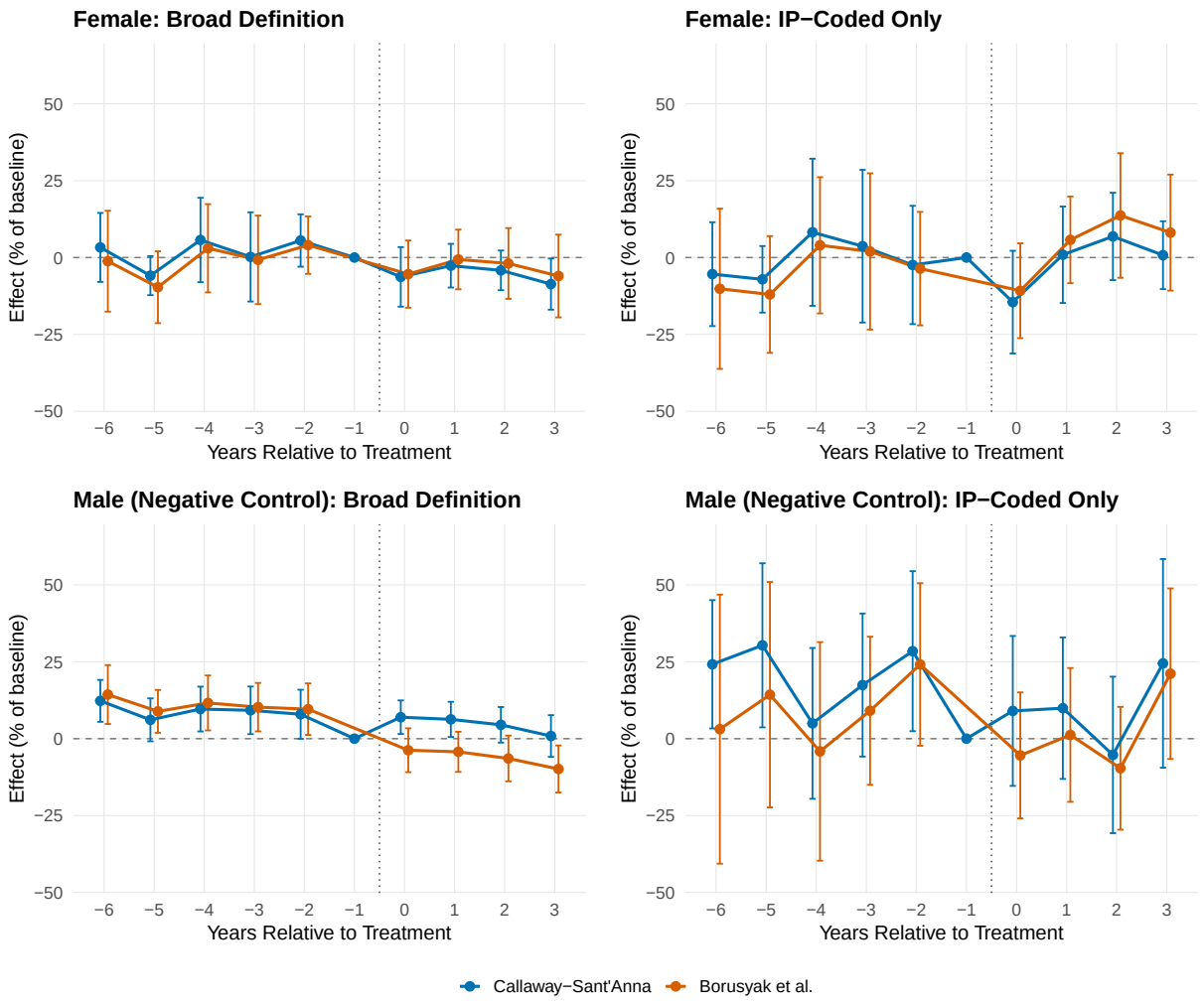


Figure 3: Event-study estimates of NFS law effects on homicide rates (SHR). The top panel shows estimates for female victims; the bottom panel shows male victims as a negative control. Columns compare the Callaway–Sant'Anna and imputation estimators. Pre-treatment coefficients for female outcomes are noisy but centered near zero. For the male negative control, the imputation estimator's pre-treatment coefficients are less stable than Callaway–Sant'Anna's on these sparse counts, but neither estimator yields a statistically significant post-treatment estimate.

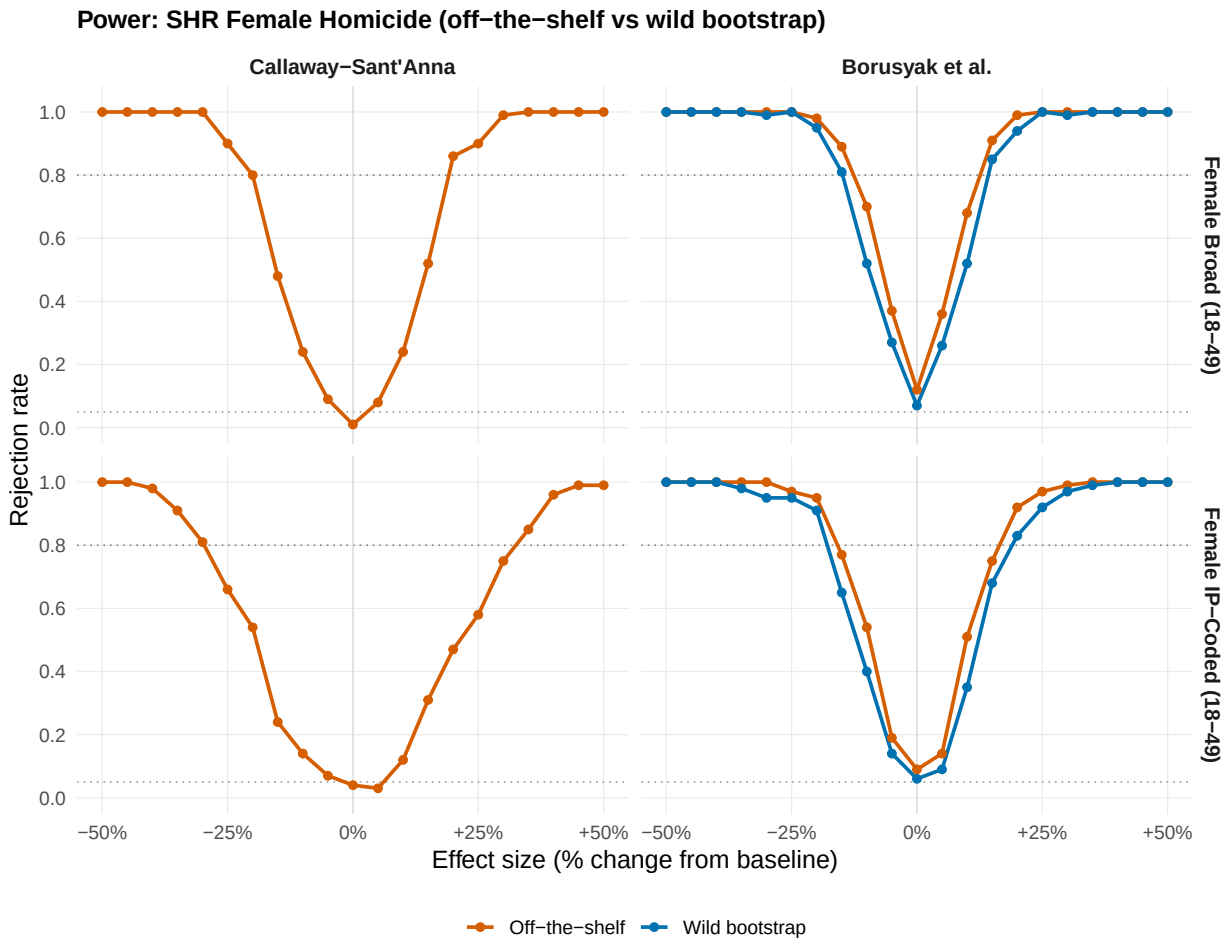


Figure 4: Power curves for SHR female intimate partner homicide. Each panel plots the rejection rate at the 5% level against the imposed effect size (percent change from baseline). Rows show the broad and IP-coded definitions; columns show the two estimators, Callaway–Sant'Anna and the Borusyak et al. imputation estimator. In the imputation panels, the two curves compare inference procedures: off-the-shelf cluster-robust inference and the restricted wild cluster bootstrap. The Callaway–Sant'Anna panels show the off-the-shelf curve alone: that test is already correctly sized on the homicide outcomes (Appendix Table 1), so no bootstrap correction is applied. Dotted horizontal lines mark 80% power and the 5% nominal size. Callaway–Sant'Anna reaches 80% power against reductions of about 20–30%; the imputation estimator's off-the-shelf curve over-rejects, so its wild-bootstrap curve is the relevant one, giving about 20%.

Appendix

A Formal Setup for the Simulation-Based Power Analysis

This appendix gives a formal, equation-based description of the simulation procedure behind the power analysis in Section 6. The goal is to build simulated datasets that (i) preserve the real staggered timing of NFS law adoption, (ii) impose a law effect of known size, and (iii) add year-to-year noise resembling the volatility of the observed data. We then re-run the paper's estimators on those datasets and record how often they detect the imposed effect, under both off-the-shelf cluster-robust inference and a wild-bootstrap correction for the few-cluster setting.

The approach builds on the simulated-power framework of Black et al. (2022), extending it to the staggered, multi-cohort setting and to a head-to-head comparison of estimators. It has five features that help interpret the null homicide result:

1. *Realistic calibration.* The simulation runs on the study's own panel, adoption timing, and estimators, so its detection rates reflect this design rather than a textbook formula.

2. *A concrete detection threshold.* It reports how large an effect must be before the design flags it reliably, so the observed near-zero estimates can be judged against that threshold rather than in the abstract.

3. *A check across identifying assumptions.* Because the simulated data can be built to satisfy either estimator's assumptions, the results show whether a conclusion holds under both or hinges on one.

4. *A check on the estimators themselves.* Building the data to suit one estimator and then applying the other reveals when an estimator misfires because its assumptions are violated—diagnosing the estimators, not just the effect size.

5. *Correctly-sized inference.* With few treated clusters, off-the-shelf cluster-robust standard errors can reject a true null too often; the analysis reports detection under a restricted wild cluster bootstrap alongside the off-the-shelf test, so the reported thresholds reflect inference that is correctly sized rather than over-rejecting.

A.1 Notation

Index states by $i \in \{1, \ldots, N\}$ and years by $t \in \{1, \ldots, T\}$. Let $G_i \in \{1, \ldots, T\} \cup \{\infty\}$ denote the first year state i adopts an NFS felony law, with $G_i = \infty$ for never-treated states. Define

$$D_{it} = \mathbb{1}\{t \geq G_i\},$$

which equals 1 for post-adoption years and 0 otherwise. Let Y_{it} denote the observed outcome (e.g., SHR homicide rate per 100,000 women; NIBRS aggravated assault share).

The simulation constructs a "no-law" outcome path for treated state-years, adds an assumed law effect of known magnitude, and then adds stochastic year-to-year fluctuation. We denote the resulting simulated outcome—the realized value the estimators are given, playing the role of the observed data—by $Y_{it}^{\text{sim}}(m)$, where m is the effect multiplier defined below.

A.2 Step 1: Constructing "No-Law" Counterfactual Outcomes

A core step is constructing what each treated state would have looked like in the absence of the law. We do this in two ways, corresponding directly to the two estimators emphasized in the paper.

Counterfactual 1: Callaway–Sant'Anna (CS) template. The CS estimator compares treated states to not-yet-treated states, measuring changes relative to the last pre-treatment year. To mirror this, we fit a smoothed set of no-law means on the untreated cells that lets each adoption cohort follow its own trend, then anchor each treated state to them.

Let S_{it} be a fitted no-law mean for each state-year, estimated by least squares on the untreated cells—never-treated states, not-yet-treated years, and treated states' pre-adoption years—allowing each state its own level a_i, a common year effect γ_t, and each cohort its own linear trend deviation:

$$S_{it} = a_i + \gamma_t + \beta_{g(i)}\,(t - \bar{t}). \tag{3}$$

The cohort slopes β_g let these means depart from a single common time path: they are *not* additively separable in state and year.

For a treated cell we set the no-law mean so the CS estimator recovers exactly zero at the

null—we take its fitted value S_{it} and subtract the group-time difference a CS fit reports on those means:

$$\mu^{(0)}_{it,\text{CS}} = S_{it} - \widehat{ATT}^{S}(G_i, t). \tag{4}$$

Equivalently, each treated cohort's post-adoption change is set to its not-yet-treated comparison group's change. Because every post-treatment group-time comparison is zero by construction, the CS parallel-trends restriction holds exactly here; because the fitted means are non-additive, imputation's two-way structure $\alpha_i + \gamma_t$ does not. This is the configuration in which the CS assumption holds and imputation's does not.

Counterfactual 2: Imputation template. The imputation estimator first fits an untreated-outcome model using only untreated observations, and then uses that model to predict what treated state-years would have looked like absent treatment. To mirror this logic, we estimate state and year effects using all observations with $D_{it} = 0$:

$$Y_{it} \approx \alpha_i + \gamma_t \qquad \text{using only } (i,t) \text{ such that } D_{it} = 0,$$

and denote the fitted effects by $\widehat{\alpha}_i$ and $\widehat{\gamma}_t$. For treated state-years ($D_{it} = 1$), the imputation-based "no-law" counterfactual mean is:

$$\mu^{(0)}_{it,\text{Imp}} = \widehat{\alpha}_i + \widehat{\gamma}_t. \tag{5}$$

Untreated state-years. The two counterfactuals above assign no-law means to treated cells only; the simulation must also assign means to the untreated cells—never-treated states, not-yet-treated years, and treated states' own pre-adoption years—since every state-year is redrawn. For the test to be fair these must follow the same no-law counterfactual the estimator assumes: an estimator applied to no-effect data must return an effect near zero, or its rejection rate at the null is not a meaningful size.

Under the imputation template the untreated mean is the same fit as in (5), $\mu^{(0)}_{it} = \widehat{\alpha}_i + \widehat{\gamma}_t$, so treated and untreated cells share one formula.

Under the CS template the untreated cells take the fitted no-law means S_{it} of (3) directly—without the $\widehat{ATT}^{S}$ subtraction, which applies only to treated cells.

A.3 Step 2: Imposing a Known Law Effect

We evaluate a grid of hypothetical effect sizes, parameterized as multiplicative changes to the post-adoption outcome path. Let

$$m \in \{0.50, 0.55, \ldots, 1.45, 1.50\}$$

denote the effect multiplier, so that $m = 0.80$ corresponds to a 20% reduction and $m = 1.20$ corresponds to a 20% increase.

For a chosen counterfactual template (CS or Imputation), define $\mu_{it}^{(0)}$ as the corresponding "no-law" mean from (4) or (5). We then impose the law effect on treated state-years by scaling the "no-law" mean:

$$\mu_{it}^{(1)}(m) = m \cdot \mu_{it}^{(0)} \qquad \text{for } D_{it} = 1. \tag{6}$$

For untreated state-years ($D_{it} = 0$), no effect is imposed, so $\mu_{it}^{(1)}(m) = \mu_{it}^{(0)}$, the fitted no-law mean from Step 1.

A.4 Step 3: Generating Stochastic Observed Outcomes

Given the effect-adjusted mean $\mu_{it}^{(1)}(m)$ from Steps 1–2, each simulated outcome is drawn from a count model—Poisson for rate outcomes, Binomial for share outcomes—and then expressed as a rate or share. A pure Poisson or Binomial draw fixes the variance once the mean is set, but the observed series fluctuate more than that. The model therefore adds two features to the pure count draw, each calibrated to the untreated series (below): extra year-to-year variance (overdispersion), and persistence, so a year's deviation carries into the next.

The latent process. Persistence is carried by a standardized latent process η_{it}, shared by both outcome types, in first-order autoregressive form—the simplest structure with a single persistence parameter ρ:

$$\eta_{it} = \rho\, \eta_{i,t-1} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathcal{N}\big(0,\, 1 - \rho^2\big), \tag{7}$$

so η_{it} has unit variance and first-order autocorrelation ρ; the AR(1) recursion carries a year's deviation into the next. Each outcome type turns this latent process into overdispersion of its own

magnitude, calibrated below.

Rate outcomes (Poisson–lognormal). For rate outcomes with population pop_{it} and scaling constant $c = 100{,}000$, the latent process enters as a multiplicative lognormal shock of log-standard-deviation σ:

$$\lambda_{it}(m) = \max\big(0,\ \mu_{it}^{(1)}(m)\big)\ \frac{\text{pop}_{it}}{c}\ \exp\big(\sigma\eta_{it} - \tfrac{1}{2}\sigma^2\big), \qquad N_{it}^{\text{sim}}(m) \sim \text{Poisson}\big(\lambda_{it}(m)\big), \tag{8}$$

with realized rate $Y_{it}^{\text{sim}}(m) = N_{it}^{\text{sim}}(m)\, c/\text{pop}_{it}$—the simulated counterpart of the rate the estimators see in the real data. The $-\frac{1}{2}\sigma^2$ term centers the shock at one, so it inflates variance without shifting the mean; at $\sigma = 0$ the draw is a pure Poisson. The floor keeps the Poisson intensity nonnegative and should bind only in essentially zero-mean cells; if it binds more than rarely, that becomes substantive, because the draw no longer preserves $\mu_{it}^{(1)}(m)$.[11]

Share outcomes (Beta–Binomial). For share outcomes with denominator n_{it}, the mean $\mu_{it}^{(1)}(m)$ is the success probability of a Beta–Binomial, the bounded and mean-preserving counterpart of the rate model. The per-cell probability is Beta-distributed with concentration ϕ,

$$p_{it}(m) \sim \text{Beta}\big(\mu_{it}^{(1)}(m)\,\phi,\ (1 - \mu_{it}^{(1)}(m))\,\phi\big), \qquad K_{it}^{\text{sim}}(m) \sim \text{Binomial}\big(n_{it},\ p_{it}(m)\big), \tag{9}$$

and $Y_{it}^{\text{sim}}(m) = K_{it}^{\text{sim}}(m)/n_{it}$. By construction $\mathbb{E}[p_{it}(m)] = \mu_{it}^{(1)}(m)$, so the draw preserves the mean, and $p_{it}(m) \in (0, 1)$ with no truncation; a smaller ϕ gives more overdispersion, and $\phi \to \infty$ recovers a pure Binomial. To carry the same persistence as the rate model, the Beta draw is coupled to the latent process through its distribution function, $p_{it}(m) = F_{\text{Beta}}^{-1}\big(\Phi(\eta_{it})\big)$, so serially correlated η_{it} yields serially correlated shares while leaving the Beta marginal—and hence the mean—unchanged.

Calibrating the noise. The persistence ρ is the pooled within-state first-order autocorrelation of the untreated residuals, from a two-way (state and year) fixed-effects decomposition of the

[11] In the panels used here the floor is essentially inert: across the two SHR rate outcomes and both counterfactual templates, the fitted no-law mean is negative in at most one of the 1,350 state-years per configuration, with a minimum fitted value of −0.12 per 100,000. Flooring these means at zero has no measurable effect on the simulated series.

untreated data. The overdispersion magnitude—σ for rate outcomes, ϕ for share outcomes—is chosen so the per-state residual standard deviation in simulated data matches the per-state residual standard deviation in the observed data.

Simulated datasets. Each replication draws a fresh outcome for every state-year from the model above. The Step-1 no-law means provide the center of each draw; Step 2 scales them on treated state-years, and Step 3 adds the calibrated noise.

A.5 Inference and the Few-Cluster Correction

We assess detection with the same inference the main analysis uses: for each estimator we take its cluster-robust standard error—clustered on state, the level at which the law is assigned and so the finest available—and reject the null of no effect when $|\widehat{\theta}/\widehat{SE}| > 1.96$. At effect multiplier m, power is the fraction of simulated datasets in which the null is rejected; at $m = 1$ the same fraction is the empirical size, which should sit near 5%.

The panels have few such clusters—47 states for the homicide outcomes (two never treated) and 29 for the enforcement outcomes—and cluster-robust inference can over-reject the null in this regime, a known small-cluster problem. Table 1 shows the pattern. On data built to satisfy its own additive structure, the imputation estimator's off-the-shelf test rejects the true null in 8–24% of samples, against a nominal 5%. The Callaway–Sant'Anna test stays at or below 8% across both data-generating processes, with the exception of the broad enforcement share, where it reaches 11%, and needs no correction.

We correct the reference distribution with the restricted wild cluster bootstrap (Cameron et al., 2008). Fit the null model $Y_{it} = \alpha_i + \gamma_t$ on the untreated cells, predict every cell to obtain fitted values $\widehat{Y}_{it}$ and residuals $\widehat{u}_{it} = Y_{it} - \widehat{Y}_{it}$. For each of B replications, assign each state a weight of +1 or −1 with probability one-half each—a Rademacher draw w_i—and form

$$Y^*_{it} = \widehat{Y}_{it} + w_i\,\widehat{u}_{it}, \tag{10}$$

which flips the sign of a state's entire residual block, preserving within-state dependence. We re-estimate the effect on Y^* and form the ratio $t^* = \widehat{\theta}^*/\widehat{SE}^*$; the critical value is the 95th percentile

of $|t^*|$ across replications—the folded statistic $|t^*|$ already combines both tails, so a two-sided 5% test takes the 0.95 quantile—and we reject when $|t_{\text{obs}}|$ exceeds it. The point estimate and standard error are unchanged; only the threshold they are compared against is corrected.

The wild bootstrap re-randomizes residual signs within the observed panel—it does not resample states or permute treatment. For the imputation estimator, the power curves report rejection under both the off-the-shelf standard error and the wild-bootstrap critical value; under the correction the $m = 1$ rejection rate returns to near 5% while power at nonzero effects is little changed.

A.6 Power and Null Size

Reading the homicide curves at 80% power gives the minimum detectable effect. The doubly-robust Callaway–Sant'Anna estimator is already correctly sized off the shelf on the homicide outcomes—so no bootstrap correction is applied and its power is read directly from the off-the-shelf curve—and reaches 80% power against SHR reductions of about 20–30%. The imputation estimator's off-the-shelf test over-rejects in this few-cluster setting, so its wild-bootstrap curve is the relevant one; read there, it reaches 80% power against reductions of about 20%, the figure we carry to the main text.

The average estimate at $m = 1$ is a separate diagnostic: a systematic departure from zero measures bias, flagging a mismatch between the counterfactual template used to impose the null and the estimator's own identifying assumption (Figures 8 and 9).

B Power Analysis Diagnostics

This appendix presents diagnostic outputs from the simulation-based power analysis described in Section 6 and formalized in Appendix A. Where informative, we report the full 2×2 grid of data-generating process (DGP)–estimator configurations, including mismatched ones, to quantify the consequences of parallel-trends mismatch.

B.1 Power Curves and Bias Heatmaps

Figure 7 shows power curves for the female NIBRS aggravated assault share, and Figures 8 and 9 present bias heatmaps showing systematic bias under the null across DGP–estimator configurations.

B.2 Null Calibration Diagnostics

Table 1 reports rejection rates under a true null for each DGP–estimator configuration. These rates are why the analysis runs the wild cluster bootstrap at all. Inference clusters on state—the level at which the law is assigned—and with few clusters, cluster-robust standard errors can understate sampling variability, so a nominal 5% test rejects a true null too often (Cameron et al., 2008). The table shows where this bites: the imputation estimator's off-the-shelf test rejects a true null in 8–24% of samples even on data generated to satisfy its own additive structure. The restricted wild cluster bootstrap (Appendix A.5) corrects the reference distribution rather than the estimate, roughly halving the over-rejection in every cell (to 4–11%) and returning most outcomes to near the nominal rate. The Callaway–Sant'Anna test is well sized off the shelf—at or below 8% under both data-generating processes, with the exception of the broad enforcement share, where it reaches 11%—so no correction is applied to it. This asymmetry is why the paper reads imputation-based power and minimum detectable effects from the wild-bootstrap curves while reading Callaway–Sant'Anna off the shelf.

C Absolute-Risk Calculation

This appendix derives the annual probability of intimate partner homicide for a woman with a prior nonfatal strangulation—the per-case benefit used in the cost-benefit analysis. Let $S = 1$ denote prior nonfatal strangulation and $H = 1$ intimate partner homicide. By Bayes' theorem,

$$P(H = 1 \mid S = 1) \;=\; \frac{P(S = 1 \mid H = 1)\, P(H = 1)}{P(S = 1)}, \tag{11}$$

where $P(S = 1 \mid H = 1)$ comes from Glass et al. (2008) and $P(H = 1)$ is the baseline annual probability of intimate partner homicide among IPV-exposed women. Because $P(S = 1)$ is not directly observed, we obtain it by marginalization,

$$P(S = 1) \;=\; P(H = 1)\, P(S = 1 \mid H = 1) \;+\; [1 - P(H = 1)]\, P(S = 1 \mid H = 0), \tag{12}$$

with $P(S = 1 \mid H = 0)$ also from Glass et al. (2008).

Inputs. The conditional strangulation shares come from the Glass et al. (2008) case-control study. Across attempted and completed homicide cases combined, 175 of 391 victims (44.8%) had a prior nonfatal strangulation, so $P(S = 1 \mid H = 1) = 0.448$; among abused controls, 41 of 426 (9.6%) reported one, so $P(S = 1 \mid H = 0) = 0.096$. Cases were women killed or nearly killed by an intimate partner, and controls were women physically assaulted or threatened with a weapon by a partner in the preceding two years, so we interpret the odds ratio as applying to this IPV-exposed population.

Baseline rate. We estimate $P(H = 1)$ by dividing annual intimate partner homicides of women aged 18–49 in the SHR by an estimate of the IPV-exposed female population in that age range. We use two prevalence assumptions: the 11.7% two-year prevalence implied by the Glass et al. (2008) survey, applied to partnered women, and the 4.5% 12-month prevalence from NISVS, applied to all women (Smith et al., 2018). (In the Glass random-digit-dialing survey, 427 of 3,637 women reported physical assault or weapon threats within the past two years; because that survey conditions on being in an intimate relationship, we apply the prevalence to partnered women

aged 18–49, approximating the partnered share as two-thirds, roughly consistent with Census tabulations for this age range. Results are not sensitive to plausible adjustments to this share.) The NISVS measure uses a shorter exposure window, implying a smaller IPV-exposed population and a higher baseline risk; reporting both assumptions brackets the baseline, and we report the resulting absolute-risk estimates as a range.

Uncertainty. We quantify uncertainty by Monte Carlo simulation with 10,000 draws, incorporating sampling variability in the Glass case-control proportions, uncertainty in IPV prevalence, and year-to-year variation in homicide counts; the medians and 95% intervals appear in the main text, and Appendix Figure 6 plots the full simulated distributions.

D Replication of Assis et al.

de Assis et al. (2026) report that NFS laws reduced intimate partner homicide (IPH) among victims aged 18–49 by 14% for women and 27% for men. Their outcome is IP-coded: single-victim incidents in which the victim–offender relationship is an intimate partner—current or former spouse, boyfriend or girlfriend, or common-law partner—with same-sex pairs excluded. They estimate it with the Gardner two-stage difference-in-differences (2SDID) estimator on homicide rates weighted by fixed 2000 state population. We re-estimate their specification, constructing rates and weights from SEER population estimates (which derive from the Census population series Assis et al. use and agree with it to within a fraction of a percent at the state, age, and sex level), and vary three choices they hold fixed. First, the estimator: we use the Borusyak–Jaravel–Spiess imputation estimator, which de Assis et al. (2026) note yields point estimates numerically equivalent to those of the 2SDID estimator they use, and compare it against the doubly-robust Callaway–Sant'Anna estimator, which identifies effects under a weaker parallel-trends assumption. Second, the weighting: fixed 2000 state-population weights, as in Assis et al., versus unweighted. Third, the rule for missing SHR data: their 21 hand-listed state-years versus the OJJDP EZASHR non-availability screen. We estimate each on their intimate-partner outcome and on a broader count of all single-victim homicides in the same age and sex group; Appendix Tables 2 and 3 report the full grid.

On the intimate-partner outcome, the imputation estimator yields point estimates in the direction and range of theirs—−0.076 per 100,000, or −15.0% of the baseline rate, for male victims aged 18–49, against the −0.090 (−27%) they report, and −0.130 (−9.4%) for female victims against their −0.169 (−14%)—though neither is statistically significant in our replication (female $p = 0.401$). The doubly-robust estimator, applied to the same data, returns essentially zero on that outcome—−0.059 for female victims ($p = 0.46$) and −0.020 for male victims ($p = 0.75$)—and reaches significance in none of its sixteen specifications at ages 18–49, nor in any of the sixty-four across all age groups and both outcome definitions. Every statistically significant estimate in the replication is produced by the imputation estimator.

At ages 18–49, no female estimate reaches conventional significance in our replication. The strongest—significant only at the 10% level—appears on the broader count of all single-victim

homicides rather than the intimate-partner outcome, and only under imputation, with population weights, and under their missing-data rule (−0.514 per 100,000, or −16.2% of baseline, $p = 0.100$). Dropping either of the remaining two choices moves the estimate further from significance: it falls to −0.190 ($p = 0.36$) unweighted, and to −0.405 ($p = 0.172$) when the OJJDP reporting-coverage screen replaces their 21-cell missing-data list.

Event studies indicate what the imputation estimator is fitting. It constructs the counterfactual by estimating a model on untreated observations and projecting it forward, so a pre-treatment trend carries into the post-treatment counterfactual. In the specifications with the largest estimates, the pre-treatment coefficients are positive and continue without a break into a negative post-period (Figure 5). The pattern is clearest for male victims on the broader outcome, where all five pre-treatment coefficients are positive and significant ($p < 0.01$) and decline toward the reference period, before a large negative post-period (−29.1%); under the doubly-robust estimator the same series shows no post-treatment effect. On the strongest female specification, the pre-treatment coefficients are again uniformly positive, though not individually significant, and give way to a negative post-period; the doubly-robust estimator applied to the same cell shows no significant pre-treatment divergence and an estimate short of conventional significance ($p = 0.08$). Assis et al.'s balance table records a significant pre-treatment difference in male-victim IPH between eventually treated and never-treated states at these ages (their Table A11), the subgroup in which the imputation effect is largest.

Population weighting concentrates the estimate on large states, and Assis et al.'s sample retains multiply-imputed values for two treated ones. Florida (adopted 2007) and Hawaii (adopted 2006) do not report to the SHR—both are absent from the raw incident files throughout the panel—so their values in the Fox–Swatt file are imputed rather than observed. Multiple imputation is a standard way to complete the SHR, and we do not fault its use in principle; for a treated state, though, it means the post-adoption outcomes enter the estimate as modeled rather than measured values. Assis et al.'s missing-data list drops only the 1990 Florida cell, leaving Florida's imputed 1991–2019 values and all thirty of Hawaii's in place. Our replication panel is built from the raw SHR and contains no imputed values—Florida and Hawaii enter none of our specifications—so the gap between our estimates and theirs partly reflects the imputed units, one of them (Florida) among the most heavily weighted states in the sample.

In sum, we reproduce the direction and approximate magnitude of Assis et al.'s estimates on the intimate-partner outcome, but statistical significance in the replication rests entirely on the imputation estimator: the doubly-robust alternative, which relies on a weaker identifying assumption, finds no significant effect at any age, for either sex, under any weighting or missing-data rule. The significant estimates coincide with positive pre-treatment trends that the imputation estimator projects forward and that the doubly-robust estimator does not, a pattern consistent with the pre-treatment differences Assis et al. report for male victims. The magnitudes depend further on population weighting, and Assis et al.'s own estimates additionally rest on imputed values for two treated non-reporting states (Florida and Hawaii) that never enter our panel.

E Additional Robustness and Coding Details

Alternative NIBRS sample constructions. Two features of the NIBRS panel could in principle shape the enforcement estimates. First, in four states—Illinois, Louisiana, Nebraska, and Vermont—exactly one agency satisfies the reporting filters, so those state-level series reflect a single agency's records. Second, the agency filter requires complete and uninterrupted reporting only over each agency's own span: agencies may enter the panel after 1991 or exit before 2019, subject to the pre/post coverage requirement, so the set of agencies underlying a state's series can change near the beginning and end of its panel. Table 6 re-estimates the headline enforcement models under three alternative samples: excluding the four single-agency states; restricting to the agencies whose reporting covers their state's entire span, which holds within-state agency composition fixed; and excluding Ohio, whose 2023 adoption makes it a within-window control state (Section 9). Across all alternatives the estimates are similar to or larger than the baseline, and the IP-coded estimate remains statistically significant throughout.

Coding procedure. For each state adopting a felony NFS statute through 2019, we searched the newspapers.com and LexisNexis news archives for coverage of the law, using the search `"domestic violence" AND "strangulation" AND (law OR bill)`, restricted to a window from one year before to one year after legislative passage (a 2015 law, for example, is searched over 2014–2016); where neither archive produced a usable article, we supplemented with a general web search. Coverage outside that window—reports on earlier failed bills or later case coverage—was excluded. Each located article was read in full from the page image and classified under criteria fixed before coding: an article counts as risk framing only if it frames strangulation as a predictor of the victim's future homicide; statements about the lethality of the act itself, the share of homicides involving strangulation, prevalence figures, and purely procedural coverage do not qualify. Risk-framing articles were classified as paper-result references if they named or invoked the underlying research or reported a numeric risk multiplier (e.g., "seven times more likely"), and as generic otherwise. Where several articles qualified for a state, we cite the one carrying the risk statistic.

Controlled specifications. We also re-estimate the main models adding time-varying controls for the state unemployment rate and overall crime rate. In this thin panel these specifications are

not identified well enough to interpret: adding covariates inflates the pre-treatment event-study coefficients' standard errors several-fold, and in the sparse homicide data these standard errors come to exceed the treatment estimate itself, so parallel trends can no longer be tested. Table 11 reports the estimates and Figures 3–4 the corresponding event studies; we rely on the no-controls specifications throughout.

Appendix Tables

Appendix Table 1: Null Calibration: Rejection Rates Under the True Null

		CS	Imputation	
Outcome	DGP	Off-shelf	Off-shelf	WCR
SHR homicide, broad	CS-valid	1%	4%	—
	Imp-valid	8%	12%	7%
SHR homicide, IP-coded	CS-valid	4%	0%	—
	Imp-valid	6%	8%	4%
NIBRS agg. share, broad	CS-valid	9%	4%	—
	Imp-valid	11%	24%	11%
NIBRS agg. share, IP-coded	CS-valid	5%	0%	—
	Imp-valid	7%	10%	4%

Notes: Rejection rate at $\alpha = 0.05$ (two-sided) under the true null (no effect), 100 simulations per cell. The DGP is fit to the untreated states: "CS-valid" generates data satisfying the Callaway–Sant'Anna post-period parallel-trends restriction exactly while being non-additive (so the imputation estimator's two-way structure is violated); "Imp-valid" generates additive two-way data. "Off-shelf" uses $|t| > 1.96$; "WCR" is the restricted wild cluster bootstrap. WCR is reported for imputation on the additive DGP, the cell where the off-the-shelf cluster-robust SE over-rejects with few treated clusters; the wild bootstrap restores nominal size. The CS estimator is well-sized off-the-shelf under both DGPs and needs no correction, and imputation on the CS-valid DGP is conservative off-the-shelf (its additive misfit inflates the analytic SE). Nominal size is 5%. Null bias is reported in Figures 8 and 9. See Appendix A.

Appendix Table 2: Replication of Assis et al. (2026): Intimate Partner Homicide by Age Band (broad definition)

	Assis A2 drops		OJJDP drops	
Outcome	CS	Imputation	CS	Imputation
Female IPH 18–49	-0.272†	-0.514†	-0.243†	-0.405
	(0.153)	(0.312)	(0.143)	(0.296)
Male IPH 18–49	-0.521	-3.577***	-0.329	-3.224***
	(0.584)	(0.995)	(0.637)	(0.968)
Female IPH 18–34	-0.263	-0.520	-0.234	-0.409
	(0.171)	(0.367)	(0.171)	(0.351)
Male IPH 18–34	-0.663	-4.699**	-0.426	-4.293*
	(0.876)	(1.697)	(0.898)	(1.680)
Female IPH 35–49	-0.278	-0.496†	-0.250	-0.389
	(0.175)	(0.265)	(0.159)	(0.248)
Male IPH 35–49	-0.131	-1.996***	0.007	-1.708***
	(0.441)	(0.503)	(0.412)	(0.470)
Female IPH 50–70	0.073	-0.098	0.112	-0.042
	(0.121)	(0.128)	(0.098)	(0.124)
Male IPH 50–70	0.040	-0.688*	0.123	-0.505
	(0.248)	(0.346)	(0.225)	(0.335)

Notes: Replication of Assis et al. (2026). The broad definition counts all single-victim homicides in the age and sex group. All regressions are weighted by fixed 2000 state population (SEER), following Assis; see the appendix text for the unweighted estimates. Entries are the ATT in homicides per 100,000; standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. CS = Callaway and Sant'Anna (2021); Imputation = Borusyak, Jaravel and Spiess (2024).

Sample: Assis A2 drops: 1,426 state-years, 48 states, 46 treated. OJJDP drops: 1,380 state-years, 48 states, 46 treated.

Appendix Table 3: Replication of Assis et al. (2026): Intimate Partner Homicide by Age Band (IP-coded definition)

	Assis A2 drops		OJJDP drops	
Outcome	CS	Imputation	CS	Imputation
Female IPH 18–49	-0.059	-0.130	-0.041	-0.078
	(0.080)	(0.155)	(0.075)	(0.143)
Male IPH 18–49	-0.020	-0.076	-0.019	-0.050
	(0.062)	(0.050)	(0.054)	(0.047)
Female IPH 18–34	-0.037	-0.072	-0.020	-0.022
	(0.092)	(0.196)	(0.098)	(0.193)
Male IPH 18–34	-0.049	-0.106**	-0.048	-0.084*
	(0.081)	(0.041)	(0.076)	(0.033)
Female IPH 35–49	-0.081	-0.195	-0.062	-0.141
	(0.109)	(0.139)	(0.090)	(0.111)
Male IPH 35–49	0.020	-0.041	0.019	-0.009
	(0.066)	(0.098)	(0.066)	(0.099)
Female IPH 50–70	0.015	-0.030	0.031	-0.014
	(0.058)	(0.057)	(0.062)	(0.054)
Male IPH 50–70	-0.022	-0.048	-0.019	-0.031
	(0.036)	(0.042)	(0.038)	(0.036)

Notes: Replication of Assis et al. (2026). The IP-coded definition counts only single-victim homicides with an intimate partner relationship code, matching Assis's outcome. All regressions are weighted by fixed 2000 state population (SEER), following Assis; see the appendix text for the unweighted estimates. Entries are the ATT in homicides per 100,000; standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. CS = Callaway and Sant'Anna (2021); Imputation = Borusyak, Jaravel and Spiess (2024).

Sample: Assis A2 drops: 1,426 state-years, 48 states, 46 treated. OJJDP drops: 1,380 state-years, 48 states, 46 treated.

Appendix Table 4: Enforcement Effects of NFS Laws (Balanced Event-Study Sample)

	Female victims	
	CS	Imputation
Panel A: Proxy Definition (all female-victim assaults)		
Aggravated assault share	0.034*	0.038**
	(0.014)	(0.014)
Total assault rate (per 100k)	-22.6	27.0
	(94.1)	(171.8)
Aggravated assault rate (per 100k)	24.6	39.8
	(25.2)	(37.1)
Clearance rate	0.027	0.034
	(0.021)	(0.025)
Arrest rate	-0.008	-0.003
	(0.012)	(0.016)
Panel B: IP-Coded Definition (relationship-verified)		
Aggravated assault share	0.051***	0.051***
	(0.015)	(0.014)
Total assault rate (per 100k)	-10.3	16.4
	(50.6)	(93.1)
Aggravated assault rate (per 100k)	23.2†	28.7
	(13.6)	(17.6)
Clearance rate	0.015	0.005
	(0.014)	(0.018)
Arrest rate	-0.009	-0.019
	(0.015)	(0.014)

Notes: As Table 1, but estimated on the balanced event-study sample—states observed at every event time in $[-6, +3]$ (19 treated states across 9 adoption cohorts)—rather than the full unbalanced panel (30 states, 27 treated, 616 state-years). Standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The aggravated assault share, clearance rate, and arrest rate are proportions (0–1); total and aggravated assault rates are counts per 100,000. No time-varying controls included.

Appendix Table 5: Effects of NFS Laws on Intimate Partner Homicide (Balanced Event-Study Sample)

	Female victims	
	CS	Imputation
Panel A: Broad Definition		
ATT	-0.135	-0.055
	(0.180)	(0.285)
% of baseline	-4.0%	-1.6%
Panel B: IP-Coded Definition		
ATT	0.151	0.202
	(0.135)	(0.175)
% of baseline	10.3%	13.7%

Notes: As Table 2, but estimated on the balanced event-study sample—states observed at every event time in $[-6, +3]$ (36 treated states across 11 adoption cohorts)—rather than the full unbalanced panel (47 states, 45 treated, 1350 state-years). Standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ATT is in intimate partner homicides per 100,000. Percent of baseline is ATT divided by the control-group mean. No time-varying controls included.

Appendix Table 6: NIBRS Enforcement Effects Under Alternative Sample Constructions

	Broad proxy		IP-coded	
Sample	CS	Imputation	CS	Imputation
Baseline panel	0.021†	0.022†	0.037*	0.034*
(states = 29)	(0.012)	(0.013)	(0.015)	(0.014)
Excluding Ohio	0.028*	0.026*	0.042**	0.036*
(states = 28)	(0.012)	(0.013)	(0.015)	(0.015)
Excluding single-agency states	0.025*	0.024†	0.040**	0.036*
(states = 25)	(0.011)	(0.014)	(0.013)	(0.015)
Full-span agencies only	0.034*	0.027†	0.046**	0.038*
(states = 29)	(0.016)	(0.015)	(0.018)	(0.017)

Notes: ATT on the female aggravated-assault share (share points); standard errors, clustered by state, in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Excluding Ohio* removes the state whose 2023 adoption makes it a within-window control (see the Discussion). *Excluding single-agency states* removes Illinois, Louisiana, Nebraska, and Vermont, whose panels rest on one reporting agency each. *Full-span agencies only* rebuilds state-year outcomes from the agencies whose reporting covers their state's entire panel span, holding within-state agency composition fixed. Estimation is identical to Table 1.

Appendix Table 7: Year Passed for Specialized Felony NFS Statutes by State

State	Year Passed	State	Year Passed	State	Year Passed
Alabama	2011	Louisiana	2006	Ohio	2023‡
Alaska	2005	Maine	2011	Oklahoma	2004
Arizona	2010	Maryland	2019	Oregon	2018*
Arkansas	2009	Massachusetts	2014	Pennsylvania	2016
California	2011	Michigan	2012	Rhode Island	2012
Colorado	2016	Minnesota	2005	South Carolina	—†
Connecticut	2007	Mississippi	2010	South Dakota	2012
Delaware	2010	Missouri	2000	Tennessee	2011
Florida	2007	Montana	2017	Texas	2009
Georgia	2014	Nebraska	2004	Utah	2017
Hawaii	2006	Nevada	2009	Vermont	2005
Idaho	2005	New Hampshire	2010	Virginia	2012
Illinois	2009	New Jersey	2017	Washington	2007
Indiana	2006	New Mexico	2018	West Virginia	2016
Iowa	2012	New York	2010	Wisconsin	2007
Kansas	2017	North Carolina	2004	Wyoming	2011
Kentucky	2019	North Dakota	2007		

Notes: Entries report the legislative year in which a specialized felony nonfatal strangulation statute was passed. * Oregon is excluded from the primary treatment coding because it had an earlier strangulation statute before the later felony reclassification. † South Carolina had no specialized felony NFS statute during the study period. ‡ Ohio enacted its statute after the main estimation windows and is therefore untreated in the primary analysis.

Appendix Table 8: Newspaper and legislative coverage of each state's nonfatal-strangulation law: discovery source, risk-framing classification, and supporting evidence.

State	Year	Source	Framing	Evidence (source article)
Missouri	2000	newspapers.com	None	General domestic-violence coverage; no strangulation future-homicide framing. *(St. Louis Post-Dispatch 2001-11-25)*
Oregon	2003	newspapers.com	None	Act-lethality framing only ("maybe attempted murder"); no future-homicide risk. *(Statesman Journal (Salem) 2003-03-15)*
Nebraska	2004	newspapers.com	None	Act-lethality framing only ("attempted murder"); no future-homicide risk. *(Telegraph (North Platte) 2004-10-19)*
North Carolina	2004	newspapers.com	None	Act-danger and legal-gap framing; no future-homicide prediction. *(The News & Observer (Raleigh) 2003-05-20)*
Oklahoma	2004	—	None	(no article located)
Alaska	2005	newspapers.com	None	Procedural coverage of the bill's passage (felony classification, vote counts, evidentiary gap under prior law); no risk framing. *(Juneau Empire 2005-04-29 (p.5, AP))*
Idaho	2005	newspapers.com	None	Bill coverage plus a general DV-murder statistic; no strangulation future-risk framing. *(AP State Wire (Boise) 2005-02-07)*
Minnesota	2005	newspapers.com	Generic	After being strangled, "the next attack is oftentimes fatal." *(St. Cloud Times 2005-08-24)*
Vermont	2005	—	None	(no article located)
Hawaii	2006	—	None	(no article located)
Indiana	2006	newspapers.com	Paper-ref.	"If a woman is choked once, she is 10 times more likely to become a victim of homicide." *(Chronicle-Tribune (Marion) 2006-10-28)*
Louisiana	2006	—	None	(no article located)
Connecticut	2007	newspapers.com	None	Procedural coverage ('Capitol Briefs': 'The bill also creates a new domestic violence crime of strangulation', protective-order and stun-gun provisions); no risk framing. *(Hartford Courant 2007-05-30 (B11, Capitol Briefs))*
Florida	2007	newspapers.com	Paper-ref.	"Victims strangled were 10 times more likely to be killed by their abuser" (Dia Kuykendall, Florida Coalition Against Domestic Violence). *(News-Press (Fort Myers) 2007-06-29)*
North Dakota	2007	newspapers.com	Generic	Strangulation is "a serious indicator that victims may die at the hands of perpetrators." *(The Bismarck Tribune 2007-01-16)*
Washington	2007	newspapers.com	Generic	Strangulation is "often a precursor to homicide by other means"; "the last warning shot" (Gael Strack). *(The News Tribune (Tacoma) 2007-03-26)*
Wisconsin	2007	newspapers.com	Paper-ref.	Victims of strangulation are "about 10 times more likely to be murdered" by their abuser. *(The Reporter (Fond du Lac) 2007-06-26)*
Arkansas	2009	newspapers.com	None	No future-homicide risk framing. *(Baxter Bulletin 2009-01-26)*
Illinois	2009	newspapers.com	Paper-ref.	Cites the Chicago Women's Health Risk Study and a 2003 Minnesota domestic-violence-murders study; "an act of strangulation was often a precursor to a domestic homicide" (Marcy Anderson, WATCH). *(Chicago Tribune 2009-03-10)*
Nevada	2009	newspapers.com	None	No future-homicide risk framing. *(Reno Gazette-Journal 2009-10-25)*
Texas	2009	newspapers.com	Paper-ref.	Nonfatal strangulation victims are "nine times more likely" to be killed by their abuser (National Coalition Against Domestic Violence). *(Fort Worth Star-Telegram 2010-08-01)*
Arizona	2010	newspapers.com	Paper-ref.	Strangulation victims are "10 times more likely to die from abuse" (U.S. Dept. of Justice). *(Arizona Daily Star 2011-01-02)*
Delaware	2010	newspapers.com	Generic	"It's a precursor to murder." *(The News Journal (Wilmington) 2010-02-01)*
Mississippi	2010	newspapers.com	Generic	Strangulation "is often the last thing before someone is killed." *(The Vicksburg Post 2010-09-20)*
New Hampshire	2010	newspapers.com	Generic	Strangulation is "one of the top indicators of a later lethal event." *(Concord Monitor 2010-03-29)*
New York	2010	newspapers.com	Paper-ref.	Strangulation victims are "10 times more likely to become homicide victims." *(Poughkeepsie Journal 2011-04-08)*
Alabama	2011	newspapers.com	None	Act-of-violence framing plus a CDC prevalence statistic; no future-homicide risk. *(The Anniston Star 2012-11-11)*
California	2011	newspapers.com	Paper-ref.	"When a woman has been strangled by her partner, she is 800 percent more likely to be killed" (Gael Strack, National Family Justice Center Alliance). *(Desert Dispatch (Barstow) 2011-07-27 (AP))*
Maine	2011	Nexis	Paper-ref.	Strangulation victims are "nine times more likely to be killed" (NIJ Journal). *(States News Service (Augusta) 2011-04-08)*

Appendix Table 8 (continued)

State	Year	Source	Framing	Evidence (source article)
Tennessee	2011	Nexis	Paper-ref.	"Victims of prior attempted strangulation are seven times more likely of becoming a victim of homicide" (Speaker Pro Tem Jamie Woodson); "one of the best predictors of future homicide." *(States News Service (Nashville) 2011-05-21)*
Wyoming	2011	Nexis	None	Procedural coverage only; no lethality or homicide framing. *(Wyoming Tribune-Eagle 2011-06-27)*
Iowa	2012	newspapers.com	Generic	The laws promote awareness of "a crime they say often precedes homicide." *(The Daily Nonpareil (Council Bluffs) 2012-05-18)*
Michigan	2012	—	None	(no article located)
Rhode Island	2012	Nexis	None	Method-share ("about 10 percent of female homicide deaths involve strangulation") and escalation language; no victim future-homicide figure. *(Targeted News Service / Rep. DaSilva 2012-04-04)*
South Dakota	2012	newspapers.com	None	Legislative/procedural coverage; no future-homicide framing. *(Rapid City Journal 2012-02-03)*
Virginia	2012	newspapers.com	Paper-ref.	Women strangled once have "about a 40 percent chance" of being killed later by their attacker (Melissa Harper, forensic nurse). *(The News & Advance (Lynchburg) 2012-01-28)*
Georgia	2014	newspapers.com	Paper-ref.	Strangulation victims are "eight times more likely to be killed" by the same abuser. *(The Atlanta Constitution 2014-09-05)*
Massachusetts	2014	newspapers.com	Paper-ref.	Strangulation is "a significant precursor to murder among domestic violence victims. One 2004 study ... found that 56 percent of those killed by their abusers had been strangled by them at some point in their relationship." *(The Patriot Ledger (Quincy) 2014-08-18 (pp. 1, 5))*
Colorado	2016	newspapers.com	Generic	Strangulation is a "major predictor of potential future homicidal behavior." *(The Daily Sentinel (Grand Junction) 2015-11-22)*
Pennsylvania	2016	newspapers.com	Paper-ref.	"A study by Johns Hopkins professor Nancy Glass found victims of non-fatal strangulation had a more than six-fold increase in odds of becoming a victim of attempted homicide." *(The Sentinel (Carlisle) 2017-10-21)*
West Virginia	2016	web search	Paper-ref.	"Someone who strangles someone is seven times more likely to kill them in the future" (Joyce Yadlosky, WVCADV, citing the Journal of Emergency Medicine). *(WV MetroNews 2016-06-13)*
Kansas	2017	newspapers.com	Paper-ref.	"A survivor of strangulation is more than seven times more likely to become a homicide victim" (Lucca Wang, KCSDV). *(The Topeka Capital-Journal 2018-04-10)*
Montana	2017	newspapers.com	Generic	"Strangulation is the last warning shot . . . something bad is going to happen" (Gael Strack). *(Great Falls Tribune 2017-11-23)*
New Jersey	2017	newspapers.com	Generic	"Nonfatal strangulation is a leading precursor to domestic homicide." *(Press of Atlantic City 2017-11-15)*
Utah	2017	Nexis	Paper-ref.	"They point to studies suggesting choking predicts future domestic violence that can end in homicide." *(Deseret Morning News 2017-03-02)*
New Mexico	2018	newspapers.com	Paper-ref.	"The mere presence of strangulation in a situation of domestic abuse increases the chances of homicide sevenfold" (Gael Strack, Training Institute on Strangulation Prevention). *(The Santa Fe New Mexican 2018-02-16)*
Kentucky	2019	newspapers.com	Generic	Domestic-violence victims are "significantly more likely to eventually die at the hands of their abusers if their abuse has included an incident of strangulation" (Tina Ward-Pugh, Louisville Metro Office for Women). *(The Courier-Journal (Louisville) 2019-03-24 (E2))*
Maryland	2019	newspapers.com	Paper-ref.	"According to a 2008 study by the Training Institute on Strangulation Prevention, a victim of strangulation is seven times more likely to end up a homicide statistic" (Aisha N. Braveboy, Prince George's County State's Attorney). *(The Baltimore Sun 2020-01-08)*

Notes. States ordered by year of adoption. *Source*: how the coverage was located—`newspapers.com`, LexisNexis (`Nexis`), or a general web search. *Framing*: *Paper-ref.* = coverage cites the underlying homicide-risk research, either by naming a study or reporting a numeric multiplier (e.g. "seven times more likely"); *Generic* = frames strangulation as a precursor to future homicide without invoking research or a figure; *None* = coverage found but no future-homicide framing (act-lethality, prevalence, or procedural only). "—" under *Source* denotes no contemporaneous coverage located. The table covers the 48 states that adopted through 2019, the window of Figure 1.

Appendix Table 9: SHR State-Years Excluded Under OJJDP EZASHR Methods (1990–2019)

State	NFS Law	*N* Dropped	Years Excluded
Alabama	Yes (2011)	10	1999, 2011–2019
Delaware	Yes (2010)	2	1994–1995
District of Columbia	No	13	1996, 1998–2008, 2012
Iowa	Yes (2012)	1	1991
Kansas	Yes (2017)	8	1993–2000
Kentucky	Yes (2019)	6	1994, 1999–2003
Louisiana	Yes (2006)	1	1991
Maine	Yes (2011)	3	1991–1993
Montana	Yes (2017)	9	1990, 1993–1994, 1996–1999, 2001–2002
Nebraska	Yes (2004)	15	1993–2003, 2005–2008
New Hampshire	Yes (2010)	1	1997
North Dakota	Yes (2007)	2	1991, 2012
South Dakota	Yes (2012)	1	1997
Wisconsin	Yes (2007)	1	1998
Total		73	

Notes: State-years flagged as non-available under the OJJDP EZASHR methods are excluded from the SHR analysis. These represent years where SHR reporting coverage falls below reliability thresholds (Office of Juvenile Justice and Delinquency Prevention, 2026). Florida and Hawaii are absent from the SHR county-level data entirely and do not appear in any specification. Oregon is excluded by design (see text). The "NFS Law" column indicates whether the state adopted a non-fatal strangulation law during the analysis period, with year of adoption in parentheses. Notable exclusions include Alabama (all post-treatment years) and Nebraska (11 of 12 pre-treatment years). The final SHR panel contains 1350 state-year observations across 47 states.

Appendix Table 10: Sensitivity of the Homicide Null to the SHR Reporting-Coverage Screen

	Excludes flagged (main)		Includes all state-years	
	CS	Imputation	CS	Imputation
Panel A: Broad Definition				
ATT	-0.155	-0.056	-0.159	-0.260
	(0.165)	(0.179)	(0.169)	(0.233)
% of baseline	-4.6%	-1.7%	-4.9%	-8.1%
Panel B: IP-Coded Definition				
ATT	0.075	0.099	0.059	-0.004
	(0.113)	(0.106)	(0.115)	(0.127)
% of baseline	5.2%	6.8%	4.2%	-0.3%

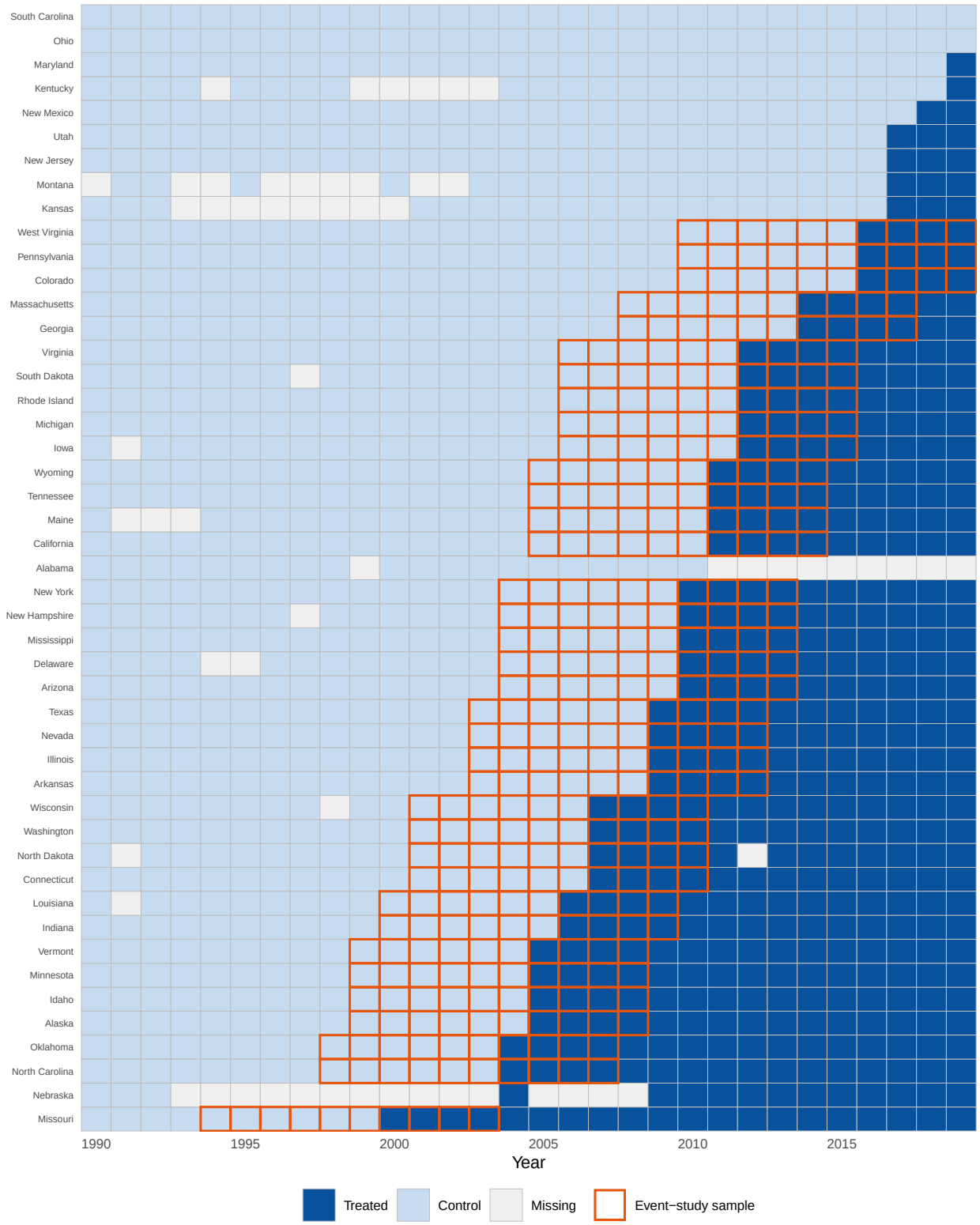
Notes: Female victims. The left columns exclude the state-years the OJJDP EZA-SHR reporting-coverage screen flags (the main-analysis panel; these columns reproduce Table 2); the right columns add those state-years back and estimate on all reported state-years. Standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ATT is in intimate partner homicides per 100,000; percent of baseline is ATT divided by the control-group mean. No time-varying controls.

Appendix Table 11: Controlled Specifications: Enforcement and Homicide Effects of NFS Laws

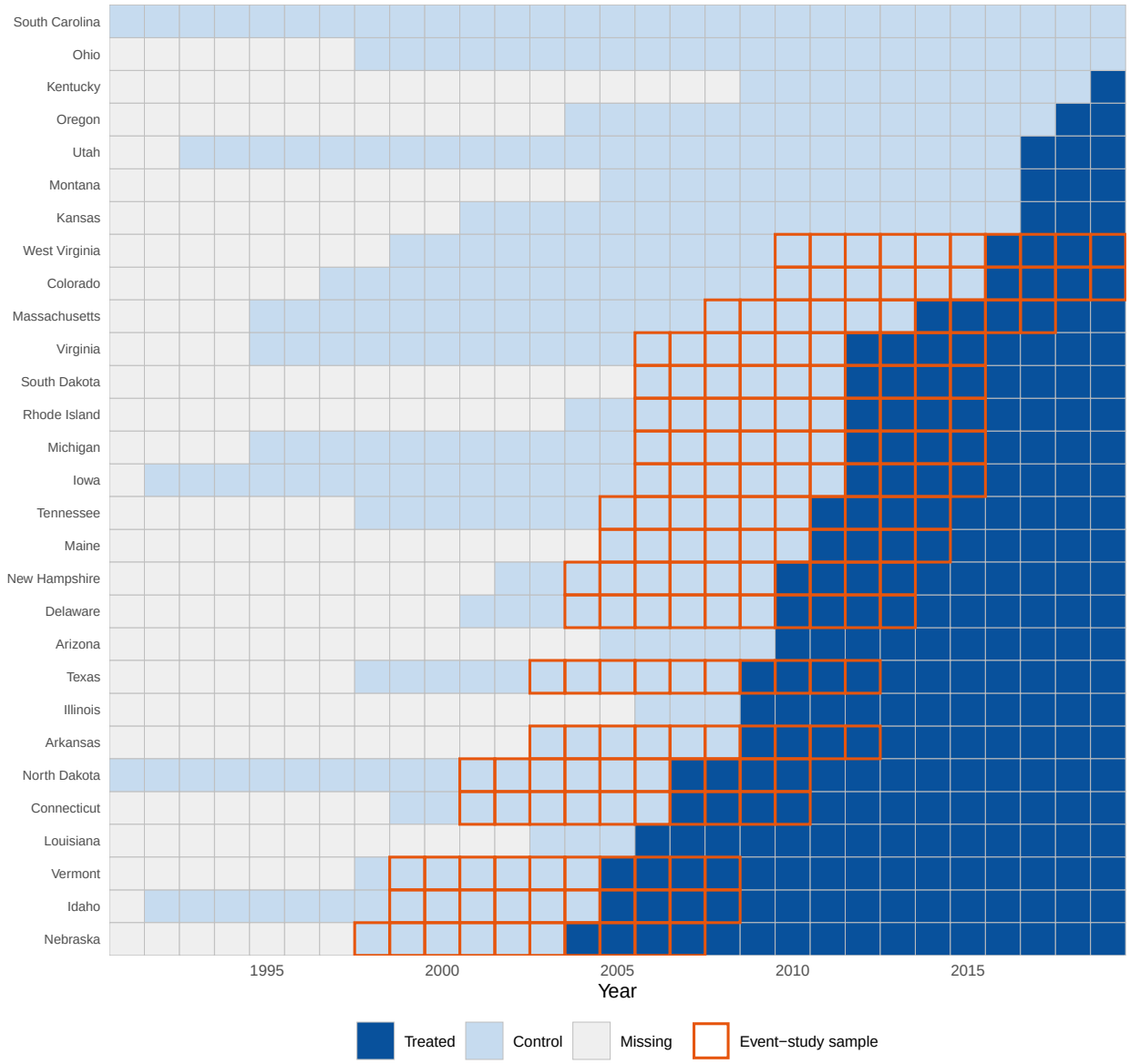
	NIBRS Aggravated Assault Share		SHR Intimate Partner Homicide	
	CS	Imputation	CS	Imputation
Panel A: Proxy/Broad Definition				
ATT	0.017	0.040***	-0.176	0.166†
	(0.025)	(0.011)	(0.221)	(0.090)
% of baseline	14.4%	32.8%	-5.3%	5.0%
Panel B: IP-Coded Definition				
ATT	0.040	0.055***	-0.134	0.142*
	(0.032)	(0.011)	(0.152)	(0.061)
% of baseline	37.4%	51.0%	-9.2%	9.8%

Notes: Standard errors clustered by state in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. NIBRS ATT is in share points (the aggravated share is a proportion, 0–1). SHR ATT is in homicides per 100,000. Percent of baseline is ATT divided by the control-group mean. Controls include state unemployment rate and overall crime rate.

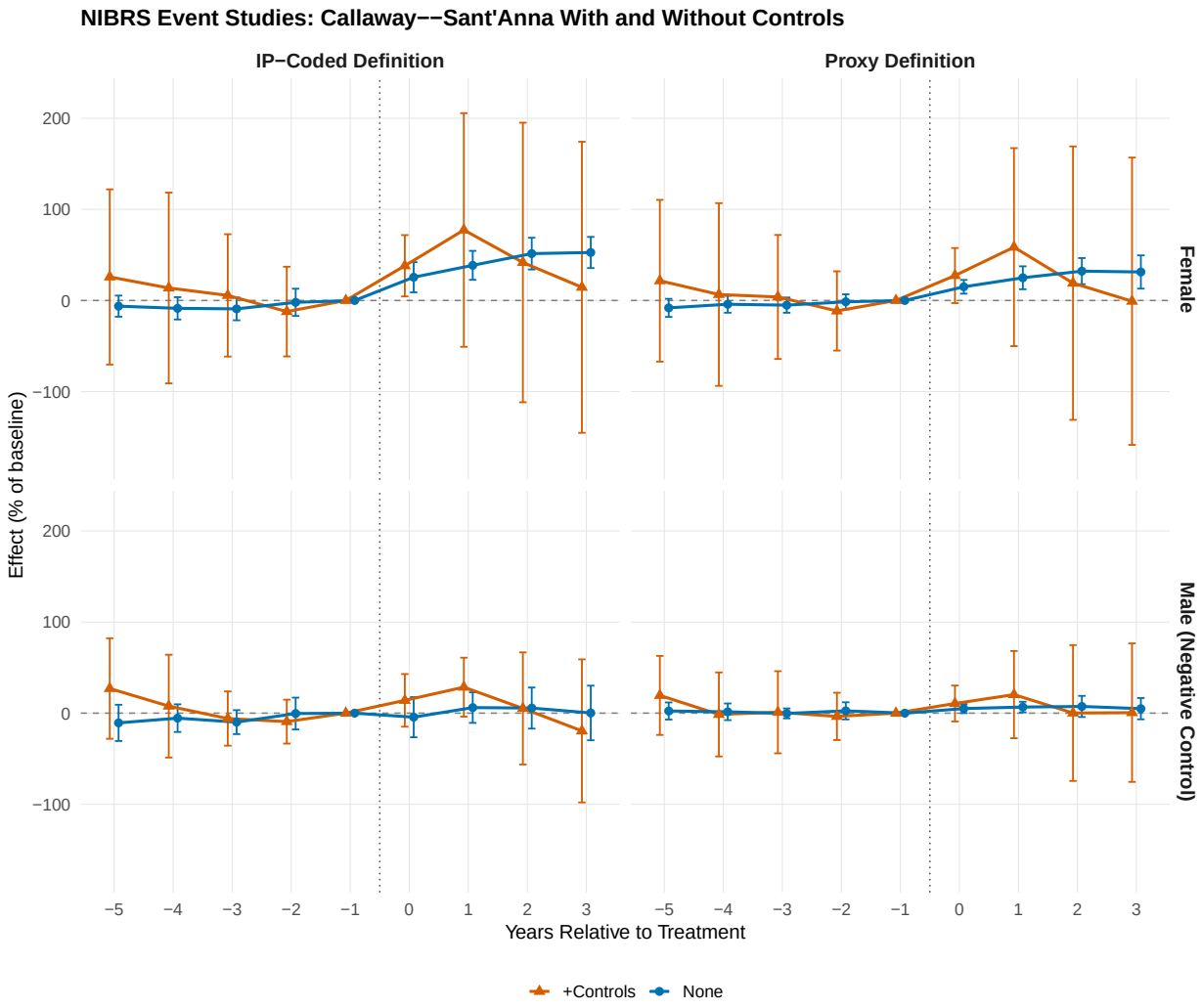
Appendix Figures



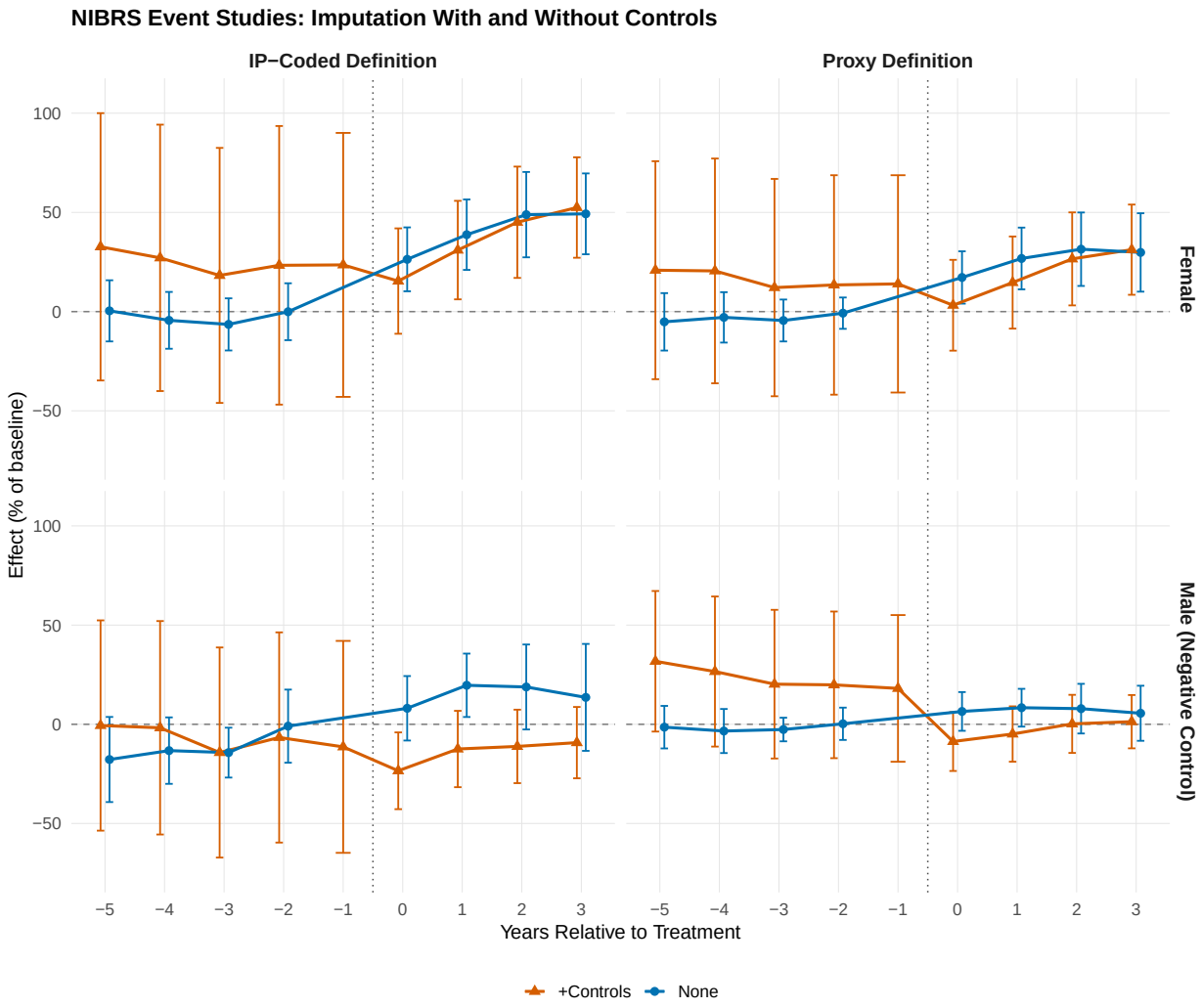
Appendix Figure 1: Staggered adoption of felony nonfatal strangulation statutes across the SHR analysis panel. Each row is a state; fill shows whether a state-year is not-yet-treated (control), treated, or excluded by the OJJDP reporting screen (missing). Cells outlined in orange form the balanced event-study sample—36 treated states across 11 adoption cohorts observed at every event time in $[-6, +3]$; all non-missing cells enter the ATT.



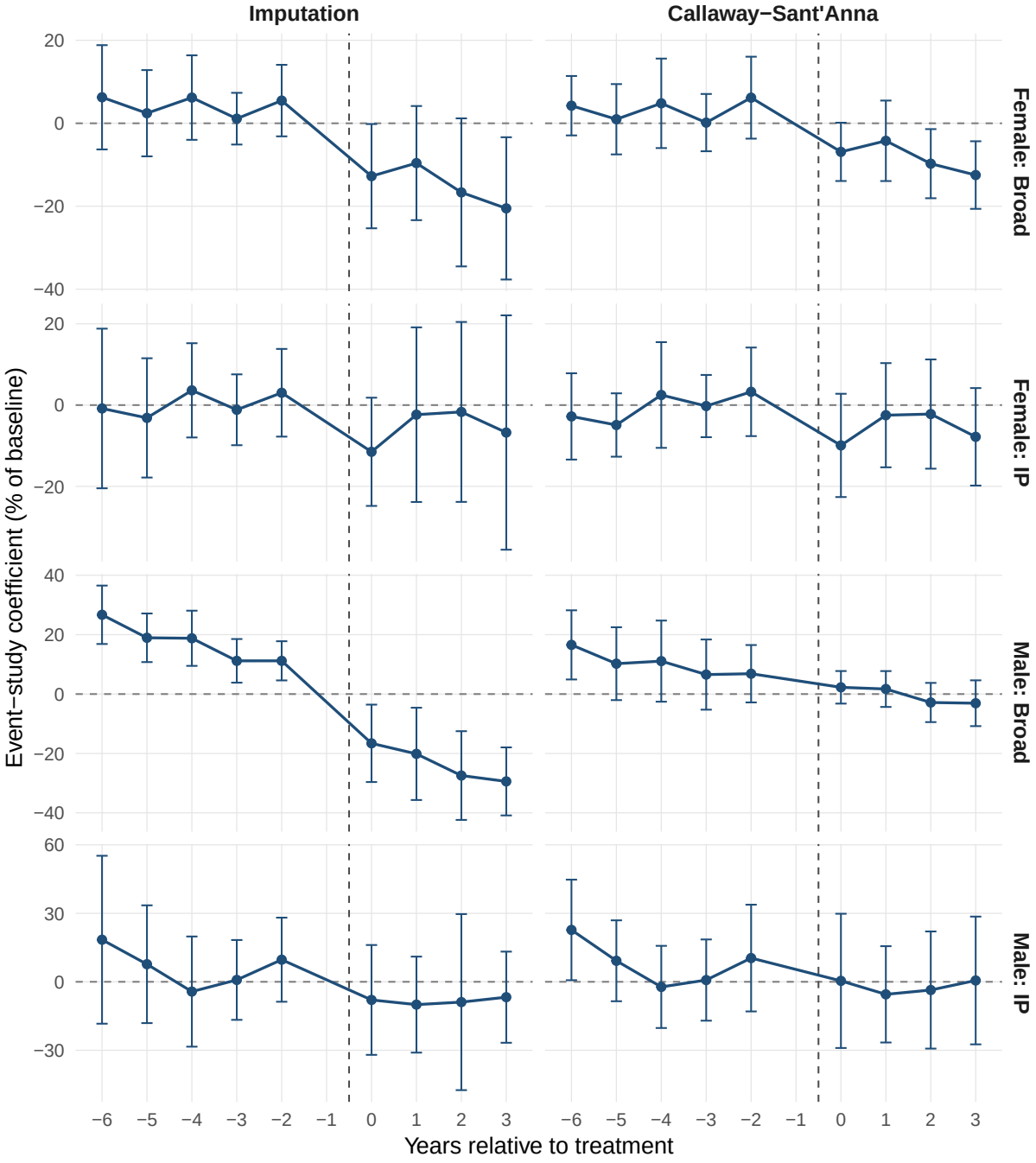
Appendix Figure 2: Staggered adoption across the NIBRS enforcement panel. Grey denotes state-years outside NIBRS coverage, which expands over the period as agencies transition to NIBRS reporting. Cells outlined in orange form the balanced event-study sample—19 treated states across 9 adoption cohorts observed at every event time in $[-6, +3]$; all non-missing cells enter the ATT.



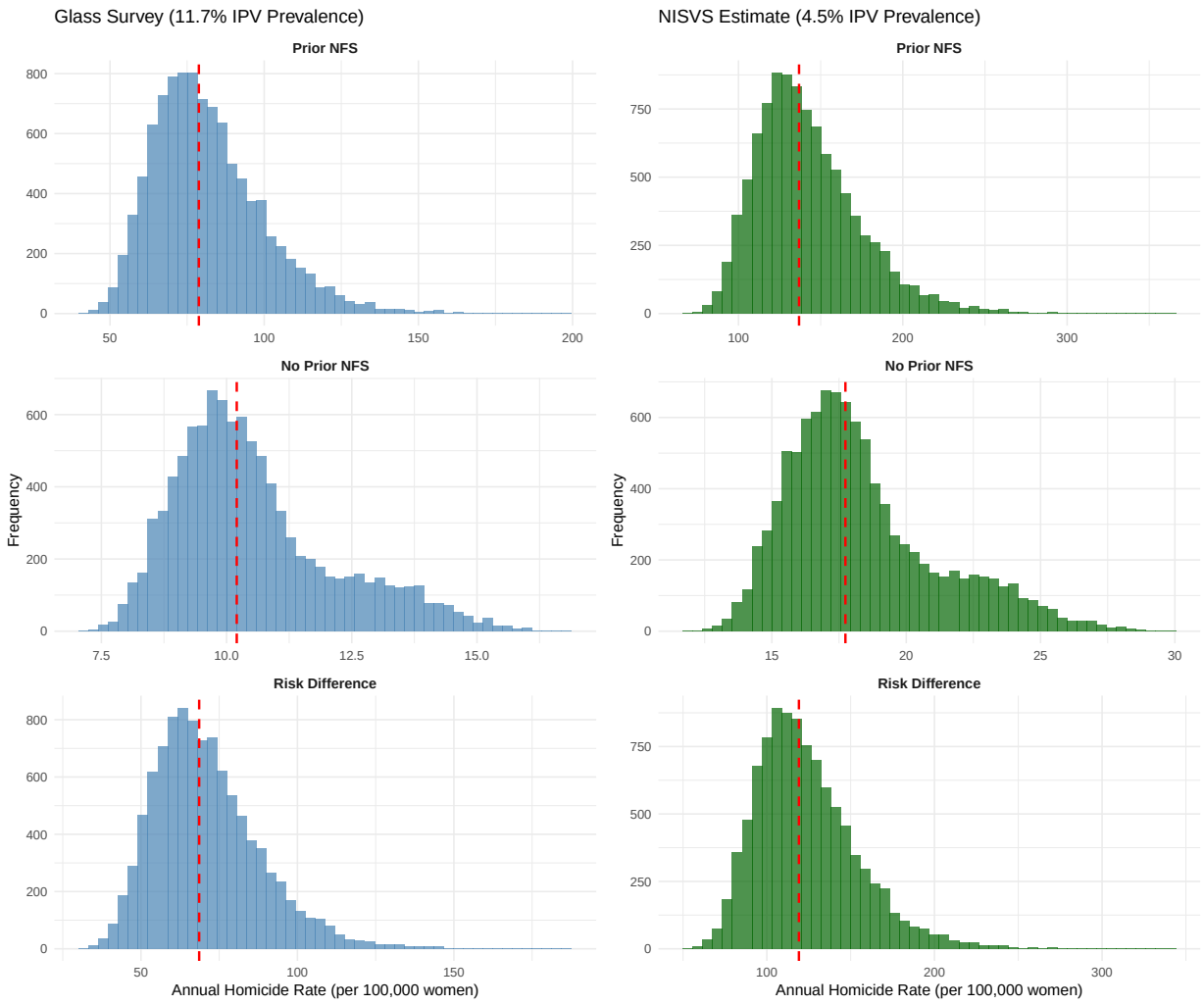
Appendix Figure 3: NIBRS event-study comparison for the Callaway–Sant'Anna estimator with and without controls. Rows show female and male negative-control outcomes; columns show the IP-coded and proxy outcome definitions. The controlled specification turns negative in later post-treatment periods for both female and male outcomes, and the male negative-control estimates become large and significant, indicating the controlled specification is not identified rather than merely noisy.



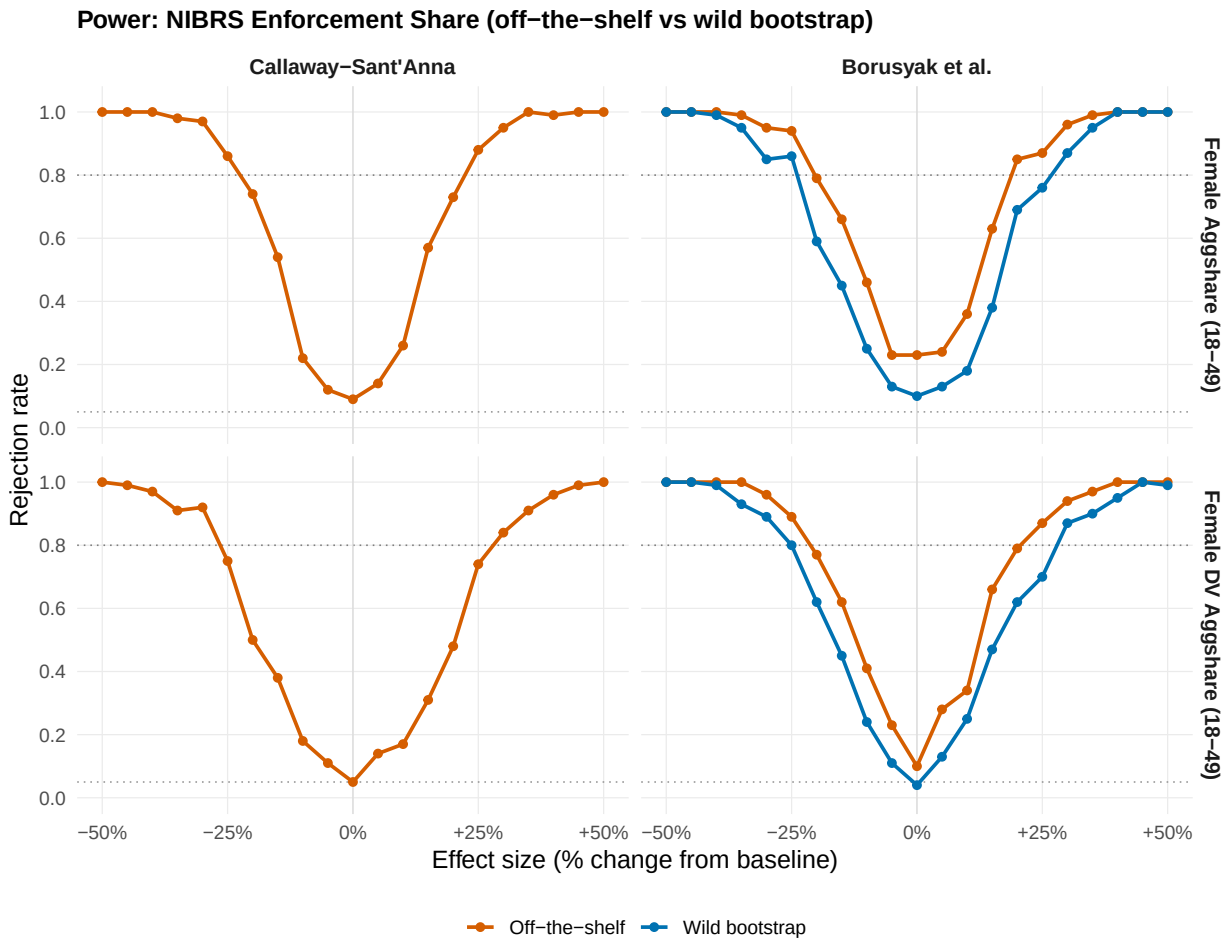
Appendix Figure 4: NIBRS event-study comparison for the imputation estimator with and without controls. Rows show female and male negative-control outcomes; columns show the IP-coded and proxy outcome definitions. The controlled female estimates remain positive, but their pre-treatment coefficients are estimated far less precisely than in the no-controls specification, so parallel trends cannot be verified; the controlled male proxy negative control drifts positive before treatment.



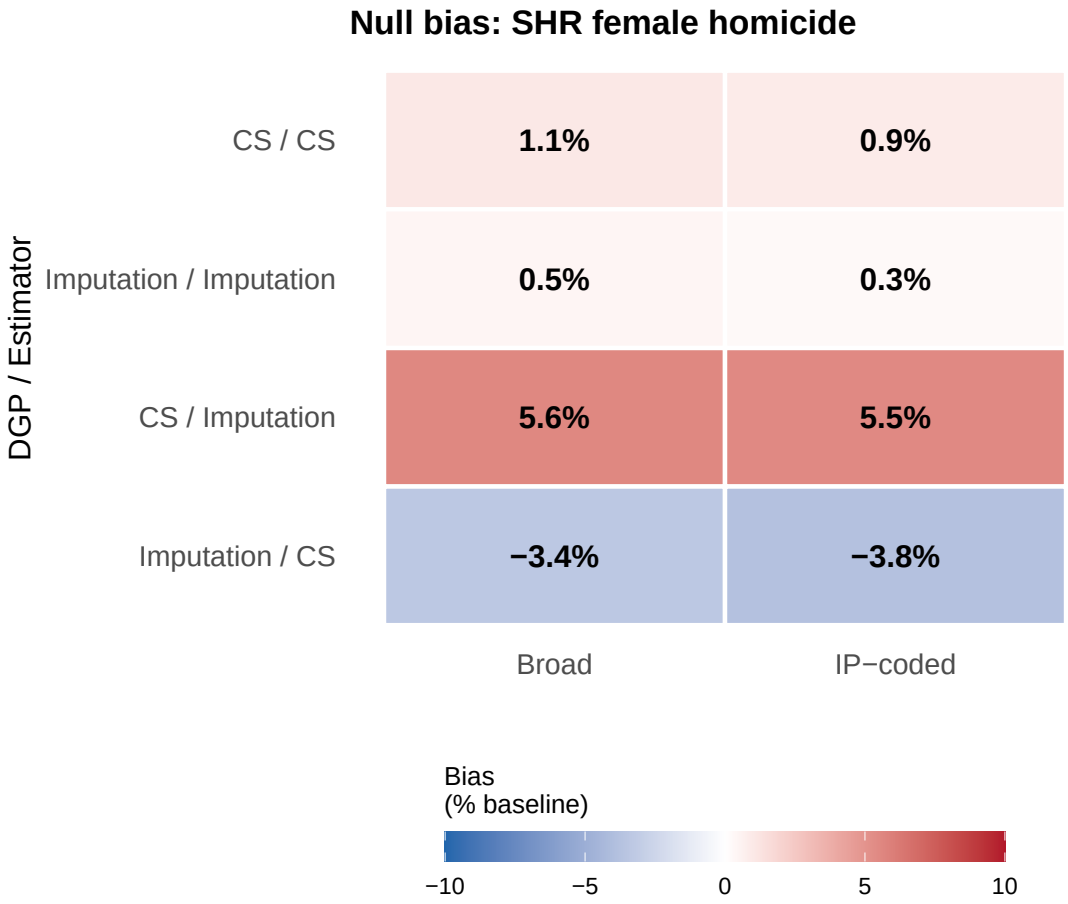
Appendix Figure 5: Event-study coefficients for the population-weighted specification on the Assis A2 missingness panel, ages 18–49, by victim sex and outcome definition (rows) and estimator (columns). Points are coefficients relative to the year before adoption (vertical line); bars are 95% confidence intervals, expressed as percent of the pre-treatment baseline rate. Under the imputation estimator, pre-treatment coefficients are positive and give way to a negative post-period; under the doubly-robust Callaway–Sant'Anna estimator, applied to the same cells, the pre-treatment coefficients show no comparable divergence, and its post-period estimates are near zero except on the broad female outcome, where the later coefficients turn negative. The vertical scale is free across rows.



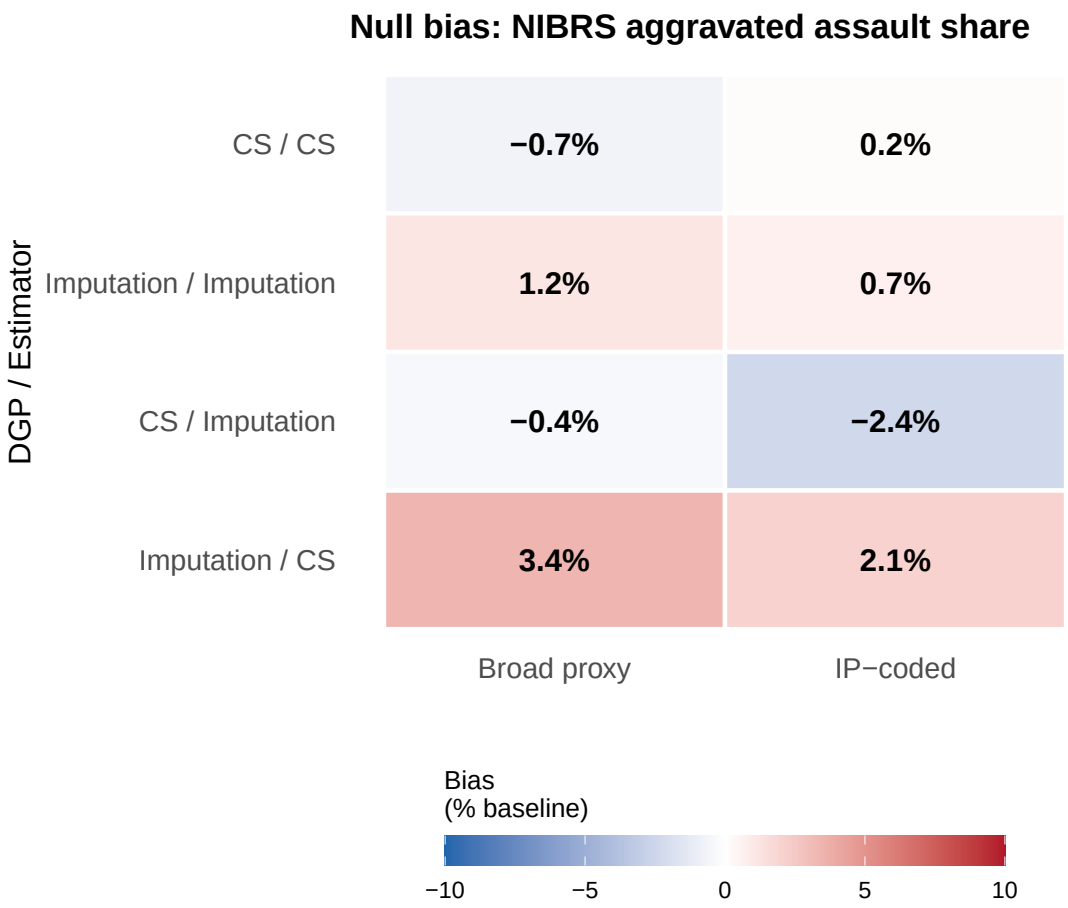
Appendix Figure 6: Monte Carlo distributions for the absolute-risk estimates (Section 8). Each panel shows the posterior distribution from 10,000 draws incorporating uncertainty in Glass et al. case-control proportions, IPV prevalence estimates, and year-to-year homicide count variation. Left column: Glass survey prevalence (11.7%); right column: NISVS prevalence (4.5%). Dashed red lines indicate medians. Under both assumptions, the absolute homicide risk among women with prior strangulation remains below 0.2%.



Appendix Figure 7: Power curves for NIBRS aggravated assault share, female victims. Layout matches Figure 4: rows show the broad and DV-coded share definitions, columns the Callaway–Sant'Anna and imputation estimators. In the imputation panels the two curves compare off-the-shelf cluster-robust inference with the restricted wild cluster bootstrap; the Callaway–Sant'Anna panels show the off-the-shelf curve alone, since that test is well sized on these shares and no correction is applied.



Appendix Figure 8: Null bias heatmap: SHR outcomes. Each cell shows the mean estimated ATT as a percentage of baseline under the null (no imposed effect). Rows are DGP–estimator configurations: in the matched rows (top two) each estimator runs on data generated under its own assumptions and bias is near zero; in the mismatched rows the estimator's assumption is violated by construction, and the systematic departure measures the cross-DGP bias.



Appendix Figure 9: Null bias heatmap: NIBRS outcomes. Layout matches Figure 8.