Research: Guns and Crime

Gary Kleck

College of Criminology and Criminal Justice, Florida State University, Tallahassee, FL, USA

Introduction

There is an enormous volume of published research on the possible effects of guns and control laws on crime and violence, but most of it is of such poor methodological quality that little credibility can be attributed to claims of causal effects. Worse still, while a more diverse set of topics has come to be addressed over the past few decades, the average technical quality has not improved, and may actually have declined. This entry tries to ameliorate this situation.

The entry is focused on nonexperimental quantitative scholarly research intended to assess the causal effects of gun availability or gun control measures on crime and violence, rather than descriptive or qualitative work. Researchers working in the area of guns and crime use the same array of methods that other criminologists use, so I have not wasted space on descriptions of methods that can be found elsewhere. This entry instead addresses the most serious methodological problems afflicting their application, and suggests ways to improve the work. I primarily rely on my own reviews to support my conclusions, and generally do not cite the numerous specific studies that were afflicted by a given problem, since these are included in those reviews.

Measurement Problems

To assert that X causes Y requires, at minimum, that one establish an association between the two, which in turn requires reasonably valid measures of the variables. Sometimes poor measurement of key variables is simply the product of the use of poor original sources of data. For example, many crime scholars have tried to estimate the effect of laws regulating the carrying of firearms using an oft-exploited county-level data set that simply aggregated up to the county level crime counts for individual law enforcement agencies, without adjusting for nonreporting of agencies. The result was that spurious "changes" in crime that were actually due to agencies ceasing to report their crime statistics (or resuming it) were misinterpreted as actual changes in crime (Maltz and Targonski 2002). Likewise, as in criminology as a whole, gun researchers' use of law enforcement-based crime statistics in general is afflicted by changes in the willingness of crime victims to report the offenses to the police, or such differences across

The Encyclopedia of Research Methods in Criminology and Criminal Justice, Volume II, First Edition.
Edited by J.C. Barnes and David R. Forde.

populations that can be misinterpreted as actual differences in crime frequency.

Measurement of the "Wrong" Dependent Variables

The goal of gun laws with regard to homicide is not to get people to kill with weapons other than guns but rather to reduce the number of killings, regardless of weapons used – that is, to save lives. To merely establish that a given gun law reduces *gun* homicide, or *gun* suicide, or *gun* violence in general does not establish that gun laws would save lives or otherwise reduce violence.

There is nothing wrong with analyzing a specific measure of gun violence (e.g. firearms homicide) in itself, as long as it is done in conjunction with analysis of the corresponding *non*gun form of violence, e.g. analyzing the effect of a control measure on nongun homicide as well as gun homicide. Indeed, doing so provides a stronger test of the hypothesis that gun levels or gun control had a causal effect – if the hypothesis is correct, gun violence should be affected more than nongun violence (Kleck and Patterson 1993, pp. 265–266). It is dubious practice, however, to assess the impact of a gun control measure on *only* gun violence.

Measurement of an Excessively Heterogeneous Gun Control Variable

It is not informative to lump a diverse batch of very different kinds of gun controls into a single index of "gun control strictness" and then see how that index is correlated with gun violence rates. This procedure makes it impossible to determine which specific types of gun control reduce violence. If 10 policies are lumped together but only one of them reduces violence while the other nine are ineffective or even counterproductive, surely it is essential that the analyst be able to discern which policy works and which do not.

Measures of "Gun Availability" that Overlap the Dependent Variable

In order to conclude that one variable causes another, the variables must at a minimum be distinct and separate, since it is tautological to assert that X causes X. Put another way, purported causes should not overlap their purported effects. In testing whether levels of gun availability affect crime rates, one should not use a measure of the former that is itself partly or entirely a crime rate. For example, some authors have tried to measure neighborhood-level prevalence of guns using data on gunshots detected by acoustic detection systems, and related this measure to neighborhood rates of gun crime. The problem is that nearly all municipalities forbid firing guns within city limits. Thus most firings of guns in a city would themselves be crimes, and their frequency would necessarily be an indicator of the residents' criminality. To find a positive association between such a measure of "gun prevalence" and crime rates would at best merely confirm the universal understanding that all crime rates are affected by the population's willingness to break the law.

A more obvious variant of this general problem is the inclusion of a common component in both the independent and the dependent variable. The percent of suicides committed with guns (PSG) is an excellent, well-validated, macro-level measure of gun ownership levels but should not be used if the dependent variable is the suicide rate, since the number of gun suicides is a component common to both the gun measure (gun suicides/total suicides) and the total suicide rate (gun suicides + nongun suicides/population). For the same reason, the percent of homicides committed with guns may be a reasonable measure of gun availability among violent criminals, but should not be used when the dependent variable is the homicide rate (Kleck 2004; Kleck and Patterson 1993).

Perhaps the most common measurement error in this field is to use PSG to measure changes in gun levels over time, in panel and other longitudinal studies. Some scholars have assumed that validation of this measure for use in cross-sectional studies also implies that it is valid in cross-temporal studies. In fact, PSG has no correlation at all over time with direct survey measures of gun ownership, and should not be used in longitudinal studies (Kleck 2004, pp. 19–26; Kovandzic et al. 2013, pp. 485–490).

Sample Bias

The degree to which researchers can generalize their results beyond the limited set of cases they studied to some larger population is a function of how representative the study sample is of the population. It is rare that researchers either are indifferent to generalization or studied entire populations and thus did not need to be concerned with the issue.

Some scholars have drawn conclusions about the entire population when their samples provided no formal basis for such broad generalizations. Kellermann and his colleagues (1993) carried out a case–control study of the impact of household gun ownership on homicide victimization. Although their samples were largely confined to people at unusually high risk of being murdered, their main conclusion pertained to the population as a whole: "people should be strongly discouraged from keeping guns in their homes" (p. 1090). Few question whether there are violence-prone subsets of the population, such as persons convicted of violent crimes, for whom gun possession elevates the risk of homicide, but this implies nothing about the population as a whole.

Studying only crimes known to the police likewise suffers from sample bias that distorts key findings regarding guns and violence. Victimization surveys show that victims rarely report to police crimes in which they were neither injured nor lost property. This distorts estimates of the effectiveness of victim self-protection actions, since it means that successful self-protection actions such as defensive gun use tend to be excluded from samples of crimes reported to police (Tark and Kleck 2004). Use of samples from victimization surveys, which include crimes not reported to the police, therefore are better, though even they too are likely to underrepresent incidents without victim harm.

Surveys of U.S. adults are commonly used to estimate the prevalence of gun ownership, and the use of guns for self-protection, but well-established sample biases of national survey samples distort these estimates. Underrepresentation of males, racial/ethnic minorities, and poor people implies underrepresentation of people at greater risk of victimization, which leads to too few people who own guns for reasons of self-defense, too few who have used guns for self-protection, and too few victims of gun crimes (Kleck 2020).

Criminologists have also studied samples of guns recovered by the police and submitted for tracing by the Bureau of Alcohol, Tobacco, Firearms and Explosives, often for the purpose of assessing the importance of gun traffickers or corrupt licensed dealers in supplying guns to criminals. Some scholars assert that certain attributes of these guns can be interpreted as indirect indicators that a gun has been trafficked. For example, some believe that if it took only a short time for a gun to move from its first retail sale to recovery by police in connection with a crime, this is one rough indication that a trafficker was involved in moving the gun. Problems arise when analysts believe they can learn about crime guns in general from studies of traced guns. Police do not randomly sample from the guns they recover and then submit a representative sample of those guns for tracing, but rather nonrandomly select a subset of guns based on their investigative needs. For example, police prefer to select guns for tracing that appear to have been first sold at retail fairly shortly before police recovery, since this provides them with fresher leads as to the sources of guns. One result is that these samples overstate the share of crime guns that show signs of having been trafficked (Kleck and Wang 2009). Police are likewise more likely to request tracing on politically "hot" types of guns, so that firearms like "assault weapons" are grossly overrepresented in samples of traced guns (Kleck 1997, pp. 112, 141–143).

Aggregation Bias

Drawing conclusions about individuals (or smaller aggregates) based on findings pertaining to large aggregates, however, can be distorted by aggregation bias – the possibility that relationships prevailing among larger aggregates are different among individuals or smaller aggregates. For example, many macro-level studies of guns and violence have used state-level data, found a positive association between gun levels and violence rates, and concluded that the former caused the latter. One problem is that the smaller areas within states that have more gun ownership are not the ones that have more violence. Gun ownership is highest in rural areas and lowest in urban areas, while violence levels are distributed in exactly the opposite way (Kleck 1997). This is confirmed by analyses of smaller, more homogenous aggregates like cities or counties, which find no significant positive association between gun levels and violence rates (Kleck and Patterson 1993; Kovandzic et al. 2013).

On the other hand, some kinds of *dis*aggregation can be problematic, such as studying some selected subsets of the population and not others. "Data-dredging" is the practice of repeatedly testing a hypothesis within selected subsets of a data set until one finds support for a favored hypothesis – and possibly reporting only the supportive findings. There is nothing wrong with separately testing a hypothesis for subsets of the population per se, e.g. first among males and then among females. Studying *only* one of them, however, raises the suspicion that the analyst indulged in data-dredging and selectively reported only findings supportive of a favored conclusion.

Causal Order – Which Is Cause and Which Is Effect?

It is crucial to establish causal order. For example, while gun ownership rates prevailing within a population might cause higher homicide rates, higher homicide rates could also motivate more people to acquire guns for self-protection. Although this question of causal order has been recognized for decades (e.g. see Kleck 1979), most macro-level studies of homicide do nothing to establish causal order (for a review, see Kleck 2015). Some researchers have tried to eliminate this problem by relating the previous year's gun rate to the current year's homicide rate, based on the appealing reasoning that nothing that happens this year can affect anything the previous year. Merely lagging the gun variable, however, actually does nothing to solve the statistical problem that arises if homicide rates affect gun levels – the gun variable, even if lagged, will be correlated with the error term of the equation predicting the homicide rate, so estimates of the gun level's effect will be biased and inconsistent (Kovandzic et al. 2013).

A more feasible statistical solution to the causal order problem is the use of "instrumental variables" (IVs) estimation procedures. For example, one might assess the effect of gun rates on homicide rates using IV methods, but they are effective only if the IVs have three properties: (i) they are exogenous (not affected by other endogenous variables – the homicide rate in this example),

(ii) they are "relevant," i.e. they affect the endogenous variable they are designed to predict or "instrument" (the gun level in this example), and (iii) they are "valid" in the special statistical sense of not having any direct effect on the other endogenous variable(s) (the homicide rate in this example). The technique, including methods for testing for exogeneity, relevance, and validity of instruments, is explained by Kovandzic et al. (2013), who unfortunately are the only scholars to properly apply these methods to research on guns and violence (Kleck 2015).

Other techniques for addressing causal order have been applied in individual-level research. For example, research on the causes of gun ownership has focused on the impact of fear of crime and perceived risk of victimization on the ownership of firearms, especially handguns. Causal order problems can arise because fear or perceived risk might initially motivate the acquisition of guns for protection, but once acquired, gun possession could reduce fear and perceived risk. One traditional way to distinguish these effects is the use of survey panel methods in which the analyst tests (i) the effect of fear measured at an earlier time on gun ownership at a later time, and (ii) the effect of gun ownership at an earlier time on fear at a later time (Hauser and Kleck 2013). A second method relates current levels of fear or perceived risk of victimization to *plans* to get a gun for protection in the future, reducing uncertainty about causal direction based on the assumption that merely planning to acquire a gun would have no measurable effect on fear (Kleck et al. 2011). A third method applied to a closely related problem used multi-level analysis, relating city-level crime rates to individual-level gun ownership, reducing uncertainty about causal direction based on the assumption that a single survey respondent's gun ownership (or lack thereof) could not have any measurable effect on an entire city's crime rates (Kleck and Kovandzic 2009).

Studies of the effects of self-protection actions on victim injury must take account of which occurred first. In most crime incidents in which the victim resisted the offender and was injured, the victim took defensive action only *after* the offender injured them. In a study of national victimization surveys, although 24% of crime incidents involved both victim self-protective actions and injury, only 3.5% of all incidents involved victims who were injured *after* taking those actions (Tark and Kleck 2004, p. 872).

Poor Control of Confounding Variables

In nonexperimental research, to separate the impact of a variable of interest from other factors that affect the same dependent variable, it is usually necessary to measure and statistically control for confounding variables. A confounder is a variable that (i) also affects the dependent variable *and* (ii) is correlated with the independent variable of interest (IVI). Only the effects of variables with both of these properties could be confused with the effects of the IVI. The more confounders one controls, the more confident one can be that any remaining association between the IVI and the dependent variable is due to the IVI's causal effects.

Most researchers studying guns and violence do a poor job of controlling confounders. A systematic review of 41 macro-level studies of the effect of gun levels on crime rates found that 14 did not control for a single confounder, and only six controlled for more than five significant control variables, not all of which were confounders (Kleck 2015, p. 44). Macro-level research on the effect of gun levels on suicide rates is even worse. A review found that in 26 of 32 analyses, researchers did not control for a single variable that was shown to be significantly related to suicide rates (Kleck 2019, p. 942). When even a few genuine confounders were controlled, any appearance that gun levels affect suicide rates disappeared (Kleck 2019, p. 943).

Individual-level studies are no better. Studies of the effect of access to guns on suicide usually employ a case–control design in which persons who committed suicide are compared with either individuals still living or persons who died from nonsuicide causes. Of 16 case–control studies of suicide, only three controlled for more than four likely confounders – out of at least 19 known confounders (Kleck 2018a, p. 316; see also Kleck and Hogan 1999 regarding case–control studies of homicide).

Suggestions for Improvement

Make a More Serious Effort to Identify Likely Confounders If one hypothesizes that X causes Y, do thorough reviews of two literatures: (i) research in which Y is the dependent variable, and (ii) research reporting the correlates of X, whether it was treated as a dependent or an independent variable. Variables that usually show significant associations in both literatures constitute a reasonable initial list of likely confounders that the researchers should try to measure and control for. Strong theory pointing to other variables as likely confounders can be used to supplement this list. Obvious though these procedures may seem, it is clear they were not used to generate the list of variables controlled in most studies of guns and violence (Kleck 2018a).

Make Procedures to Establish Causal Order a Central Element of the Research Design It is not satisfactory to merely mention the absence of such procedures when listing limitations of the research at the end of the research report, or to speculate one's way around the problem. If one's research design does not satisfactorily address causal order, the researcher should employ one that does.

Use Validated Measures of Gun Levels in Macro-Level Research The percent of suicides committed with guns has been validated for use in cross-sectional studies but is *not* valid for use in longitudinal studies. It also should not be used in studies of suicide rates due to the common components problem (Kleck 2004; Kovandzic et al. 2013).

Expand the Body of Data Establish annual rates of state-level household firearms prevalence by making gun ownership questions a part of the standard questionnaire used in the national Behavioral Risk Factor Surveillance Surveys (BRFSS). Currently the gun ownership questions are asked only in states that choose to use an optional module of questions, which few states do.

Analysis of gun crime incidents could be greatly enhanced if researchers could combine hospital information about patients treated for gunshot wounds and police information about the crime and offenders. Medical confidentiality laws preclude most researchers from accessing hospital information, so they should be amended to allow wider researcher access.

Minimize Aggregation Bias in Macro-Level Studies Use the smallest, most homogenous aggregates for which the requisite data are available. Thus, studies are better done at the city or county level than at the level of states, which are in turn preferable to regions or nations. Sometimes data availability is greater for larger aggregates, and an essential variable is measured only for such an aggregate, in which case the costs of aggregation bias may be outweighed by the benefits of measuring more variables, but it is otherwise generally problematic to analyze large aggregates.

Understand the Likely Consequences of Problems that Cannot Be Fixed or Avoided For example, it has been shown that many gun owners deny their gun ownership in surveys (Kellermann et al. 1990; Kleck 1997, pp. 64–68), while no evidence indicates that nonowners falsely claim gun ownership.

Thus, researchers should understand that surveys are likely to understate gun ownership. The most commonly used national survey mode, telephone surveying, is unfortunately the least effective in getting respondents to report sensitive behaviors (Kleck and Roberts 2012). Likewise, research on survey methodology indicates that nearly all of the documented sources of error in surveys, including both sample bias and response error, tend to depress estimates of defensive gun use frequency, implying survey estimates of this frequency are likely to be too low (Kleck 2018b, 2020).

References

Hauser, W. and Kleck, G. (2013). Guns and fear: a one-way street? *Crime and Delinquency* 59: 271–291.

Kellermann, A.L., Rivara, F.P., Banton, J. et al. (1990). Validating survey responses to questions about gun ownership among owners of registered handguns. *American Journal of Epidemiology* 131: 1080–1084.

Kellermann, A.L., Rivara, F.P., Rushforth, N.B. et al. (1993). Gun ownership as a risk factor for homicide in the home. *New England Journal of Medicine* 329: 1084–1091.

Kleck, G. (1979). Capital punishment, gun ownership, and homicide. *American Journal of Sociology* 84 (4): 882–910.

Kleck, G. (1997). *Targeting Guns: Firearms and Their Control*. NY: Aldine.

Kleck, G. (2004). Measures of gun ownership levels for macro-level crime and violence research. *Journal of Research in Crime and Delinquency* 41 (1): 3–36.

Kleck, G. (2015). The impact of gun ownership rates on crime rates: a methodological review of the evidence. *Journal of Criminal Justice* 43 (1): 40–48.

Kleck, G. (2018a). The effect of firearms on suicide. In: *Gun Studies: Interdisciplinary Approaches to Politics, Policy, and Practice* (eds. J. Carlson, K. Goss and H. Shapira), 309–329. New York: Routledge.

Kleck, G. (2018b). Response errors in survey estimates of defensive gun use. *Crime & Delinquency* 64 (9): 1119–1142.

Kleck, G. (2019). Macro-level research on the effect of firearms prevalence on suicide rates: a systematic review and new evidence. *Social Science Quarterly* 100 (3): 936–950.

Kleck, G. 2020. The prevalence of defensive gun use: possible sources of error and the results of 21 national surveys. *Social Science Research Network* at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3607221.

Kleck, G. and Hogan, M. (1999). A national case-control study of homicide offending and gun ownership. *Social Problems* 46 (2): 275–293.

Kleck, G. and Kovandzic, T. (2009). City-level characteristics and individual handgun ownership: effects of collective security and homicide. *Journal of Contemporary Criminal Justice* 25 (1): 45–66.

Kleck, G. and Patterson, E.B. (1993). The impact of gun control and gun ownership levels on violence rates. *Journal of Quantitative Criminology* 9 (3): 249–287.

Kleck, G. and Roberts, K. (2012). What survey modes are most effective in eliciting self-reports of criminal or delinquent behavior? In: *Handbook of Survey Methodology* (ed. L. Gideon), 415–439. New York: Springer.

Kleck, G. and Wang, S.-Y. (2009). The myth of big-time gun trafficking and the overinterpretation of gun tracing data. *UCLA Law Review* 56 (5): 1233–1294.

Kleck, G., Kovandzic, T., Saber, M., and Hauser, W. (2011). The effect of perceived risk and victimization on plans to purchase a gun for self-protection. *Journal of Criminal Justice* 39 (4): 312–319.

Kovandzic, T., Schaffer, M., and Kleck, G. (2013). Estimating the causal effect of gun prevalence on homicide rates: a local average treatment effect approach. *Journal of Quantitative Criminology* 28 (4): 477–541.

Maltz, M.D. and Targonski, J. (2002). A note on the use of county-level UCR data. *Journal of Quantitative Criminology* 18: 297–318.

Tark, J. and Kleck, G. (2004). Resisting crime: the effects of victim action on the outcomes of crimes. *Criminology* 42: 861–909.