The Small Arms Survey Estimates of National Civilian Firearms Ownership: An Assessment

Gary Kleck

College of Criminology and Criminal Justice

Florida State University

Tallahassee, Florida 32306-1127

(850)559-0922

gkleck@fsu.edu

September 25, 2020

Keywords: Small Arms Survey; gun ownership rates; cross-national homicide studies

Abstract

Researchers doing cross-national research on the associations between gun ownership rates and violence rates have begun to use the Small Arms Survey **(**SAS) to measure national rates of gun ownership. The main problem with the SAS is that it counts apples for some nations, oranges for others. That is, even though its developers clearly believe it generates estimates that are comparable across nations, they are not, because the methods measures different quantities for different nations. A direct test of the validity of SAS civilian gun totals, using the percent of suicides committed with guns as a criterion, indicates that the SAS estimates have little validity as measures of national firearms prevalence. A better measure is the percent of suicides committed with guns.

The Small Arms Survey **(**SAS) is a research project housed at the Graduate Institute of International and Development Studies in Geneva, Switzerland, and is partially funded by Swiss national and cantonal governments. It focuses on a number of issues related to small arms and armed violence, but one of its main data-gathering efforts is to measure national totals of firearms, including the majority of the world's guns that are owned by civilians. SAS publishes estimates of civilian gun totals for over 200 countries and territories (Karp, 2018). Researchers doing cross-national research on the associations between gun ownership rates and violence rates have begun to use SAS to measure national rates of gun ownership. This is unfortunate because SAS estimates have poor validity and are not uniformly comparable across nations. Although SAS developers clearly believe that it generates estimates that are at least roughly comparable across nations, SAS use of fundamentally difference estimation procedures for different countries effectively means that it counts apples for some nations, oranges for others.

Methods Used to Create SAS Estimates – Poorly Documented and Noncomparable

There are two publicly available descriptions of the methods used by SAS to estimate national civilian firearms totals (Small Arms Survey, 2007; Karp, 2018). Both descriptions are poorly documented, often vague, and extremely incomplete (see Small Arms Survey 2007, Chapter 2 "Completing the Count: Civilian Firearms" [pp. 55-59]; Karp, 2018). The following is the best that can be inferred from these descriptions of the estimation methods.

SAS uses four radically different methods for estimating the number of civilian guns in a given country.

1. For the minority of nations for which national governmental counts of registered guns are available, estimates are based the count of registered firearms, combined with a huge upward adjustment for unregistered guns, which the authors call "extrapolation from registration data." The authors claim that the number of unregistered guns is typically 2.6 times the number of registered guns (p. 55), implying that the total number (registered and unregistered) is 3.6 times the number registered. This claim was apparently based on a bivariate regression analysis in which the dependent variable was a "comprehensive" independent estimate of unregistered guns and the sole predictor was the number of officially registered guns, using data from the nations for which both kinds of information were available (Small Arms Survey, 2007, pp. 55-56). The estimated parameters of this regression equation (constant and coefficient) were not reported – only its R^2. The authors do not explain in what sense the estimates of unregistered guns were "comprehensive," and offer no evidence that they were in fact comprehensive in the sense of counting all unregistered guns.

   The same upward adjustment is applied to all countries with registration systems, regardless of differences across the nations in the completeness of registration, efforts to enforce registration requirement, or willingness of gun owners to comply with registration requirements. SAS staff acknowledge that this method does not work equally well for all of the nations to which it was applied (Small Arms Survey, 2007, p. 55), but did not seem to recognize that this implies that resulting estimates are not comparable across nations, since the degree to which registered gun counts understate the total civilian guns differs from country to country. Perhaps in response to this problem, SAS staff state that estimates for some nations "have been adjusted," but do not report how

this was done (Small Arms Survey, 2007, p. 55). This sort of incomplete and undocumented description of methods pervades the SAS reports.

The so-called "comprehensive" independent estimates of unregistered guns were computed as an average of estimates generated by a variety of supposedly "expert" sources. By the authors' own admission, some of these "independent estimates" were little more than guesses (Small Arms Survey, 2007, p. 53). If the "comprehensive" estimates were anything but comprehensive, the upward adjustment for unregistered guns would be inaccurate. Further, the kinds of sources on which these "comprehensive" estimates were based differed from nation to nation and thus were not comparable across nations. This necessarily means that registration counts are adjusted upward too little for some nations and too much for others, reducing comparability across countries.

SAS staff also arbitrarily drop some registration figures based on their subjective judgment that they "appeared suspiciously low" (Small Arms Survey, 2007, p. 67, fn. 3). Since the authors concede that registration figures in general are wildly low as counts of total civilian firearms (Small Arms Survey, 2007, p. 55), it is unclear by what standard the rejected figures were considered *especially* inadequate while other registration figures were accepted.

2. For countries (n=56 in 2017) for which there were national surveys asking about firearms ownership, the number of gun-owning households or individual gun owners is computed, based on the most recent available survey. Since most surveys do not ask how many guns were owned by each household or person, SAS staff arbitrarily assume that each gun-owning household contains exactly 1.5 guns - notwithstanding likely between-nation

differences in numbers of guns per owner attributable to differences in income and gun control regulations.

Besides 28 countries covered by a European Union survey in 2014, the rest of the surveys used by the SAS were necessarily a hodge-podge of mostly one-time surveys that were not standardized across countries. The authors do not claim that all the surveys asked the gun ownership question the same way (to the extent feasible in light of translation difficulties), or that any adjustments were performed to take account of differences. For example, they do not claim that all surveys covered guns kept in vehicles, garages, sheds, and other locations outside the home, in addition to those kept in the owner's home. Nor do they indicate that all the surveys explicitly cautioned respondents to not report air guns, CO2 guns, BB guns, and similar guns that are not firearms. To the extent that surveys differed in these respects regarding the gun ownership questions, different subsets of civilian guns were being estimated in different countries, further reducing cross-national comparability.

Since available surveys were fielded in different years, additional adjustments needed to be made to take account of changes over time in gun ownership. Thus, if the most recent survey was fielded in 2014, to get estimates for 2017 the result needed to be adjusted to take account of three years of change. SAS staff arbitrarily assumes that national gun counts increase by exactly one percent each year, for nearly all countries – even though they acknowledge that rates of firearms growth differ radically across countries (Small Arms Survey 2011). This fixed adjustment means that gun counts will be adjusted upward too little for some countries, too much for others. Recognizing this problem, staff deviate it from in some cases where it is known that countries had

experienced unusual "deflation" (gun loss) or "inflation" (growth) rates. Since information of this sort will be available for some countries and not others, this means that the possibly erroneous one percent growth rate will be applied to some nations and not others, thereby introducing yet another potential source of noncomparability.

3. A third, perhaps most dubious, method used in the SAS is to compute averages of what SAS staff generously refer to as "expert estimates." SAS staff insist that "expert estimates are important and should be considered seriously" (Karp, 2018, p. 5) but do not explain why. Further, they do not define what it takes for a person to be defined as an expert, and do not identify any experts by name or by position. They do not require that "expert estimates" have any particular empirical foundation – apparently they can be little more than guesses. SAS staff acknowledge that some expert estimates can "differ by a factor of ten" (Small Arms Survey 2011, p. 2), without concluding that this suggests they are of highly dubious reliability. The only SAS adaptation to such inconsistencies is to discard the "highest and lowest expert estimates," even though one of these extreme estimates could actually be the most correct one. Even this procedure is not followed consistently or objectively, since staff do this only if these lowest or highest estimates were considered "too extreme," according to some unstated, possibly subjective assessment by staff (Karp, 2018, p. 8). Even when these estimates are similar to each other, it might only mean the experts providing them are simply parroting each other's previous claims, or that they share similar misperceptions about the prevalence of guns in their country. One "expert" may tell a reporter there are X million guns in the country, solely because she or he has seen a previous "expert" guess that there were X million. The result is a false sense of consensus or consistency. Finally, the average of the expert

estimates not regarded as being "too extreme" is adjusted to inflate the estimate up to the estimation year – presumably again based on the arbitrary, fixed one percent annual increase assumed by staff.

When both survey-based and average "expert estimates" are both available for a given country, they are averaged (Karp, 2018, p. 8) – even though the survey-based estimate might be more accurate and probably more firmly grounded in empirical evidence. Assuming this step involves simply computing the arithmetic average, it effectively gives equal weight to the results of a scientifically designed probability sample survey and an average of "expert estimates" that may be little more than guesses. SAS staff in general invest too much confidence in the value of averaging estimates, taking it on faith that this will generally improve estimates. Averaging good and bad estimates, however, is not a scientifically sensible procedure, as it serves to dilute the influence of the best available information.

4. Finally, the most vaguely described SAS method is used when none of the first three data sources are available for a country. SAS staff will only say that number of guns are "estimated using analogous rates from comparable countries and territories where the research team, guided by available research and media reporting, appraised whether analogous comparisons were plausible" (Karp, 2018, p. 8). No further detail is provided on how staff identified "comparable" countries, nor do SAS staff say what they do if they cannot identify any comparable countries. Do they use whatever country is most similar, even if it is not very similar? Is comparability based on similar levels of gun ownership? Similar trends in earlier periods? Location in the same region? Similar level of economic development? Do staff do anything to insure that the same procedures are used

to identify comparable countries, or can procedures vary? Are there large doses of subjectively in how the procedures are applied? These too would reduce cross-nation comparability of SAS gun estimates.

SAS staff appear to be unfamiliar with scientific standards such as replicability and the need to fully document one's methods. There is no way an outside researcher would be able, on the basis of published SAS descriptions, to reproduce SAS estimates or determine whether SAS staff actually applied the procedures they outline. Their descriptions of the methods are too sketchy and incomplete, and the methods seem to entail subjective judgements of SAS staff that other researchers might not share.

The underlying problem in the SAS effort to estimate gun totals may be an ill-conceived effort to achieve comprehensive coverage of all nations, including those for which no genuinely usable gun data are available. It certainly sounds impressive when SAS boasts that their estimates cover over 200 countries and territories, and this probably is a big part of the appeal of SAS data to users. This comprehensive coverage, however, comes at the expense of accuracy and comparability of gun counts. For many countries, it would have been better for SAS to simply recognize that no meaningful estimate could be produced, and to say, in effect, "we don't know." Instead, SAS staff appear to operate on the idea that any old estimate is better than none at all.

At no point do SAS staff ever caution readers that their estimates of civilian gun totals are not comparable across nations. Quite the contrary, they convey exactly the opposite impression, using their estimates to rank nations by their rates of guns per capita (see especially Small Arms Survey 2007, Annexe 4). Nevertheless, it is clear that these estimates cannot be legitimately

used to compare gun levels across nations, either when the estimates were created using different methods among their three broad methodologies, or even when the estimates were all created by the same one of these three methodologies.

No Validity Tests

The SAS estimates have never been empirically validated. That is, SAS staff have not publicly demonstrated a strong correlation between their estimates and other previously accepted measures of national gun levels used for cross-national comparisons – even measures that they themselves used in generating their own estimates, such as direct survey measures of household gun prevalence. In particular it would have been easy for them to have conducted validity checks using the percent of suicides committed with guns (PSG) as a criterion measures, since they were aware of the measure (Karp, 2018, p. 6) and the requisite data are easily available for over 190 countries. They did not perform this or any other validity check.

Indeed, SAS staff seem to go out of their way to persuade readers that PSG itself is not valid, based entirely on the speculation that its validity would be worse for less developed nations than it has been shown to be for developed nations (Karp, 2018, p. 6). Their sole support for this dubious conjecture was a single article (Adjaccio-Gross et al. 2006) which actually did not say a thing about less developed nations.

At times, SAS staff appear to have misunderstood just what this proxy measure is, at one time describing it as "firearm suicide" (SAS 2007, p. 54). PSG is neither the count of firearm suicides nor the per capita rate of firearm suicides. The idea that SAS staff wrongly believed that the suicide-related proxy is one of these two measures is supported by their stated rationale for rejecting its use. They assert that it "is especially weak in societies where suicide is anathema and routinely concealed, disguised as a natural death or an accident, or just not reported at all"

(SAS 2007, p. 54). This would indeed weaken confidence in simple counts or per capita rates of total suicides or firearm suicides, in countries where suicide is "routinely concealed." There is no reason, however, to believe that the stigma of suicide applies more (or less) to firearms suicides than to nonfirearms suicides. The validity of PSG does not depend on an assumption that all suicides, or all firearm suicides, are reported. It only relies on the assumption that any underreporting that may afflict suicide statistics applies to both gun suicides and nongun suicides to roughly equal degree. Thus, the SAS speculation is irrelevant to the validity of PSG – accurately understood – as a proxy for gun levels.

A Test of Validity

A validity test is therefore offered here. The concurrent validity of the SAS-estimated civilian guns/100 population is measured by computing its correlation with the percent of suicide committed with guns (PSG). PSG is used as a criterion because it has previously been shown to have a near-perfect correlation (r=.95) with direct survey measures of the percent of households reporting gun ownership generated by the International Crime Survey (Kleck, 2004, p. 18). Thus, using PSG as a criterion of validity is virtually equivalent to using direct survey measures of household gun prevalence. Direct survey measures themselves, however, are available only for a small (less that one fourth of all countries) unrepresentative subset of most highly developed nations. SAS personnel report that there are survey data for only 56 "countries/territories," a total that may include internal components of countries, such as Scotland, Wales, or Northern Ireland – Karp, 2018, p. 8).

National mortality data on firearms suicides and total suicides for computing PSG, derived from the World Health Organization, are available for 194 nations (Our World in Data, 2020; Global Health Data Exchange, 2020). The percent of suicides committed with firearms

was computed by dividing the age-adjusted rate of *firearms* suicide by the age-adjusted rate of *total* suicide. SAS gun rates are available for all but one of these countries (Small Arms Survey 2020). Thus, the validity test is based on all 193 nations for which data on both SAS civilian firearms per 100 people and PSG measures are available. Both the SAS and PSG data pertain to 2017, the latest year available. SPSS statistical software was used for computations.

Results

The Pearson product-moment correlation between the SAS estimate of civilian firearms per 100 persons and the PSG validity criterion was computed to be r=.586. Since $r^2 = .34$, the proportional reduction in error interpretation is that knowing the SAS gun rate reduces error in predicting PSG by just 34%. Put another way, 66% of the variation in PSG is <u>not</u> shared with variation in the SAS gun rate. Thus, the SAS measure appears to be a very poor measure of the civilian availability of firearms for nations.

Conclusions

The easy availability of the SAS data, and its extensive coverage of over 190 nations, is likely to tempt unwary researchers into using the SAS civilian gun ownership estimates for studies of the relationship of gun ownership rates and violence rates. Indeed, at least one set of researchers has already done so (Konty and Schaefer 2012). Nevertheless, these estimates should not be used for cross-national research on the relationship between gun levels and violence or crime rates. SAS estimates are neither strongly valid nor comparable across nations, so any findings based on their use would be meaningless because the estimates do not measure the relative availability of firearms across nations.

Instead, a better proxy for national prevalence of civilian firearms possession in cross-national studies of homicide or crime is PSG. (PSG cannot be used as a predictor of suicide rates

because of the common components problem – National Research Council 2005, pp. 168-170). PSG has been validated, in that it has been shown to be almost perfectly correlated with direct survey measures of household firearms prevalence, while SAS has not. Further, PSG is measured in the same way in all nations, and thus is more comparable than the SAS across countries. Finally, the data for computing PSG are available for over 190 countries (Global Health Data Exchange 2020).

Supplementary cross-national research can also be conducted using direct survey measures of gun ownership, albeit only on a few dozen countries. Those using this measure would have to confine their conclusions to highly developed, mostly European, nations because such survey data are available almost exclusively for such countries. Further, surveys probably reflect mostly legal gun ownership among mostly noncriminal people, so interpretations of research results would need to properly reflect this limitation.

Neither PSG nor (most) surveys can tell us how *many* guns a country has. Instead, these sources are useful for measuring relative national levels of gun *prevalence* – the fraction of the population with immediate access to firearms. When one considers the purposes to which gun data might be put, however, this is probably a minor limitation. If one seeks to assess the links between gun availability and violence, it is prevalence of guns that is relevant. Notwithstanding rare mass shootings, virtually all violent gun crimes are committed with a single gun. It only takes one gun to commit a homicide, suicide, assault, or robbery, and it matters little how many additional, surplus guns the violent actor might own but not use. Consequently, the SAS futile effort to estimate numbers of firearms probably has little utility for studies of violent crime.

References

Ajdacic-Gross, Vladeta, et al. 2006. "Changing times: a longitudinal analysis of international firearms suicide data." American Journal of Public Health 96:1752–55.

Global Health Data Exchange. 2020. International mortality data website at http://ghdx.healthdata.org/gbd-results-tool. Accessed 9-13-20.

Karp, Aaron. 2018. Estimating Global Civilian-held Firearms Numbers. Briefing Paper. Small Arms Survey. Available at http://www.smallarmssurvey.org/fileadmin/docs/T-Briefing-Papers/SAS-BP-Civilian-Firearms-Numbers.pdf.

Kleck, Gary. 2004. "Measures of gun ownership levels for macro-level crime and violence research." Journal of Research in Crime and Delinquency 41:3-36.

Konty, Mark, and Brian Schaefer. 2012. "Small arms mortality: access to firearms and lethal violence." Sociological Spectrum 32:475-490.

National Research Council. 2005. Firearms and Violence: A Critical Review. Washington, D.C.: National Academics Press.

Our World in Data. 2020. Website with suicide data from the World Health Organization, at https://ourworldindata.org/grapher/suicide-rate-by-firearm?tab=table. Accessed 9-11-20.

Small Arms Survey. 2007. The Small Arms Survey – 2007: Guns and the City. "Chapter 2, Completing the count: civilian firearms." Available at http://www.smallarmssurvey.org/publications/by-type/yearbook/small-arms-survey-2007.html.

Small Arms Survey. 2011. Estimating Civilian Owned Firearms. Available online at http://www.smallarmssurvey.org/fileadmin/docs/H-Research_Notes/SAS-Research-Note-9.pdf.

Small Arms Survey. 2020. Civilian Firearms Holdings, 2017. Available online at http://www.smallarmssurvey.org/fileadmin/docs/Weapons_and_Markets/Tools/Firearms_holdings/SAS-BP-Civilian-held-firearms-annexe.pdf.