CrossMark

After the gun: examining police visits and intimate partner violence following incidents involving a firearm

Dylan S. Small[1] · Susan B. Sorenson[2,3] · Richard A. Berk[1,3,4]

Abstract Laws have been enacted to keep firearms out of the hands of abusers. In this study, we examined one such effort—removal of a firearm at the scene of intimate partner violence (IPV)—to assess the subsequent occurrence and number of IPV incidents responded to by police and subsequent risk of injury to the victim. Using the 28,977 IPV calls in one large U.S. city to which officers responded during the 2013 calendar year, we identified 220 first-time incidents in which offenders used (i.e., brandished, pistol whipped, shot) a pistol, revolver, rifle, or shotgun. Officers reported removing a firearm from 52 (24%) of the offenders. After using full propensity score matching to control for potential confounders, logistic and Poisson regressions were used to assess differences between those from whom a firearm was removed and those whose firearm was not removed. Firearm removal at the scene of an IPV incident appears to increase the likelihood of subsequent IPV reports to police and suggestive evidence that subsequent injury to the victim might increase as well. The offender shifting from threats with a firearm to physical violence and a change (an increase as well as a decrease) in victim willingness to summon police may account for the findings.

Keywords Domestic violence · Firearms · Intimate partner violence · Policy · Propensity score matching

✉ Susan B. Sorenson
sorenson@upenn.edu

Dylan S. Small
dsmall@wharton.upenn.edu

Richard A. Berk
berkr@sas.upenn.edu

1 Department of Statistics, Wharton School, University of Pennsylvania, Philadelphia, PA, USA

2 School of Social Policy and Practice, University of Pennsylvania, 3815 Walnut Street, Philadelphia, PA 19104, USA

3 Ortner Center on Violence and Abuse in Relationships, University of Pennsylvania, 3815 Walnut Street, Philadelphia, PA 19104, USA

4 Department of Criminology, University of Pennsylvania, Philadelphia, PA, USA

Introduction

In 1994, Congress passed the Violence Against Women Act (VAWA), which contained, among many other stipulations, provisions addressing domestic violence abusers and firearms. Specifically, VAWA prohibited a person from purchasing or possessing a firearm or ammunition if she or he was subject to a domestic violence restraining order issued by a court after a hearing at which the person to be restrained had the opportunity to appear and defend himself or herself. The prohibitions are removed when the restraining order expires. The 1996 Lautenberg Amendment extended the federal purchase and possession restrictions to those convicted of a domestic violence related misdemeanor. (The 1968 Gun Control Act already prohibited firearm purchase and possession by those convicted of aggravated assault and other felonies as well as several other categories of persons). Under each law, the prohibitions were applied if the individuals involved were currently or formerly married, currently were cohabiting or had formerly cohabitated, or had a child in common.

Springer

Some states already had similar restrictions in place and others, following the lead of Congress, enacted enabling legislation. Each state had its own criteria for who was subject to the law (e.g., persons in a same-sex relationship were able to obtain a domestic violence restraining order in some states but not others).

The laws were important because they explicitly recognized that firearms pose a particular risk in an intimate relationship. Namely, among all weapons, firearms are particularly lethal and should not be in the hands of someone who the courts have determined to pose a risk to the safety of another, specific person. Data from 2016, the most recent available, document the continued importance of firearms in homicides perpetrated by an intimate partner: Although few (3.5%) of men's murders were perpetrated by an intimate partner, in about half (49.7%) the murder weapon was a firearm; for women, the corresponding numbers are 34.7% and 58.7%, respectively.[1] Data from the past 30 years shows that, when the victim-suspect relationship is known, women are two to two-and-a-half times as likely to be killed by a male intimate wielding a firearm than to be shot, stabbed, strangled, bludgeoned to death or killed in any other way by a stranger (Kellermann & Mercy, 1992; Sorenson, 2006). And, most of the intimate partner violence homicides are preceded by a history of abuse and risk is higher when the abuser has access to a firearm (Campbell et al., 2003).

Prohibitions on the legal purchase of a firearm were implemented through the National Instant Check System (NICS) and state-administered background check systems: applications to purchase a firearm are compared to a database of disqualifying conditions, including domestic violence restraining orders and misdemeanors. The integrity of the systems depend largely on the prompt submission of all records regarding disqualifying conditions and the nature of state law (e.g., who can apply for a domestic violence restraining order; whether sales between private parties, in addition to those involving a gun dealer, are subject to background checks). During the past 20 years, NICS denied over 200,000 firearm purchases because of a domestic violence misdemeanor or qualifying domestic violence restraining order (FBI, 2018).

How to address firearms already in the possession of an abuser was less clear. Some states direct "prohibited persons" to turn in their firearms. At least one state (California) and a handful of local jurisdictions (e.g., Seattle) actively attempt to remove firearms from persons after their domestic violence court cases prohibit them from pur-

chasing and possessing a firearm. Implementation and rearrest information about one such effort in California can be found elsewhere (Wintemute et al., 2014, 2015).

Some states, intending to prevent escalation of the violence, require an officer at a domestic violence scene to remove the weapon. Pennsylvania, the state where the present investigation was conducted, requires the arresting officer to "seize all weapons used by the defendant in the commission of the alleged offense" if a probable-cause arrest (i.e., without a warrant) can be made under certain statutes (e.g., simple assault) and the officer has observed "recent physical injury to the victim or other corroborative evidence" [18 Pa. Cons. Stat. x 2711(b)]. In addition, officers exercise their judgment in all incidents to which they respond and may remove weapons when not required by law to do so. State laws—whether addressing firearm purchase and possession after determination by a court or at the scene of an IPV incident—have improved over time but remain a patchwork (Diez et al., 2017).

Evaluations of the effectiveness of the laws, all using ecological designs, have found limited to modest effects on rates of intimate partner homicide (Bridges et al., 2008; Diez et al., 2017; Dugan, 2003; Vigdor & Mercy, 2006; Zeoli & Webster, 2010; Zeoli et al., 2018). Evaluations of laws that prohibit firearm purchase and possession by those convicted of a domestic violence misdemeanor have not found an association with a reduction in intimate partner homicide. Evaluations of such laws regarding domestic violence restraining orders generally have shown modest positive results. The most recent research, which took into account various potential confounders (e.g., demographic characteristics of the population, economic indicators, per capita alcohol consumption rates), found that laws allowing the removal of a firearm at the scene of domestic violence are associated with a statistically nonsignificant 6.4–8.0% reduction in domestic violence homicides and firearm-related domestic violence homicides (Diez et al., 2017; Zeoli et al., 2018). The authors point out that a single law operates in a context of multiple laws that might combine for greater effect. The association between firearms prohibitions (officer removal at the scene and other laws) and violence against an intimate partner that does not result in death has not been studied. The present investigation was designed to address that gap in the literature.

Moreover, to date, evaluations of firearm prohibitions have not included an indicator of one potentially powerful predictor—the degree to which the laws were implemented and enforced. This is a notable shortcoming because officers have and use considerable discretion when responding to calls, including domestic violence calls, for assistance. Domestic violence is widely believe to be the single most common call for law enforcement assistance and many police departments have directives that require mandatory

---

[1] Calculations were conducted using the FBI's Uniform Crime Reporting Program Data: Supplementary Homicide Reports, United States, 2016, the first offender-victim, and the relationship categories of boyfriend, girlfriend, husband, wife, common-law husband, common-law wife, ex-husband, ex-wife, and homosexual.

arrest in such cases, however, the number of arrests fall short of the number of calls for assistance. Not involving people further with the criminal justice system may well be appropriate in certain cases, nonetheless, it would be helpful to have a sense of compliance with the domestic violence firearm-related laws and the effectiveness of the action taken.

Given the structure of law enforcement in the U.S.—county sheriff's offices, municipal police departments, and state and federal law enforcement agencies (versus, for example, France's single national police force)—it is not feasible to assess the degree to which all officers follow a given policy. Thus, we turn to individual agencies in individual jurisdictions.

A previous study, using the same law enforcement response data upon which the current investigation is based, found that most intimate partner violence (IPV) incidents involving a gun were male-on-female (79.5%) and consisted of threatening the victim with the gun (69.1%) (Sorenson, 2017). An offender was arrested in 25.8% of the gun-involved IPV incidents, a percentage that likely was affected by the fact that 71.4% of the offenders had left the scene by the time officers arrived. When an arrest was made in a gun-involved IPV incident, officers reported that they removed the weapon 41.6% of the time. When an arrest was made following the likely infliction of visible injury to the intimate partner via the firearm (was shot or pistol whipped), officers reported removing a firearm in 44.8% of the incidents. The realities of implementation of the law (e.g., whether the offender flees and presumably takes the weapon with him, the legal requirement of "physical injury to the victim or other corroborative evidence") intersect with officers' unverifiable recording of firearm removal,[2] suggesting that most abusers who use a firearm against an intimate partner in an incident to which law enforcement is summoned retain their weapon.

But what happens when a firearm is removed? The present investigation addresses three questions, namely, whether firearm removal following an IPV incident to which police were called is associated with: (a) the subsequent occurrence of non-fatal IPV reported to police; (b) the number of subsequent non-fatal IPV-related calls to police; and (c) the subsequent risk of non-fatal injury to the victim.

---

[2] Perhaps officers removed a firearm but did not record it on one of the places on the department-mandated form that would have documented their action. Although the opposite—that officers reported that they removed a gun when they didn't—is possible, it is believed to be far less likely; as the case proceeded through the criminal justice system, it quickly would become obvious that no firearm had been taken into evidence.

Methods

A retrospective cohort study design was employed using all, not a sample of, IPV cases to which police were called in Philadelphia during the full calendar year of 2013. The study was reviewed and approved by the Institutional Review Board of [name added after review].

Data

Data were drawn from the forms that the police department requires all officers to complete and file when responding to a domestic violence call for assistance. Completion of the form, in its entirety, is mandated whether or not an arrest is made. Form entries were recorded by police officers in one of three formats: (a) marking a checkbox (e.g., whether the victim had visible injuries), (b) providing specific hand-written information (e.g., offender name), and (c) a hand-written narrative describing what happened in the incident. As is not unusual in data collected for administrative purposes, many of the entries were problematic. Common difficulties included writing the required information in the wrong place, leaving some fields empty, and providing different information from that requested. The forms, collected in hard copy, were converted to electronic data that could be used for research purposes; see (reference added after review) for details on this process and the quality control measures used. Many months were spent cleaning and coding the data electronically and by hand. The form referred to the parties as "offenders" or "complainants" rather than offenders and victims or suspects and complainants. For the sake of simplicity, we refer to them herein as offenders and victims.

There were 28,977 IPV calls for assistance in 2013 in which IPV was reported, the victim-offender relationship indicated, and a dispute or altercation occurred (vs. was a request for information only). In almost all cases in which an offender had a repeat incident, the victim was the same: a sample of 100 repeat cases for an offender indicated that 97.5% of the time, the victim was the same. Our focus on subsequent incidents required that offenders, not incidents, be the unit of analysis so that key features of subsequent incidents could be attached to the appropriate offender.

Because the incident forms did not contain a unique identification number for each offender, we followed offenders across incidents using their date of birth, last name and leading initial of the first name. Data cleaning addressed, as much as possible, questionable spellings (e.g., names spelled phonetically), different credible spellings for what could be the same name (e.g., McNally vs. MacNally), the use of nicknames (e.g., Bob vs. Robert), birth dates in different formats (e.g., 8/15/13 vs. 15-8-

2013), and other issues. After multiple quality control checks we arrived at 23,076 first-time calls for an offender and 5901 repeat calls (i.e., another call for the same offender for a subsequent incident after the first one). We undoubtedly missed some repeat incidents and incorrectly flagged some incidents as repeats. With nearly 30,000 cases and an absence of ground truth it is not possible to provide a proper estimate of the error rate; our informal assessment (e.g., checking the birthdate-name match with the street address) suggests that over 90% of the repeat cases were correctly identified.

Weapon use was classified into four categories based on information provided in the narrative portion of the police form: (a) verbal abuse but no abuse with a bodily (hands, fists, or feet) or external (bat, knife, firearm, etc.) weapon; (b) abuse with a bodily weapon but no abuse with an external weapon; (c) abuse with a nonfirearm external weapon (either alone or in conjunction with a bodily weapon) but no abuse with a firearm;[3] and (d) abuse with a firearm (alone or in conjunction with a bodily weapon, nonfirearm external weapon, or both). Abuse with a firearm was defined as brandishing or threatening with the weapon, pistol whipping, and firing the weapon. Two examples of narratives are: (a) "Comp [*complainant*] states during a verbal dispute w/her boyfriend, male pushed her several times, pulled out a gun and placed it to her head. Comp told offender she was leaving him, offender told her the only way she was leaving was in a body bag. Comp was able to get away from offender when she agreed to walk with him to the store". and (b) "The compl. states her ex-boyfriend punched her in the mouth with a closed fist then used a black gun and hit her in the back of her head".

As is common in administrative data, information was missing for certain characteristics of a minority of the cases. When information was missing about whether a particular feature of an IPV incident occurred, we took it to mean that the feature in question was not present. For example, information on whether the offender was gone when the police arrived was missing in 10 of the 220 cases; we coded these 10 cases as the offender being on the scene. On occasion, statistical methods were used for imputation; for example, an offender's missing age was imputed from a regression of offender ages on victim ages.

In our study, we consider only *offenders for whom abuse with a firearm was reported during the first IPV incident reported to law enforcement* during 2013. By focusing in this subset of offenders, it is likely there was a firearm that could have been removed by responding officers. There were 220 such offenders. A firearm was removed from 52

offenders (24%) and not removed from 168 offenders (76%). We refer to the offenders as the cases and the characteristics (e.g., demographic characteristics and characteristics of the IPV incident) on the first call for the offender as the characteristics of the cases.

Statistical analysis

The statistical analysis proceeded in four steps. First, univariate summary statistics were examined to set the context. Second, a modest amount of imputation was employed to adjust for missing data. Third, full matching was employed so that valid comparisons could be made between incidents in which a firearm was removed and incidents in which a firearm was not removed. Finally, average treatment effect estimates were computed for the matched cases using one of two form of the generalized linear model: logistic regression for the binary outcome of *any* repeat incident, and Poisson regression for the *number* of repeat incidents. The four steps are described in this section. Empirical results are considered in the section that follows.

Addressing potential confounding

To evaluate the role of potential confounders, we compared characteristics of the cases in which a firearm was removed from the offender to the characteristics of the cases for which a firearm was not removed. Standard descriptive statistics were applied coupled with tests of the null hypothesis of no difference between the two groups. Fisher's exact tests were employed for discrete variables; *t* tests were employed for continuous variables. As will be seen, there were a few important differences (e.g., a firearm was far less likely to be removed when the offender had left the premises before officers arrived).

Missing data can become a confounder if systematically related to the response variable. Fortunately, for most variables we could not reject the null hypothesis of no difference in the rates of missing data between the firearm-removed versus firearm-not-removed cases (Supplementary Table 1). Missingness does not appear to be a confounder for these data.

We used full matching on several measured confounding variables to estimate the causal effect of a firearm being removed on future IPV incidents. In full matching, each matched set consists of either: (a) one treated case (i.e., removing a firearm) and one or more control cases (i.e., not removing a firearm) or (b) several treated cases and one control case. A full match is the optimal way to stratify subjects in the sense that subjects in the same stratum are as similar as possible under the requirement that every stra-

---

[3] A firearm was defined as a pistol, revolver, rifle, or shotgun. All other guns (e.g., BB guns, Taser [stun gun]) were classified as external non-firearm weapons.

Table 1 Matching variables

Offender
  Male
  Black
  Age
  History of substance abuse
  Under court supervision (on probation or pending criminal case)
  Prior history of domestic violence
  Prior history of domestic violence to which police were called
  Gone when the officers arrived
  Arrested at the scene
Victim
  Female
  Black
  Emotional reactions (sum of victim in tears or crying, distraught, shaking)

tum contain at least one treated subject and one control (Rosenbaum, 1991). We used propensity score caliper matching with a robust Mahalanobis distance as recommended by Rosenbaum (2010, Chapter 8). The matching controlled for the potential confounding variables gathered on the scene by responding officers that are reported in Table 1.

Propensity scores for the 220 cases were estimated from a logistic regression model that included these potential confounding variables (Rosenbaum & Rubin, 1983). Full matching was undertaken using a propensity score caliper of 0.4 standard deviations for the logit propensity score combined with a robust Mahalanobis distance using the variables offender's age, prior history of domestic violence, prior history of domestic violence reported to police and offender under court supervision. These variables were selected because they were thought to potentially be particularly important for the outcome. The full matching was performed using the optmatch package in R (Hansen & Klopfer, 2006; Hansen, 2007). The empirical results are addressed below.

Covariates

Covariate balance was assessed using two procedures: (a) standardized differences weighted by the number of treated cases in the matched set (Rosenbaum & Rubin, 1985), and (b) tests of the null hypothesis that each confounding variable had the same distribution within matched sets between the treated and control subjects. For binary variables, Mantel–Haenszel tests were applied; for continuous covariates, t tests were applied for regressions of each covariate on the treatment variable. As will be seen in the results section, balance was quite good.

As a precaution, we also assessed balance for variables in the data *not* used for matching. Variables for which there was statistically significant ($p < 0.05$) imbalance were included as regressors in later diagnostic analyses of causal effect estimates. Only asymptotic p values are reported for those analyses because the permutation distribution does not hold if unmatched covariates are assumed to affect the probability of treatment assignment.

Outcomes

For estimates of the effects of firearm removal, we employed the binary outcome of whether there was a future IPV incident and the count outcome of the number of future IPV incidents. We also considered the corresponding binary and count outcome for a future IPV incident with visible injuries.

For the binary outcome, we estimated the odds ratio of a future IPV incident when the offender's firearm was (vs. was not) removed. The average causal effect on the treated was estimated from a weighted, conditional logistic regression (Austin & Stuart, 2017). Treated subjects were assigned a weight of 1, while each control subject was weighted proportional to the number of treated subjects in its matched set divided by the number of controls in the matched set (Austin & Stuart, 2017). Given variation in the number of treatment and comparison cases over matched groups, this approach allows for proper comparisons. To test the null hypothesis of no treatment effect, we employed a distribution-free permutation test version of the Mantel–Haenszel test (Mantel & Haenszel, 1959).

For the count outcome, we applied a propensity score, weighted robust Poisson regression. Because the study ended on December 31, 2013, time at risk for a given offender depended on when the first IPV incident occurred (i.e., offenders whose first incident was later in the year had less time in which to re-offend). Varying time at risk was addressed by employing an offset in the regression analysis that, in effect, converted the count into a rate per day. We employed robust standard errors in response to the clustering created by our matching strategy (White, 1984; Liang & Zeger, 1986; Austin & Stuart, 2017). This also addressed possible overdispersion in case the repeat incidents did not meet the Poisson assumption of conditional independence (Cameron & Trivedi, 2009). In addition to computing an asymptotic treatment p value for the Poisson regression, we computed a distribution-free p value by permuting the treatment assignment within matched sets. The 220 cases likely are sufficient for valid asymptotic results, but the (likely less powerful) permutation test provided a useful check.

Table 2 Characteristics before matching, 220 cases in which the offender used a firearm in the first reported IPV incident for the offender, by whether the firearm was removed, 2013

Variable	Firearm collected	Firearm not collected	p value for difference between groups
Overall	0.24 (52)	0.76 (168)	
Offender			
Male	0.87 (45)	0.88 (148)	0.81
Race			0.11
Black	0.69 (36)	0.80 (135)	
Hispanic	0.10 (5)	0.03 (5)	
White	0.17 (9)	0.15 (25)	
Asian	0.02 (1)	0.01 (1)	
Age (mean, in yrs)	37.0	31.9	0.02
History of substance abuse	0.14 (7)	0.18 (30)	0.53
History of domestic violence (DV)	0.27 (14)	0.42 (71)	0.05
History of DV reports to police	0.15 (8)	0.26 (44)	0.14
Protection from abuse order (PFA) ever	0.02 (1)	0.10 (16)	0.08
PFA that could make gun subject to removal	0.02 (1)	0.07 (12)	0.31
On probation	0.02 (1)	0.08 (14)	0.20
Has criminal case pending	0.00 (0)	0.02 (3)	1.00
Fled scene	0.37 (19)	0.80 (135)	0.00
Victim			
Female	0.83 (43)	0.89 (150)	0.23
Race			0.16
Black	0.65 (34)	0.77 (129)	
Hispanic	0.06 (3)	0.04 (6)	
White	0.21 (11)	0.17 (28)	
Asian	0.04 (2)	0.01 (1)	
Age (mean, in yrs)	33.5	30.2	0.06
Relationship			< 0.001
Currently Married	0.21 (11)	0.10 (16)	0.03
Formerly married	0.02 (4)	0.01 (1)	
Currently together but not married	0.48 (25)	0.38 (64)	0.20
Formerly together but never married	0.25 (13)	0.49 (84)	0.00
Breaking Up	0.00 (0)	0.02 (3)	1.00
Incident			
Witnesses present	0.29 (15)	0.35 (58)	0.50
Children present	0.31 (16)	0.20 (34)	0.13
PFA was registered	0.39 (20)	0.25 (42)	0.08
Offender behavior			
Threatened victim	0.79 (41)	0.83 (139)	0.50
Punched victim	0.25 (13)	0.26 (44)	1.00
Strangled victim	0.06 (3)	0.17 (28)	0.07
Kicked victim	0.00 (0)	0.05 (9)	0.12
Slapped victim	0.06 (3)	0.08 (13)	0.77
Pulled hair of victim	0.04 (2)	0.09 (15)	0.37
Pushed or shoved victim	0.37 (19)	0.39 (66)	0.75
Grabbed victim	0.44 (23)	0.38 (63)	0.42
Stabbed victim	0.02 (1)	0.02 (4)	1.00
Bit victim	0.00 (0)	0.04 (6)	0.34
Sexually abused victim	0.00 (0)	0.02 (3)	1.00
Imprisoned victim	0.10 (5)	0.04 (6)	0.13

Table 2 continued

Variable	Firearm collected	Firearm not collected	p value for difference between groups
Broke In	0.06 (3)	0.04 (6)	0.44
Threw objects	0.12 (6)	0.07 (12)	0.38
Damaged property	0.14 (7)	0.10 (17)	0.61
Stalked victim	0.02 (1)	0.07 (11)	0.30
Injured children	0.00 (0)	0.02 (4)	0.57
Injured pets	0.00 (0)	0.00 (0)	1.00
Violated PFA	0.00 (0)	0.07 (12)	0.07
Scene			
Location			
Furniture disarrayed	0.10 (5)	0.04 (6)	0.14
Property damaged	0.14 (7)	0.14 (23)	1.00
Blood at scene	0.10 (5)	0.08 (14)	0.78
Victim			
Distraught	0.71 (37)	0.53 (88)	0.02
Tearful/crying	0.50 (26)	0.65 (109)	0.07
Frightened	0.67 (35)	0.67 (113)	1.00
Shaking	0.50 (26)	0.41 (69)	0.27
In pain	0.23 (12)	0.18 (31)	0.55
Clothing disarrayed	0.10 (5)	0.09 (15)	0.79
Visible injuries	0.21 (11)	0.19 (31)	0.69
Injuries not visible	0.08 (4)	0.08 (14)	1.00
Taken to hospital	0.08 (4)	0.07 (11)	0.76
Offender			
Polite	0.33 (17)	0.11 (19)	0.00
Cooperative	0.46 (24)	0.14 (23)	< 0.001
Apologetic	0.06 (3)	0.01 (2)	0.09
Angry	0.37 (19)	0.11 (18)	< 0.001
Threatening	0.50 (26)	0.52 (87)	0.87
Officer			
Gathered			
Statements	0.21 (11)	0.13 (22)	0.18
Statements from child(ren)	0.06 (3)	0.05 (9)	1.00
Photo evidence	0.04 (2)	0.02 (4)	0.63
Evidence from texts	0.00 (0)	0.01 (1)	1.00
Other evidence	0.04 (2)	0.05 (8)	1.00
Checked for			
Offender gun permit	0.31 (16)	0.11 (18)	0.00
Victim gun permit	0.27 (14)	0.02 (4)	< 0.001
Active PFA	0.39 (20)	0.25 (42)	0.08
Provided			
DV resources card	0.19 (10)	0.28 (47)	0.28
Information about PFAs	0.27 (14)	0.41 (69)	0.07
No medical treatment to victim	0.67 (35)	0.76 (128)	0.21
Information about DV advocate	0.12 (6)	0.12 (20)	1.00
Phone number for DV assistance office	0.15 (8)	0.20 (34)	0.55
Arrested offender	0.75 (39)	0.22 (37)	< 0.001

In addition to the outcome of future IPV incidents reported to the police, we used the same analyses described aConsisbove to examine whether the subsequent incidents resulted in visible injury to the victim. We planned to assess the outcome of subsequent IPV incidents reported to the police that involved a firearm, but there were no such incidents in 2013 for the 220 cases where a firearm was used in the first incident.

Results

Group comparability

Table 2 compares, before matching, the characteristics of cases in which a firearm was removed and cases in which a firearm was not removed. Some notable, statistically significant differences surfaced. For example, a firearm was less likely to be removed when the offender was gone when police arrived and when the offender had a history of domestic violence. A firearm was more likely to be removed when the police arrested the offender. Thus, we undertook full matching to reduce the serious risk of confounding between some covariates and the treatment or comparison condition.

The propensity score model coefficients, which indicate how predictors used to construct the propensity scores are related with whether a firearm was removed, are reported in Table 3. Consistent with the bivariate finding reported in Table 2 and controlling for the other variables in the
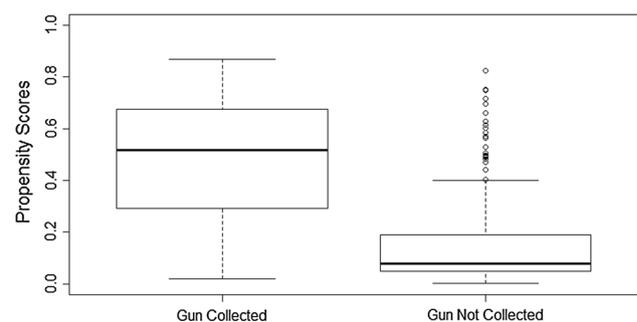
propensity score model, the log odds multiplier of 1.82 indicates a strong association between firearm removal and arrest. This finding is an initial indication that the propensity score model is performing in a sensible fashion. Side-by-side box plots of the propensity scores for when a firearm was and was not removed (see Fig. 1) show that, for the most part, the distribution of propensity scores for the two groups overlap, which bodes well for matching. Another encouraging indicator is that the propensity scores have a substantial spread. Substantial differences are found between cases in the probability that a firearm was removed by responding officers.

Covariate balance before and after matching is reported in Table 4. For example, the mean time at risk, which could be a very powerful confounder, was similar for the firearm removed and matched firearm not removed cases (189 days vs. 208 days, respectively). After matching, all of the standardized differences have an absolute value less than 0.2, which is considered to be very small (Silber et al., 2001). Furthermore, the p values for tests of the null hypothesis of no difference on a covariate after matching between the treated and matched control groups are all greater than 0.4. These results suggest that valid treatment effect comparisons between the treated and comparison groups can follow (Cochran, 1965).

A common concern with propensity score matching is that a few cases with extreme propensity score values will dominate the results. In this instance, 40 matched sets were formed. Two control (firearm not removed) cases were not matched because their propensity score was more than 0.4 logits below that of the nearest treated case. A histogram of the weights of the control units for estimating the treatment-on-treated effect of removing (vs. not removing) a firearm on the risk difference for a subsequent IPV incident are shown in Fig. 2. No control case receives a very high weight; thus, our analysis is not strongly driven by a few cases.

Table 3 Propensity score model coefficients for probability of the firearm being removed

	Log odds ratio estimate	SE	p value
Offender			
Male	0.53	0.78	0.50
Black	− 0.46	0.63	0.46
Age	0.03	0.03	0.38
Under court supervision	− 1.39	1.15	0.23
History of substance abuse	− 0.04	0.60	0.94
History of domestic violence	− 1.09	0.54	0.04
History of domestic violence reports to police	0.74	0.64	0.25
Fled scene	− 0.92	0.45	0.04
Arrested at scene	1.82	0.47	< 0.01
Victim			
Female	− 0.56	0.75	0.46
Black	− 0.64	0.62	0.30
Age	− 0.01	0.03	0.85
Index of emotional reactions	0.22	0.15	0.13

Fig. 1 Box plot of propensity scores

Outcomes

As a reminder, in the following section—as well as the entire paper—"IPV" refers to IPV to which law enforcement officers were summoned and responded.

Any subsequent IPV

For the binary outcome of having one or more subsequent IPV incidents during the follow-up period, we estimate that removing the firearm multiplies the odds of a future IPV incident by 3.30 (95% confidence interval (0.80, 14.03); two-sided asymptotic p value = 0.11; two-sided permutation p value = 0.09). The odds multiplier is big enough to imply an important *increase* in the risk of repeat incidents caused by firearm removal. In probability units, the estimated probability of a repeat incident for the comparison groups is 0.04 whereas the estimated probability of a repeat incident for treatment groups is 0.15. The p value, however, is large by conventional standards: We do not reject the null hypothesis at the 0.05 level of an average causal effect multiplier of 1.0.

Number of subsequent incidents

We applied a related approach to the count variable. The robust Poisson model coefficient estimate of the effect of removing a firearm compared to not removing a firearm on the incidence rate of future IPV is 6.45 (95% asymptotic CI

1.32, 31.43; two-sided asymptotic $p = 0.02$; two-sided permutation $p = 0.01$). Thus, one can reject the null hypothesis and conclude that firearm removal leads to an increase in future reported IPV incidents after controlling for the confounding variables in Table 4.

Recall that balance was examined for covariates directly relevant to firearm removal and subsequent incidents. Covariates expected to be irrelevant to the analysis also were examined for balance. By and large, balance was good for both. But, as a diagnostic, we undertook the analyses again including several of the variables thought to be unimportant for which the null hypothesis of balance was, nevertheless, rejected. The same pattern of results was found. Whatever confounding impact they may have had was minimal.
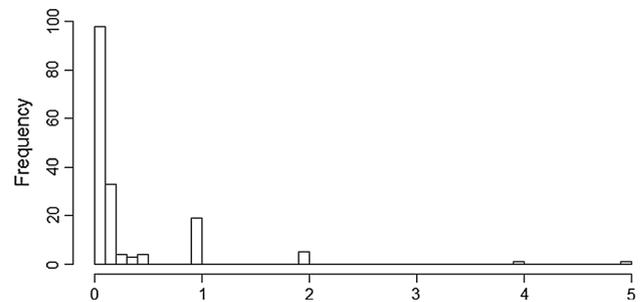
Fig. 2 Histogram of weights for control cases in estimating treatment on treated effect of removing a firearm versus not removing a firearm

Table 4 Balance before and after matching

Covariate	Firearm collected	Firearm not collected		Standardized differences		p value for no difference after matching
		Before matching	After matching	Before matching	After matching	
Offender						
Male	0.87	0.88	0.89	− 0.05	− 0.08	1
Black	0.69	0.8	0.77	− 0.26	− 0.19	0.49
Age (mean in yrs)	36.83	32.06	35.17	0.41	0.14	0.58
History of substance abuse	0.13	0.18	0.18	− 0.12	− 0.13	0.8
Under court supervision	0.02	0.1	0.04	− 0.35	− 0.07	1
History of domestic violence	0.27	0.42	0.27	− 0.32	0	0.81
History of domestic violence reported to police	0.15	0.26	0.2	− 0.27	− 0.11	0.49
Fled scene	0.37	0.8	0.36	− 0.99	0	1
Arrested at scene	0.75	0.22	0.74	1.24	0.03	1
Victim female						
Female	0.83	0.89	0.86	− 0.19	− 0.1	1
Black	0.65	0.77	0.68	− 0.25	− 0.06	0.81
Age (mean in yrs)	33.52	30.19	32.22	0.32	0.12	0.85
Index of emotional reactions (range: 0–4)	2.02	1.49	1.98	0.38	0.03	0.63

Victim injury in subsequent incident(s)

For the binary outcome of having at least one future IPV incident with visible injuries at the scene in the follow up period, we estimate that removing the firearm multiplies the odds of a future IPV incident with visible injuries by 17.49 (95% confidence interval (0.59, 1780.60); two-sided asymptotic $p = 0.04$; two sided permutation $p = 0.06$). The adjusted odds ratio is large because the odds for the control group are very small. For the comparison group, the estimated probability of a subsequent incident involving visible injuries to the victim is close to zero (0.0005) whereas for the treatment group, the estimated probability is 0.06. It is formally impossible to reconcile the two p values because the smaller value rests on untestable assumptions.

The robust Poisson model coefficient estimate of the effect of removing (vs. not removing) a firearm on the incidence rate of subsequent IPV in which the victim had visible injuries is 114.21 (95% asymptotic CI 11.18, 1166.84 31.43; two-sided asymptotic p value < 0.01; two-sided permutation p value $= 0.09$). For the comparison group, the estimated number of repeat incidents with visible injuries per 12 months is essentially 0. For the treatment group, the estimated number is 0.12. We again have dueling p values with no way to reconcile them. Perhaps the most prudent conclusion that can be drawn from these findings regarding victim injury is that the impact of firearm removal on subsequent IPV in which there is visible injury merits further investigation.

Discussion

We found no evidence that firearm removal at the scene of an intimate partner violence (IPV) incident reduces reports to police of subsequent IPV. If anything, it appears that firearm removal increases the likelihood of subsequent IPV reports to police and perhaps subsequent injury. Some may find these findings counterintuitive; a bit of context is needed.

A central goal of the firearms purchase and possession restrictions of VAWA was to reduce women's risk of injury and death caused by an abuser with a firearm. According to the same data as analyzed herein, threatening the victim with the gun is the most common use of a firearm in IPV (reference added after review). An unintended consequence of firearm removal at the scene of an IPV incident might be to shift from threats to physical violence. One could reasonably expect that after a firearm is removed from an abuser the abuser might attempt to obtain another firearm; our findings do not support this hypothesis in that, for the study population (first incidents that involved a firearm), no firearms were reported to have been used in subsequent

incidents during the same calendar year. Alternatively, in the absence of a firearm, an abuser might turn to other weapons (hands, fists, feet, bats, knives, etc.) that would result in physical injury, but these data do not support that idea, either. Although firearm removal was not statistically associated with subsequent visible injury to the victim, the direction and magnitude of the finding suggests that it could be a possibility. Further research with a larger population and a longer follow-up period is warranted.

Why these findings

In considering possible explanations for the findings, we acknowledge that firearm removal at the scene has implications beyond law enforcement officers taking possession of a gun. Such implications include the manner in which the firearm was removed (e.g., apologetically, angrily), the prospect for when it might be returned, and whether the abuser possesses or has access to other firearms. Although we don't know any of these things, they are part of the intervention and its context and may affect what happens afterwards. It is important to note that the findings cannot be explained by the association between firearm removal and whether the offender was arrested and detained (and, thus, unable to have direct contact with the victim). Arrest was one of the matching variables, thus, could not account for the differences between the groups.

The existing data do not allow us to determine the why of the findings but at least three possibilities merit mention. The first two are related to the victim's willingness to seek police assistance. First, IPV victims might feel safer and less fearful of retribution without the firearm in the home and be more comfortable to call for assistance. Second, IPV victims might perceive the police as having responded in a helpful way in the index incident (i.e., they removed the firearm) and, thus, have more confidence in police and be more willing to call for assistance.

The third possibility is related to the offender who uses a firearm against an intimate. Specifically, an abuser can say almost nothing when brandishing a firearm and the partner likely will acquiesce. If the abuser no longer has access to a firearm after one is removed at the scene, more visible domination strategies are needed: he has to threaten more often and more aggressively, which leads to more reported incidents. This possibility is consistent with the idea of firearm threats facilitating coercive control (e.g., Lynch & Logan, 2018; Sorenson & Schut, 2018). If domination is the goal of the abuser, policy would be wise to focus on that domination as the underlying mechanism, something that is much harder to fix than "simply" removing a firearm. Countries such as Scotland have incorporated coercive control into their criminal code (BBC, February 1, 2018).

How not to use these findings

The findings reported herein should not be used to conclude that firearm removal does not "work". We could not study the outcomes to which the intervention was largely directed. We do not know, for example, whether homicide was reduced. There were, thankfully, far too few to systematically analyze. At our request, the police department provided limited information about the IPV homicides that occurred during 2013 and 2014. Even for a time period twice as long as the 12 months we studied, the IPV homicide count was small ($n = 20$). Six of the subsequently fatal victim-offender pairings were involved in a total of 10 IPV incidents reported to police in 2013 and firearm removal at the scene was not an option: none involved a firearm according to the 2013 police reports. As previously noted, a necessary condition for firearm removal is the presence of a firearm and, in some states, its use in the IPV incident to which officers were summoned.

Study strengths and limitations

Limitations

The law enforcement forms used in this research are an exemplar of "messy" administrative data. Officers completed the forms by hand which undoubtedly introduced errors (e.g., leaving a checkbox blank) that could have been substantially reduced if not precluded by the use of a hand-held device. Unfortunately, the lack of hand-held devices by officers in the field is common in 21st century policing in the U.S. Thus, these data reflect the realities associated with administrative data that are recorded by hand. Moreover, it is not possible to verify the validity of the recorded information. Another weakness was the modest number of cases and the small to modest base rates, especially for the outcome of visible injury to the victim. Unmeasured confounders might also be relevant.

Although assessing subsequent non-fatal IPV reports to police during one calendar year provides an indicator of the use of publicly-funded services following firearm removal, it does not capture the full range of what happens after a firearm is removed. To have a more complete picture of the nature and scope of subsequent abuse, researchers would need to follow up with victims directly rather than rely on reports to police. Ideally, more than one calendar year of follow-up could be conducted for each case.

It also does not, obviously, capture subsequent non-fatal IPV that occurs and is not reported to law enforcement. If, after firearm removal, victims are less likely to call police when an incident occurs, the incident would not appear in the data base. Recall that the state law specifies removal of weapons used in the commission of the alleged offense, not all weapons in the offender's possession. In the U.S., gun owners own a median of 2 firearms (Azrael et al., 2017), thus, the removal of a single firearm may be of limited utility.

Strengths

One positive aspect to using data gathered for other-than-research purposes is that they provide an economical way to study a phenomenon. Data from the then-fifth largest city in the nation provided a large number of cases that comprised the population of incidents reported to law enforcement in one calendar year. And, importantly, the data—all from the same source—provided real follow-up such that the time ordering of cause and effect was clearly established. The data included a rich set of potentially relevant predictors by which to construct propensity scores and very good balance resulted; in addition, the scores were spread out and yet there were very few outliers. Despite involving different outcomes and different ways of estimating treatment, all of the findings pointed in the same direction.

Policy implications

Implications for policy would be simple if we had found dramatic reductions in IPV, but we didn't. That said, such an outcome likely is no surprise to those who have worked with abusers. IPV is motivated by a desire to dominate. Removing one means to dominate may simply call forth another means of domination. If the alternative means prove to be less lethal, harm reduction is accomplished. However, such success does not get at root causes of IPV especially given that the vast majority of incidents do not involve fatalities. The full range of other IPV interventions (that is, more than law enforcement) must be available as well.

References

Austin, P. C., & Stuart, E. A. (2017). Estimating the effect of treatment on binary outcomes using full matching on the

propensity score. *Statistical Methods in Medical Research, 26,* 2505–2525.

Azrael, D., Hepburn, L., Hemenway, D., & Miller, M. M. (2017). The stock and flow of U.S. firearms: Results from the 2015 National Firearms Survey. *The Russell Sage Foundation Journal of the Social Sciences, 3,* 38–57.

BBC. (2018). *New domestic abuse law 'could change Scotland'.* Retrieved February 1, 2018 from https://www.bbc.com/news/uk-scotland-42890990

Bridges, F. S., Tatum, K. M., & Kunselman, J. C. (2008). Domestic violence statutes and rates of intimate partner and family homicide: A research note. *Criminal Justice Policy Review, 19,* 117–130.

Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using Stata.* College Station, TX: Stata Press.

Campbell, J. C., Webster, D., Koziol-McLain, J., Block, C., Campbell, D., Curry, M. A., et al. (2003). Risk factors for femicide in abusive relationships: Results from a multisite case control study. *American Journal of Public Health, 93,* 1089–1097.

Cochran, W. G. (1965). The planning of observational studies of human populations. *Journal of the Royal Statistical Society. Series A (General), 128,* 234–266.

Díez, C., Kurland, R. P., Rothman, E. F., Bair-Merritt, M., Fleegler, E., Xuan, Z., 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,* 536–543.

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

Federal Bureau of Investigation (FBI). (2018). *Federal denials. Reasons why the NICS Section denies.* Retrieved November 30, 2018 from https://www.fbi.gov/file-repository/federal_denials.pdf/view

Hansen, B. B. (2007). Flexible, optimal matching for observational studies. *R News, 7,* 18–24.

Hansen, B. B., & Klopfer, S. O. (2006). Optimal full matching and related designs via network flows. *Journal of Computational and Graphical Statistics, 15,* 609–627.

Kellermann, A. L., & Mercy, J. A. (1992). Men, women, and murder: Gender-specific differences in rates of fatal violence and victimization. *Journal of Trauma, 33,* 1–5.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73,* 13–22.

Lynch, K. R., & Logan, T. K. (2018). "You better say your prayers and get ready": Guns within the context of partner abuse. *Journal of Interpersonal Violence, 33,* 686–711.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22,* 719–748.

Rosenbaum, P. R. (1991). A characterization of optimal designs for observational studies. *Journal of the Royal Statistical Society. Series B (Methodological), 53,* 597–610.

Rosenbaum, P. R. (2010). *Design of observational studies.* New York: Springer.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70,* 4–55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39,* 33–38.

Silber, J. H., Rosenbaum, P. R., Trudeau, M. E., Even-Shoshan, O., Chen, W., Zhang, X., et al. (2001). Multivariate matching and bias reduction in the surgical outcomes study. *Medical Care, 39,* 1048–1064.

Sorenson, S. B. (2006). Firearms use in intimate partner violence: A brief overview. *Evaluation Review, 30,* 229–236.

Sorenson, S. B. (2017). Guns in intimate partner violence: Comparing incidents that involve a gun, other weapon, or no weapon. *Journal of Women's Health, 26,* 249–258.

Sorenson, S. B., & Schut, R. (2018). Non-fatal gun use in intimate partner violence: A systematic review of the literature. *Trauma, Violence, and Abuse, 19,* 431–442.

United States. Federal Bureau of Investigation. Uniform Crime Reporting Program Data: Supplementary Homicide Reports, United States, 2016. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2018-06-28. https://doi.org/10.3886/icpsr37064.v1

Vigdor, E. R., & Mercy, J. A. (2006). Do laws restricting access to firearms prevent intimate partner homicide? *Evaluation Review, 30,* 313–346.

White, H. (1984). *Asymptotic theory for econometricians.* San Diego: Academic Press.

Wintemute, G. J., Frattaroli, S., Claire, B. E., Vittes, K. A., & Webster, D. W. (2014). Identifying armed respondents to domestic violence restraining orders and recovering their firearms: Process evaluation of an initiative in California. *American Journal of Public Health, 104,* e113–e118.

Wintemute, G. J., Frattaroli, S., Wright, M. A., Claire, B. E., Vittes, K. A., & Webster, D. W. (2015). Firearms and the incidence of arrest among respondents to domestic violence restraining orders. *Injury Epidemiology, 2,* 14.

Zeoli, A. M., McCourt, A., Buggs, S., Frattaroli, S., Lilley, D., & Webster, D. W. (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,* 1449–1455.

Zeoli, A. M., & Webster, D. W. (2010). Effects of domestic violence policies, alcohol taxes and police staffing levels on intimate partner homicide in large U.S. cities. *Injury Prevention, 16,* 90–95.