DISCUSSION PAPER SERIES

No. 5357

**GUN PREVALENCE, HOMICIDE
RATES AND CAUSALITY:
A GMM APPROACH TO
ENDOGENEITY BIAS**

Tomislav Kovandzic,
Mark E Schaffer and Gary Kleck

PUBLIC POLICY

CEPR

Centre for Economic Policy Research

www.cepr.org

GUN PREVALENCE, HOMICIDE RATES AND CAUSALITY: A GMM APPROACH TO ENDOGENEITY BIAS

Tomislav Kovandzic, University of Alabama at Birmingham
Mark E Schaffer, Heriot-Watt University and CEPR
Gary Kleck, Florida State University

ABSTRACT

Gun Prevalence, Homicide Rates and Causality: A GMM Approach to Endogeneity Bias*

The positive correlation between gun prevalence and homicide rates has been widely documented. But does this correlation reflect a causal relationship? This study seeks to answer the question of whether more guns cause more homicide, and unlike nearly all previous such studies, we properly account for the endogeneity of gun ownership levels. We discuss the three main sources of endogeneity bias – reverse causality (higher crime rates lead people to acquire guns for self-protection), mis-measurement of gun levels, and omitted/confounding variables – and show how the Generalized Method of Moments (GMM) can provide an empirical researcher with both a clear modeling framework and a set of estimation and specification testing procedures that can address these problems. A county level cross-sectional analysis was performed using data on every US county with a population of at least 25,000 in 1990; the sample covers over 90% of the US population in that year. Gun ownership levels were measured using the percent of suicides committed with guns, which recent research indicates is the best measure of gun levels for cross-sectional research. We apply our procedures to these data, and find strong evidence of the existence of endogeneity problems. When the problem is ignored, gun levels are associated with higher rates of gun homicide; when the problem is addressed, this association disappears or reverses. Our results indicate that gun prevalence has no significant net positive effect on homicide rates: ceteris paribus, more guns do not mean more homicide.

Tomislav V. Kovandzic
Department of Justice Sciences
University of Alabama at Birmingham
UBOB 210
1530 3rd Avenue South
Birmingham, Alabama 35294-4562
USA
Email: tkovan@uab.edu

Mark E Schaffer
Centre for Economic Reform
and Transformation, Dept of
Economics
Heriot-Watt University
Riccarton
EDINBURGH
EH14 4AS
Tel: (44 131) 451 3494
Fax: (44 131) 451 3294
Email: m.e.schaffer@hw.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=163631

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=110682

Gary Kleck
College of Criminology& Criminal
Justice
Florida State University
Tallahassee
Florida 32306-1127
USA
Email: gkleck@mailer.fsu.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=163630

1.   Introduction

As is well known, guns are heavily involved in violence in America, especially homicide. In 2002, 63.4 percent of homicides were committed by criminals armed with guns (U.S. Federal Bureau of Investigation 2002, p. 23). Probably an additional 100,000 to 150,000 individuals were medically treated for nonfatal gunshot wounds (Kleck, 1997, p. 5; Annest et al., 1995). Further, relative to other industrialized nations, the United States has higher rates of violent crime, both fatal and nonfatal, a larger private civilian gun stock (about 90 guns of all types for every 100 Americans), and a higher fraction of its violent acts committed with guns (Killias, 1993; Kleck, 1997, p. 64). These simple facts have led many to the logical conclusion that America's high rate of gun ownership must be at least partially responsible for the nation's high rates of violence, or at least its high homicide rate (e.g., Sloan et al., 1990; Killias, 1993; Zimring and Hawkins, 1999).[1] This belief in a causal effect of gun levels on violence rates, and not merely on criminals' choice of weaponry, has likewise inclined some to conclude that limiting the availability of guns would substantially reduce violent crime, especially the homicide rate (e.g., Clarke and Mayhew, 1988, p. 106).

While there is a considerable body of individual-level research relevant to these questions (summarized in Kleck, 1997), macro-level research on the possible links between gun availability and homicide is also essential to assessing these assumptions. This is true partly because it is obviously useful to have multiple approaches to testing a given hypothesis. Perhaps more importantly, macro-level analysis enables estimation of the net effects of community gun availability on homicide rates. While gun possession among aggressors in violent incidents may serve to increase the probability of a victim's death, gun possession among victims may reduce their chances of injury or death. Individual-level research (e.g. Kleck and McElrath, 1990; Kleck and DeLone, 1993; Tark and Kleck 2004) can assess such effects of gun use in crime incidents, but it is less useful for detecting deterrent effects of gun ownership among prospective victims. Because criminals usually cannot visually distinguish people carrying

---

[1] Detailed studies using cross-national data are, however, generally unsupportive of this conclusion, and suggest instead that there is no significant association between national gun ownership rates and rates of homicide, suicide, robbery, or assault (Kleck, 1997, p. 254; Killias, van Kesteren, and Rindlisbacher, 2001, pp. 436, 440).

concealed weapons from other people, or residences with gun-owning occupants from other residences, deterrent effects would not be limited to gun owners, and might not even differ between owners and nonowners (Kleck, 1988; Kleck and Kates, 2001, pp. 153-154; Lott, 2000). Because the protective effects of gun ownership may spill over to nongun owners, the aggregate net impact of homicide-increasing and homicide-decreasing effects of gun availability can be quantified only through macro-level research.

Such macro-level studies must, however, take account of a number of potential pitfalls. The most important of these are reverse causality in the guns-crime relationship, errors in and validation of measures of gun prevalence, and omitted and confounding variables.

First, gun levels may affect crime rates, but higher crime rates may also increase gun levels, by stimulating people to acquire guns, especially handguns, for self-protection. At least ten macro-level studies have found effects of crime rates on gun levels (Kleck, 1979; Bordua and Lizotte, 1979; Clotfelter, 1981; McDowall and Loftin, 1983; Kleck, 1984; Magaddino and Medoff, 1984; Kleck and Patterson, 1993; Southwick, 1997; Duggan, 2001; Rice and Hemley, 2002), and individual-level survey evidence (not afflicted by simultaneity problems) directly indicates that people buy guns in response to higher crime rates (summarized in Kleck, 1997, pp. 74-79). Alternatively, higher violent crime rates, especially gun crime rates, could discourage some people from owning guns, by reminding them of the dangers of guns.

Thus, causality in the guns-crime relationship may run in either or both directions. If such a simultaneous relationship exists, but analysts fail to take account of it using appropriate methods, their results will be almost meaningless. What is asserted to be the impact of gun levels on crime rates will in fact also include the impact of crime rates on gun levels. Indeed, in their estimations the crime=>guns relationship could quantitatively dominate the guns=>crime relationship, in which case the analysts will misinterpret an effect of crime on gun levels as an effect of gun levels on crime.

Second, direct measurement of gun levels is subject to well-documented problems (Kleck, 2004), and many researchers have responded by using a diverse set of proxy measures. At the most basic level, researchers must either validate their chosen proxy against other measures – i.e., establish that it is correlated with other measures of gun levels – or rely on the validation investigations of others. Even a

valid proxy will, however, still suffer from measurement error. Measurement error can lead to biased estimates of the impact of gun levels on crime. Again, unless appropriate variables and methods are used, the analyst may commit either Type I or Type II errors when testing whether gun levels have an impact on crime rates.

Third, analysts must be aware of and take measures to accommodate possible confounding variables. Omitted variable bias is a particular problem for macro-level studies of the guns-crime relationship. Omission of confounding variables that are known to be correlated with both gun prevalence and crime rates (e.g., poverty and unemployment) will contaminate any estimate of the impact of gun levels on crime. Investigators need at least to include an appropriate range of controls, and ideally should adopt estimation methods that address the problem of confounding variables that are unobservable (e.g., "social capital").

This study seeks to answer the question of whether there is a causal effect of gun levels on violence rates. Unlike nearly all previous such studies, we properly account for the endogeneity of gun ownership levels. We set out a formal analysis of the main sources of endogeneity bias, and discuss how an empirical researcher can address these problems using estimation and specification testing procedures in a Generalized Method of Moments (GMM) framework for a linear model. A county-level cross-sectional analysis was performed using data on every U.S. county with a population of at least 25,000 in 1990. Gun ownership levels were measured using the percent of suicides committed with guns, which recent research indicates is the best measure of gun levels for cross-sectional research. Our estimation techniques allow us to address the problems of reverse causality, measurement error and unobservable confounding variables, and we include a wide range of controls in our estimations.

The paper is organized as follows. Section 2 formalizes the three sets of problems discussed above – reverse causality, mismeasurement of gun ownership levels, and omitted/confounding variables – using a simple modeling framework, and discusses how to address these problems. Section 3 critically reviews the macro-level research on the gun-homicide relationship in light of these three sets of problems. Section 4 sets out a GMM-based estimation strategy and describes the tests to be used, and Section 5

discusses the data and the specific estimation strategy used in this study. The results are presented in Section 6, and Section 7 concludes.

2.   Reverse Causality, Mismeasurement, and Confounding Variables

Consider a researcher who wants to estimate the impact of gun availability on the homicide rate. The researcher has available cross-sectional data on localities (we consider the potential alternative of longitudinal data in the next section). The researcher estimates the following simple linear model using ordinary least squares,

$$hom_i = \quad \beta_0 guns_i \quad + \beta_1 control_i \quad\quad\quad\quad\quad\quad\quad\quad + u_i \quad\quad\quad (1)$$

where, in self-evident notation, hom_i is the homicide rate in locality i, $guns_i$ is the level of gun ownership, $control_i$ is a variable that controls for some characteristic of locality i, u_i is the error term for the homicide equation, and for expositional convenience the constant term is suppressed. The key parameter of interest to the researcher in equation (1) is β_0, the impact of gun levels on the homicide rate. What are the potential pitfalls facing this estimation strategy?

Reverse causality

Say, for the moment, that equation (1) is well-specified, but also that it captures only part of the picture – causality in the guns-crime relationship also runs from crime to guns. In equation (2),

$$guns_i = \quad \gamma_0 hom_i \quad\quad\quad\quad\quad\quad + \gamma_3 X_i \quad\quad\quad\quad + u_{gi} \quad\quad\quad (2)$$

gun levels are influenced both by homicide rates – we expect $\gamma_0 > 0$, i.e., people buy guns in response to higher crime rates – and by some other covariate X.

If the researcher estimates (1) by OLS, but the true set of relationships is captured by (1) and (2) together, then the estimated $\hat{\beta}_0$ will not be "consistent" – it will suffer from "endogeneity" or "simultaneity" bias.[2] The reason is that the regressor *guns* is itself endogenous in a system of

---

[2] The term "bias" is used here as a shorthand for "asymptotic bias", i.e., the difference between the probability limit (as the sample size goes to infinity) of an estimator and the true value of the parameter. An estimator is "consistent" if its asymptotic bias is zero. IV/GMM estimators are in general unbiased only asymptotically. See e.g., Hayashi (2000), pp. 94-5 and chapter 3.

simultaneous equations, making it correlated with the error term u_i in (1). In this case, $\hat{\beta}_0$ will be biased

upwards by the positive γ_0. Indeed, if the reverse causality is strong enough, i.e., γ_0 is large relative to β_0,

the researcher could find that $\hat{\beta}_0 > 0$ and conclude that more guns means more crime even if the true

impact of guns on crime is negligible or negative.

The standard answer to this problem is to estimate equation (1) using the method of instrumental

variables (IV) or the more modern framework of GMM. This requires the researcher to have a variable

that is correlated with *guns* ("instrument relevance") and that is also uncorrelated with the error term u_i in

the homicide equation ("instrument validity"). The covariate X is potentially such a variable because it

appears in equation (2) as a determinant of gun levels, but is excluded from equation (1) (hence the term

"exclusion restriction"). Note that even if equation (2) is misspecified and itself suffers from endogeneity

or other problems, the researcher can still obtain consistent estimates of the parameters of equation (1) so

long as X satisfies these requirements.

Measurement and mismeasurement of gun levels

Let us maintain the assumption that equation (1) is a well-specified description of the impact of

guns on homicide rates. Say also that there is no reverse causality issue – homicide levels have no impact

on the propensity of people in the locality to acquire guns, i.e., $\gamma_0 = 0$. However, the level of gun ownership

cannot be measured exactly, and the researcher must make use of a proxy. Instead of estimating

equation (1), the researcher estimates equation (1a),

$$hom_i = \text{b}_0 prox_i + \text{b}_1 control_i + u_i \tag{1a}$$

again by OLS. The (unobservable) relationship between *guns* and *prox* is given by equation (3):

$$prox_i = \delta_0 guns_i + u_{pi} \tag{3}$$

where the term u_{pi} is the measurement error that degrades the proxy.

The consequences of measurement error for the OLS estimate $\hat{\text{b}}_0$ depend on the nature of u_{pi}. If it

is a textbook case of purely random measurement error, then the estimated relationship between the gun

proxy and homicide will be biased towards zero. It is also possible, however, that the proxy for guns is

itself endogenous, via its direct dependence on an endogenous *guns* or some other route that generates a correlation between *prox* and u_i.[3] If so, the researcher faces an endogeneity bias problem as well, and the net bias on the OLS estimate $\hat{b}_0$ can be positive or negative.

The researcher must in any case bear in mind that the coefficient b_0 is *not* the quantitative impact of gun levels on crime rates; this is given by $b_0\delta_0$, and of course δ_0 is typically not observed directly. A test of the estimated $\hat{b}_0$ may enable the researcher to say if there is a statistically significant non-zero impact of guns on crime; but without an estimate of δ_0, the researcher will be unable to say anything about the practical significance of the impact.

The usual approach when employing a proxy subject to measurement error is threefold. First, the researcher needs to validate the proxy against other, albeit imperfect, measures of gun availability. Second, estimating equation (1a) using IV/GMM techniques will generate a consistent estimate $\hat{b}_0$, assuming that instruments for guns are available that satisfy the conditions of relevance and validity. Third, if the estimated coefficient on the guns proxy $\hat{b}_0$ is significantly different from zero, the researcher will have a qualitative estimate of the impact of guns on crime; but s/he must also have some idea of the magnitude of δ_0 (e.g., from a validation exercise) in order to obtain a quantitative estimate of the impact.

Omitted/confounding variables

Now let us abandon the assumption that equation (1) is a well-specified description of the impact of guns on homicide rates. The true relationship is one in which there is an additional characteristic of localities that determines of homicide rates, *confound*, as shown in equation (1b):

$$hom_i = \beta_0 guns_i + \beta_1 control_i + \beta_2 confound_i + u_i \qquad (1b)$$

Say again that there is no reverse causality issue. Gun levels are, however, also influenced by the variable *confound*, as in equation (2a):

$$guns_i = \gamma_2 confound_i + \gamma_3 X_i + u_{gi} \qquad (2b)$$

---

[3] For example, omitted variable bias as discussed below could also operate via a confounder that is omitted from the homicide and proxy equations.

If the researcher estimates the original equation (1) using OLS, the estimated $\hat{\beta}_0$ will again be biased. This time, the endogeneity bias is an omitted variable bias. The absence of *confound* from the estimated homicide equation (1) and its role as a determinant of *guns* in equation (2b) means that *guns* will be correlated with the error term in (1). In other words, omitting *confound* from the homicide equation makes *guns* an endogenous regressor.

How to address this problem depends on whether the omitted variable is observable. If *confound* is a variable that is available to the researcher, then equation (1b) may be estimated using OLS; the researcher simply includes it alongside *control* as a second control variable. Often, however, confounding variables are unobservable (e.g., pro-violence subcultural norms or social capital) and hence cannot be included as explicit regressors. In this case, the standard approach would be the same as that for the reverse causality problem: estimate (1) using IV/GMM, with the covariate *X* as the excluded instrument.

Specification testing

In all three sets of problems discussed above – reverse causality, measurement error, and omitted/confounding variables – IV/GMM methods can, in principle, enable the researcher to avoid the biases that would contaminate the OLS estimate of the impact of guns levels on homicide rates. Given the mass of empirical evidence and theoretical considerations cited above, the natural starting point of the researcher should be that these biases are likely to be present and that IV or GMM estimation is the appropriate technique.

For the results to have any credibility, the IV specification should be both plausible and subject to rigorous testing. The *ex ante* exclusion restrictions that identify the model must be consistent with prior theory and evidence – here, the exclusion of the covariate *X* from the homicide equation and its presumed correlation with *guns* or *prox*. For example, in a model of burglary rates, the percent of an area's population that resides in rural areas should not be omitted if prior research and theory indicates that this variable influences burglary rates (e.g., Cook and Ludwig 2003, pp. 106-107). The *ex post* specification

testing should include tests of both instrument relevance and, if the model is overidentified,[4] instrument validity; the latter test is sometimes called a test of overidentifying restrictions.

Although the presumption of the researcher should be that OLS estimates are likely to be biased, the possibility that these biases are small or negligible cannot be ruled out. If this is indeed the case, then *guns* (or *prox*) can be treated as an exogenous regressor, and estimation by OLS would be preferred to IV because it is the more efficient (lower variance) estimator. The standard approach to this question is to conduct a test of the endogeneity of *guns*. Such a test relies implicitly on a comparison of an estimation in which *guns* is treated as exogenous and one in which it is treated as endogenous. For the test to have any meaning, it is therefore essential that the OLS estimation be contrasted with a well-specified IV estimation.

Summary

The discussion above indicates that macro-level investigations of the guns-homicide relationship need to take account of a range of potential pitfalls. Both the empirical evidence and standard practice suggest that the starting point should be the presumption of possible reverse causality, measurement error, and omitted variables. We can summarize the basic requirements and procedures for addressing these pitfalls as follows:

1. The researcher requires a proxy for gun levels that has been properly validated against other available measures.

2. One or more instruments for guns is required. Any such instrument needs to satisfy the two requirements of instrument validity and instrument relevance.

3. Instrument validity means an instrument should be exogenous in the econometric sense of the term, i.e., uncorrelated with the error term in the homicide equation. Prior reasoning should also suggest that the variable has no direct impact on homicide rates, i.e., that the variable is properly excluded from the homicide equation. If the researcher has more than one instrument available and the model is overidentified, then the validity of the instruments can and should be tested and the results reported.

4. Instrument relevance means an instrument should be plausibly correlated with gun levels. The relevance should be tested and results reported.

---

[4] I.e., the number of excluded instruments exceeds the number of endogenous regressors.

5. If gun levels are to be considered exogenous, such a specification should be supported by a properly conducted endogeneity test that uses a comparison with a well-specified IV/GMM estimation, i.e., one that uses instruments for guns that are both relevant and valid. Testing for the endogeneity of guns by comparing OLS to a misspecified IV estimation cannot provide evidence that OLS is acceptable.

6. The homicide equation should include a reasonably full set of control variables, so as to reduce the problem of omitted variables and confounding factors.

In the next section, we use these criteria to review the macro-level studies that investigate the guns-homicide relationship.

3.   Prior Research

Table 1 summarizes macro-level studies of the impact of gun ownership levels on crime rates. Many of these studies found a significant positive association between crime or violence rates and some measure of gun ownership, but all of these studies share at least one, and usually several, of the methodological problems we have outlined (see Kleck, 1997, Chapter 7 for a more extensive review of the pre-1997 research). The studies are also commonly characterized by small samples and other problems, but it is the flaws in measurement and modeling that are most clearly consequential.

Invalid measurement of levels and changes in gun ownership

Table 1 (see note b) shows that past macro-level guns-violence studies have used a large and diverse set of proxies for gun levels. With few exceptions (e.g., Cook, 1979; Kleck and Patterson, 1993), researchers using these measures failed to validate them using any criterion, such as establishing that they correlate well with more direct survey measures of gun prevalence. The validity of over two dozen of these measures has recently been systematically assessed by measuring correlations between the proxies and direct survey measures (Kleck, 2004). The results indicate that most of the measures used in prior cross-sectional research and *all* of those used in time-series or pooled cross-section studies have poor validity. The best indicator of gun levels was the percent of suicides committed with guns (PSG), which correlated strongly (with correlations ranging between 0.85 to 0.95) with direct survey measures across cities, states, and nations.

Despite its excellence as an indicator of cross-sectional variation in gun levels, PSG has no validity whatsoever as a cross-temporal indicator. Kleck (2004, p. 24) demonstrated that not only do changes in PSG fail to strongly correlate with direct survey measures of changes in gun prevalence over time; PSG actually generally has weak negative correlations with these criterion measures. Further, Kleck found that *none* of the measures used in past research or any of the rest of the two dozen indicators he assessed were valid measures of cross-temporal variation in gun levels (pp. 20-21). The implication for the potential alternative approach of longitudinal analysis is clear: unless some new, heretofore unknown proxies are developed that are valid indicators of gun trends, meaningful longitudinal analysis of the impact of gun levels on violence rates is impossible.

The recent study by Moody and Marvell (2005) illustrates this. The authors generated a state-level panel dataset covering 1977-98 by combining direct survey measures of gun levels taken from the General Social Surveys (GSS) with imputed values based on PSG. Their analysis used both IV methods and Granger causality tests to detect whether changes in gun levels led to changes in violence rates or visa-versa. This use of direct survey measures of gun levels fails, however, because (a) some states contribute few or no respondents to the GSS sample in a given year, (b) the GSS asked gun ownership questions in only 17 of the 22 years analyzed, and (c) since only about 1,400 people are asked the gun questions in a typical year (not 3,000, as the authors report), GSS samples for any one state in any one year average only about 29. As a result, random sampling and response error could easily account for virtually all observed cross-temporal variation in the state gun prevalence figures; indeed, it is unlikely that any of the year-to-year changes in state gun prevalence are statistically meaningful.

This becomes apparent if we examine and extend the validation exercise the authors use to justify imputing missing state-level GSS data on percent of households with a gun (HHG) or handgun (HHGG) for a given year using state-level data on PSG;[5] the exercise is crucial to their study since imputation of missing HHG and HHGG observations effectively doubles their estimation sample. The authors report that in their state-level panel, PSG is highly correlated with both HHG and HHGG, and much more

---

[5] The data were downloaded from http://cemood.people.wm.edu/research.html.

correlated than alternative proxies such as gun magazine subscriptions. What the authors fail to note is that this correlation is driven entirely by the cross-sectional correlation between PSG and the GSS gun measures; the cross-temporal correlations are tiny. This can be easily seen by examining separately the cross-sectional and cross-temporal correlations of HHG and PSG using their data. The left-hand panel of Figure 1 shows that state-average PSG and state-average HHG are highly correlated – the correlation coefficient is 0.83 and statistically highly significant. The right-hand panel, on the other hand, shows that the correlation of annual changes (first-differences) in state PSG and state HHG is essentially nil – the correlation coefficient is 0.06 and statistically insignificant. Note also that the very small GSS state sample sizes mean that most of the annual changes in GSS state-level gun prevalence are implausibly large (the standard deviation is 18). Since the authors' analysis is based entirely on cross-temporal variation in HHG/HHGG[6] with missing values imputed from the cross-temporal variation in PSG, their null findings on the guns/crime link are not surprising. The authors were essentially modeling noise in the gun data, and their analysis says little about the merits of the guns-cause-crime hypothesis.[7]

Figure 1: PSG and General Social Survey HHG, State-level Data 1977-98

---
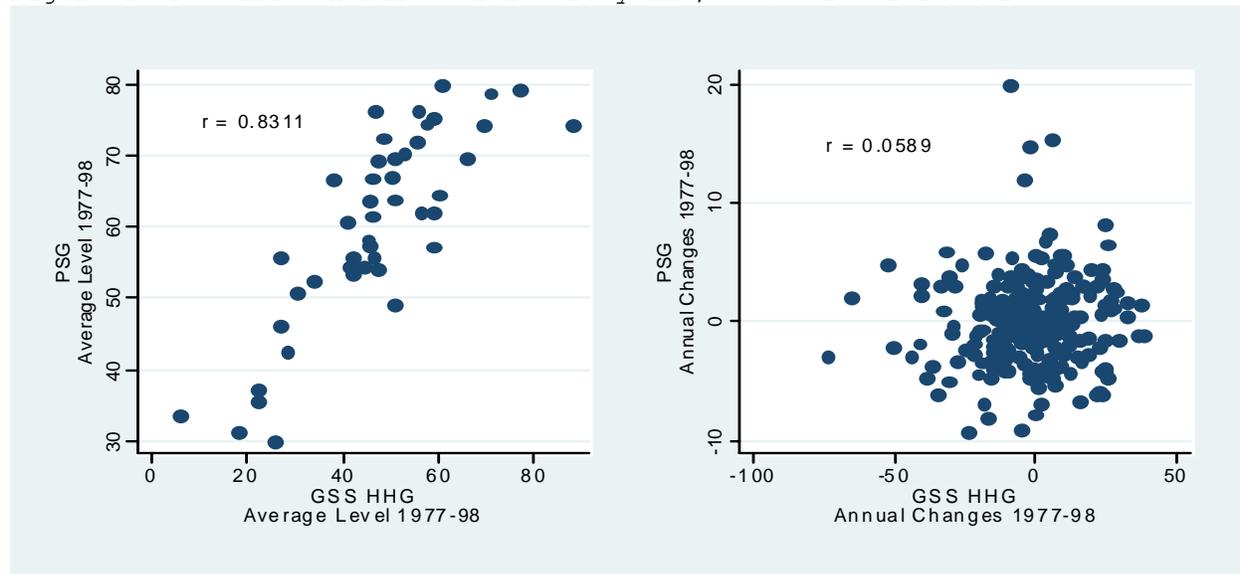
[6] The estimations are either in first-differences or in levels with fixed effects; in the latter case, the fixed effects absorb all the between-state variation and leave only the within-state (cross-temporal) variation in crime levels to be explained by the corresponding variation in guns.
[7] The authors try in the paper to address the issue of attenuation bias with a calibration exercise, but they underestimate the scale of the bias by calibrating to the wrong gun measure (gun levels, instead of changes in gun levels as used in their regressions.)

In sum, with the exception of the few studies that used PSG, indexes including PSG, or direct survey measures in cross-sectional research (e.g., Cook, 1979; Kleck and Patterson, 1993), the supposed gun-crime associations estimated in nearly all past research are uninterpretable on the simple grounds that gun levels were not actually measured.

Many measures have flaws that go beyond merely being imperfect indicators of gun levels. Measures such as the percent of homicides (or robberies, or aggravated assaults) committed with guns are vulnerable to the possibility of artifactual associations with crime rates. For example, Hemenway and Miller (2000) used Cook's (1979) "gun density" index, which is the average of (1) the percent of homicides committed with guns, and (2) PSG. The reason for the significant associations found between the Cook measure and homicide rates across 26 nations (and the absence of such when just PSG was used), is that both national homicide rates ([gun homicides + nongun homicides]/population) and the percent of homicides committed with guns ([gun homicides/total homicides] x 100%) contain a common component in their numerators: the number of gun homicides. In other words, the "gun density" index is endogenous: when used as a proxy for gun levels, u_p, the error term that degrades the proxy in equation (3), is correlated by construction with the homicide rate *hom*, and hence with u, the error term in the homicide equation (1). Had the authors employed instrumental variables or a related technique, they might, in principle, have been able to obtain unbiased estimates. But of course the better approach, regardless of estimation technique employed, is to use a proxy that is not biased by construction.

The "percent of crimes with a gun" proxy has another flaw. This indicator reflects not only the availability of guns but also the preference of the criminal population for using guns (Brill, 1977, pp. 19-20). While availability certainly affects how often criminals use guns, the "lethality" of offenders, i.e., their willingness to inflict potentially lethal injury on others, affects weapon choice as well – criminals willing to use lethal weapons are also more willing to inflict lethal injury (Cook, 1982, p. 248; Wright and Rossi, 1983, pp. 189-211). Consequently, these macro-level indicators can confound gun availability with the average lethality of the criminal population, producing guns/homicide associations that are virtual tautologies, reflecting nothing more than the truism that populations that are more lethal are more likely

to commit lethal acts. Here we have an example in which the same problem can be described as either measurement error or omitted variable bias. Average lethality is an omitted and unobservable variable (*confound* in equation (1b)), and is also a component of the measurement error that degrades the proxy (u_p in equation (3)). Again, this problem can be countered using instrumental variable estimation techniques, but again, it is preferable to use a better proxy whatever the estimation technique employed.

Finally, virtually all studies using a proxy for gun levels have failed to calibrate the proxy used to a survey-based measure of gun prevalence. That is, they failed to adjust for the fact that there is not necessarily a one-to-one correspondence between the gun proxy and actual gun levels. Moody and Marvell (2003) showed that the failure of two such studies (Duggan, 2000; Cook and Ludwig, 2003) to calibrate their proxy led the authors to make claims of policy relevance that were simply unfounded.

Endogeneity and reverse causality

Most guns-violence studies do nothing whatsoever to deal with this problem (e.g., Hemenway and Miller, 2000; Killias, 1993; Miller, Azrael, and Hemenway, 2002). Table 1 shows that of 31 total studies, twenty had no research design features that would help the analyst distinguish the possible positive effects of gun levels on crime rates from the possible positive effects of crime rates on gun levels. Three studies (Southwick, 1999; Duggan, 2003; Moody and Marvell, 2005) used the weaker notion of Granger causality, in which longitudinal data are used to establish whether past gun levels help predict current crime rates, but their findings are uninterpretable for the reasons cited above: there are no useable longitudinal data on gun levels that have been shown to be reliable enough for statistical analysis.

Of the 10 studies that attempted to address the causality problem using a structural approach, eight clearly failed because they used inappropriate methods, e.g., estimating unidentified models. This means that if crime rates do influence decisions to acquire guns, the findings of all but a handful of prior studies (e.g., Kleck and Patterson, 1993) are uninterpretable on the grounds that the statistical models on which they were implicitly based were unidentified. Even the studies that did estimate identified models using appropriate techniques failed to report tests of instrument relevance and instrument validity, a problem that we might label "underreporting" of structural causality methods.

Hoskin (2001), for example, estimated a simultaneous equations model of homicide rates and gun availability across 36 nations, but the model was almost certainly underidentified. The exclusion restrictions used to identify the model were arbitrary and implausible (indeed, he never made them explicit), and directly contradicted the author's own theoretical assertions.[8] Nor did Hoskin report any tests of instrument validity that would indicate their adequacy, or provide a discussion of the requirement of instrument relevance.

The Stolzenberg and D'Alessio (2000) study is another example of the underreporting problem. They used an appropriate test for endogeneity, the Hausman test, to support a specification of a crime equation in which gun levels is treated as exogenous, but the meaning of their test is doubtful (p. 1475). The test's utility depends crucially on the specification of the IV estimation in which gun levels are instrumented, and these authors did not report what instrumental variables they used, let alone whether they passed tests of validity and relevance. As a result, it is impossible to place any confidence in their Hausman test results, and therefore in the conclusions they draw from their estimated crime equation.

Omitted/confounding variables

Most studies also use minimal or no controls for possible confounding variables. As a point of reference, Kleck and Patterson (1993, pp. 259-260) controlled for as many as 36 potential confounders, beyond their gun level measure and 19 dummy variables representing gun laws. In contrast, Cook and Ludwig's (2003) county-level instrumental variable analysis included just three control variables (beyond county and year dummies), Hoskin (2001, p. 586) included just three controls in his homicide equation, and Hemenway and Miller (2000) did not control for a single confounding variable in their homicide models.

Due to uncertainty about exactly which macro-level attributes of places affect crime rates, it is impossible to know exactly which variables in a guns-violence study might be confounding variables, i.e.,

---

[8] To achieve identification in his homicide equation, Hoskin excluded (1) population density, (2) the percent of the population that was male and aged 15 to 34 (p. 584), and (3) an East Asia dummy. Yet just a few pages earlier he had asserted, quite plausibly, that the first two of these variables should affect homicide rates, and his discussion of the third is limited to the remark that both homicide and firearms ownership rates are low in East Asia (pp. 580-1).

factors that affect crime rates but are also correlated with gun levels or gun control laws. That is, it is uncertain which variables might either generate spurious associations among these variables or suppress or distort genuine causal effects. The consequences of failing to control confounders can be quite serious – biased parameter estimates – while the consequences of wrongly including irrelevant variables are more mild – somewhat inflated standard errors of coefficients. Thus, the most sensible procedure is to control for as many likely and observable crime determinants as is reasonable, within the limits imposed by sample size and assuming that multicollinearity does not preclude doing so, and to employ IV/GMM techniques in order to address the problem of unobservable and hence omitted confounders.

4.   IV/GMM, Instrument Validity, and Instrument Relevance

In this section we outline an estimation strategy using modern econometric techniques and GMM methods in particular. GMM can be applied to nonlinear as well as linear problems; we use a linear model for simplicity of exposition as well as ease of implementation. We begin by illustrating the two requirements of instrument validity and instrument relevance in the simplest version of model (1). There is a single explanatory variable, *guns*,

$$hom_i = \quad \beta_0 guns_i \hspace{6cm} + u_i \hspace{2cm} (1c)$$

and a single excluded instrument, X. The standard IV estimator of β_0 is $\hat{\beta}_0 = \sum X_i hom_i \Big/ \sum X_i guns_i$.

The proof that $\hat{\beta}_0$ is a consistent estimator of β_0 is straightforward:

$$\text{plim}\hat{\beta}_0 = \text{plim}(\sum X_i hom_i \Big/ \sum X_i guns_i) \hspace{2cm} = \text{plim}(\sum X_i (\beta_0 guns_i + u_i) \Big/ \sum X_i guns_i)$$

$= \beta_0 + \text{plim}(\sum X_i u_i \Big/ \sum X_i guns_i) = \beta_0 + \text{plim}(\frac{1}{N}\sum X_i u_i) \Big/ \text{plim}(\frac{1}{N}\sum X_i guns_i)$. This will be equal to β_0 if the second term is zero, which will be the case if X satisfies two conditions. First, if $E(X_i u_i)=0$ (X is uncorrelated with the error term; validity), then $\text{plim}\frac{1}{N}\sum X_i u_i = 0$. Second, if $E(X_i guns_i)\neq0$ (X is correlated with *guns*; relevance), then $\text{plim}\frac{1}{N}\sum X_i guns_i \neq 0$. Both of these are "moment conditions"; they relate to statistical moments of *guns*, X and *u*.

A modern and increasingly popular approach to the problem of estimation with endogenous regressors is the Generalized Method of Moments or GMM (Hansen, 1982).[9] GMM provides a unified framework for estimation and testing that is naturally suited to empirical problems where endogeneity and instrument validity are central. The issue of instrument relevance is one that has attracted a great deal of econometric research in recent years and new findings are appearing regularly. Nevertheless, there are enough established results to provide empirical researchers with some practical guidelines for how to detect and address problems of instrument relevance.

Estimation, testing and instrument validity in the Generalized Method of Moments

As its name implies, GMM is a generalization of the method of moments (MM), a much older technique introduced by Karl Pearson in 1894. The essence of MM is straightforward. The researcher has theoretical priors which imply theoretical or population moments, i.e., characteristics of the population that are implied by the researcher's model. The researcher also has data available, and can calculate sample moments using these data. The sample moments depend not only on the data, but also on the unknown parameters of the model that the researcher wants to estimate. The researcher's estimates of the parameters are the values that make the sample moments match the assumed population moments.

We use model (1c) to illustrate MM estimation. There is a single explanatory variable, *guns*. The researcher also has a single theoretical moment condition, namely that the variable X is exogenous: it is orthogonal to, i.e., uncorrelated with, the error term u, $E(X_i u_i)=0$. This is referred to as an "orthogonality condition". Note that the researcher has not imposed any priors about the exogeneity of *guns*; in particular, by choosing $E(X_i u_i)=0$ as the single theoretical orthogonality condition, the researcher is allowing for the possibility that that *guns* may be endogenous, i.e., $E(guns_i u_i) \neq 0$. The MM estimate of the parameter β_0, is the value $\hat{\beta}_0$ that makes the sample moment corresponding to $E(X_i u_i)=0$ also equal to zero. The sample counterpart to the error term u_i is the residual, defined in the usual way as

---

[9] The literature on GMM is now vast and many good expositions are available. Wooldridge (2001) provides an easily accessible introduction to GMM and its applications. Hayashi (2000) is an advanced text that sets out many of the tests and results used and cited in this paper. Baum et al. (2003) set out the basics of IV and GMM estimation and specification testing, and describe the set of extended Stata estimation and testing routines used here.

$\hat{u}_i \equiv hom_i - \hat{\beta}_0 guns_i$. The sample moment condition is therefore that the sample mean of $(X_i \hat{u}_i)$ is zero, i.e., $\frac{1}{N}\sum X_i \hat{u}_i = 0$. This is just one equation in one unknown, i.e., the model is exactly identified. To solve the equation for $\hat{\beta}_0$ we substitute for the residual to obtain $\frac{1}{N}\sum X_i(hom_i - \hat{\beta}_0 guns) = 0$, and after simplifying and rearranging, we have the MM estimate of the impact of guns on homicide: $\hat{\beta}_0 = \sum X_i hom_i / \sum X_i guns_i$. The MM estimator in this single-regressor model allowing for possible endogeneity of *guns* is, in fact, the standard IV estimator.

The above illustrates two features of MM which carry over to GMM and which make the method attractive to empirical researchers and to users of their work. First, many commonly used estimators are special cases of GMM estimators, and the apparatus of GMM can be used with these estimators. Both IV and OLS, for example, are GMM estimators. Second, the MM/GMM approach makes very clear what the assumptions of the researcher are. Thus estimating model (1) above with two regressors using OLS is equivalent to using MM with two theoretical orthogonality conditions, namely E(*guns_i u_i*)=0 and E(*control_i u_i*)=0. These conditions must hold – both guns and the control variable must be exogenous – if consistent estimates of the parameters β_0 and β_1 are to be obtained in this way. If either theoretical orthogonality condition does not hold, then the estimates of both parameters will be inconsistent.

A natural question is, what if there are more orthogonality conditions than there are parameters to be estimated? Say that the researcher wants to estimate (1), believes that *control* (but not *guns*) is exogenous, and has two more variables, *X1* and *X2*, that s/he believes are also exogenous. (In the terminology of IV estimation, *control* is an "included instrument" and *X1* and *X2* are "excluded instruments".) This gives three orthogonality conditions, E(*control_i u_i*)=0, E(*X1_i u_i*)=0 and E(*X2_i u_i*)=0, but still only two parameters to estimate, β_0 and β_1. Because the model is overidentified – there are three sample moment equations but only two unknowns – in general it will be impossible to find estimates of β_0 and β_1 that set the sample moment conditions exactly to zero.

In GMM, estimates of β_0 and β_1 are chosen such that the three sample moment conditions are as "close" to zero as possible. More precisely, GMM proceeds by defining an objective function J(.) that is a function of the data, the parameters and a set of weights in a weighting matrix W. The GMM objective function is a quadratic form in the sample moment conditions, i.e., the sample moment conditions are

weighted using the weights in W and summed to produce a scalar that is minimized. J(.) can be thought of as the "GMM distance" – the distance from zero, which is the value the objective function would take if all the sample moment conditions were satisfied – and the definitions of the GMM estimators of β_0 and β_1 are those values that minimize J given W and the data.

There are as many different GMM estimators as there are different possible Ws to use in J, and in fact any GMM estimator using a non-trivial W will generate consistent estimates of the parameters. Where GMM comes into its own is when an optimal weighting matrix is chosen. The optimal GMM weighting matrix is the inverse of the covariance matrix of moment conditions S; in our example, the variances and covariances of ($control_i\, u_i$), ($X1_i\, u_i$) and ($X2_i\, u_i$). GMM estimation with an optimal weighting matrix has the following features. First, when S^{-1} is used as the optimal weighting matrix, the GMM estimator is efficient, i.e., it has the smallest asymptotic variance. Although the true covariance matrix S is unknown, it can be easily estimated in a prior step,[10] and the two-step GMM estimator that uses the estimated $\hat{S}^{-1}$ as the weighting matrix is also efficient. Second, more orthogonality conditions mean more efficient estimation, as long as the additional orthogonality conditions are satisfied. Third, efficient GMM readily accommodates forms of errors that are often encountered by empirical researchers in social science: heteroskedasticity, autocorrelation, and clustering (spatial correlation). Thus if the estimated $\hat{S}$ is robust to arbitrary heteroskedasticity, i.e., heteroskedasticity of unknown form, the efficient GMM estimator that uses $\hat{S}^{-1}$ as the weighting matrix will be efficient in the presence of arbitrary heteroskedasticity, and the GMM standard errors of the parameters will be robust to arbitrary heteroskedasticity. The same applies to autocorrelation and clustering. This robustness to arbitrary violations of homoskedasticity and independence is appealing to empirical researchers, not least because it means obtaining valid estimation results does not require a researcher to model these violations explicitly and correctly.

---

[10] The reason that it is easy is that an estimate of S can be obtained from any consistent estimator of the equation. The usual GMM estimator is therefore "two-step feasible efficient GMM": (1) in the first step, some consistent but possibly inefficient estimator of the parameters (e.g., IV) is used to obtain $\hat{S}$; in (2) in the second step, $\hat{S}^{-1}$ is used to minimize J and obtain the efficient GMM estimates of the parameters.

Fourth, and most importantly for this paper, efficient GMM provides a straightforward framework for testing the validity of orthogonality conditions when the equation is overidentified, i.e., when there are more orthogonality conditions than there are parameters to be estimated. Under the null hypothesis that the all the orthogonality conditions are valid – all the variables that were assumed to be exogenous are indeed exogenous – the minimized value of J is distributed as χ^2 with degrees of freedom equal to the degree of overidentification.[11] This test of overidentifying restrictions is known in the literature as the Hansen or Sargan-Hansen J statistic and is, conveniently, an automatic by-product of GMM estimation.

Note that the orthogonality condition corresponding to an excluded instrument may fail because the exclusion restriction itself is invalid. Say the model of equation (1) is misspecified in the sense that *X1* should actually be an included exogenous regressor because it has a direct impact on homicide rates. If the researcher estimates equation (1) and uses *X1* and *X2* as excluded instruments, *X1* will be correlated with the error term of (1), and the J statistic will be tend to be large, indicating a failure of orthogonality conditions. In other words, the incorrect exclusion of *X1* from equation (1) makes *X1* endogenous.

The same framework can be used to test the validity of a subset of orthogonality conditions, i.e., to test whether or not selected instruments are exogenous. Consider the J statistic resulting from two different efficient GMM estimations: J1 is the J statistic from an efficient GMM estimation that uses all the moment conditions, and J2 is the J statistic from an efficient GMM estimation that does not use the moment conditions corresponding to the suspect instruments (if the suspect variables are included instruments, i.e., regressors, in the J2 estimation they are treated as endogenous; if they are excluded instruments, in the J2 estimation they are not used at all). Under the null hypothesis that the suspect variables are valid instruments, the quantity J1-J2 is distributed as χ^2 with degrees of freedom equal to the number of instruments being tested. This test is known variously in the literature as a C test, a distance GMM test, or a difference-in-Sargan test. It is important to note that for the C test to be valid, the

---

[11] In the exactly-identified case, the minimized value of J is zero, and the orthogonality conditions cannot be tested.

orthogonality conditions corresponding to the instruments *not* being tested must also be valid, i.e., the J2 statistic (which is a test of the validity of these orthogonality conditions) should be small.

The GMM framework usefully encompasses and extends a number of older and well-known procedures. For example, if the error term u is homoskedastic and independent:[12] (a) OLS is the efficient GMM estimator when regressors are exogenous; (b) IV is the efficient GMM estimator when some regressors are endogenous; (c) if regressors are exogenous, OLS is more efficient (has a lower variance) than IV because it makes use of more orthogonality conditions;[13] (d) the J statistic for the IV estimator is numerically identical to the Sargan's (1958) NR^2 overidentification statistic;[14] (e) the Hausman-Wu test for the endogeneity of regressors is numerically identical to the C test that uses the Sargan-Hansen J statistics from IV and OLS estimations of a model. GMM provides a straightforward method to generalizing the above estimators and tests to situations where homoskedasticity and/or independence do not hold.

Useful as it is, GMM is not a panacea, and several caveats to its use should be mentioned. First, the use of large numbers of orthogonality conditions in the form of excluded instruments can generate finite sample bias problems, and the general advice here is to be parsimonious with excluded instruments. Second, there is some evidence that the standard errors for some efficient GMM estimators may be biased downwards in finite samples, i.e., testing coefficients may be prone to Type I errors. A conservative estimation strategy adopted by some researchers is consequently to use inefficient GMM estimators (e.g., OLS or IV in the presence of heteroskedasticity), rely on robust standard errors for inference, and use the efficient GMM J statistic for specification testing.[15] Third, Sargan-Hansen tests can have limited power, i.e., may be prone to Type II errors, and a J or C test that fails to reject the null should be treated with caution.

---

[12] See Hayashi (2000) for details.

[13] More precisely, when errors are homoskedastic and independent, the additional orthogonality conditions corresponding to the endogenous regressors in IV improve the efficiency of OLS, and the conditions corresponding to the excluded instruments in IV become redundant for OLS.

[14] Basmann's (1960) overidentification statistic is a close relative of, and asymptotically equivalent to, Sargan's statistic, and is also invalid when errors are not homoskedastic.

[15] Thus Wooldridge's (1995) robust overidentification statistic for IV estimation is numerically equal to the J statistic for two-step efficient GMM (Baum et al. 2003).

The GMM procedure that our hypothetical researcher would employ for testing the exogeneity of variables, and that we use below, is as follows:

1. Estimate equation (1) using the full set of four orthogonality conditions: $E(guns_i\, u_i)=0$, $E(control_i\, u_i)=0$, $E(X1_i\, u_i)=0$ and $E(X2_i\, u_i)=0$, and obtain the J statistic for the efficient GMM estimator. Allow for possible heteroskedasticity or spatial correlation if suggested by priors or evidence from the data. (The corresponding efficient GMM estimator is due to Cragg (1983) and is also known as HOLS, "heteroskedastic OLS".) If the J statistic is large, take this as evidence that one or more moment conditions is invalid, i.e., one or more of the four variables is endogenous, and proceed to Step 2. If the J statistic is small, take this as evidence that the orthgonality conditions are jointly valid – all the variables are exogenous. However, prior research suggests that the assumption that *guns* is exogenous is questionable, so proceed to Steps 2 and 3.

2. Estimate equation (1) without the assumption that *guns* is exogenous, i.e., using the three orthogonality conditions $E(control_i\, u_i)=0$, $E(X1_i\, u_i)=0$ and $E(X2_i\, u_i)=0$, and obtain the J statistic for the efficient GMM estimator. Allow for possible heteroskedasticity or spatial correlation. If the J statistic is small, take this as evidence that the moment conditions are valid – *control*, *X1*, and *X2* are all exogenous – and proceed to Step 3. If the J statistic is large, take this as evidence that one or more of the remaining orthogonality conditions is invalid, i.e., either the control variable or one of the excluded instruments is invalid, and stop – consistent estimation is not possible.

3. Test whether *guns* is endogenous using a C test using J1-J2, where J1 is the J statistic using the full set of orthogonality conditions (Step 1) and J2 is the J statistic that does not assume exogeneity for *guns*. If the C statistic is small, take this as evidence that *guns* may be exogenous along with the other variables, and proceed to Step 4. If the C statistic is large, take this as evidence that *guns* is endogenous and go to Step 5.

4. (*guns* is exogenous) *Consider* estimating the equation by efficient GMM (i.e., HOLS) or OLS. The former is consistent and efficient; the latter is consistent, but inefficient if errors are not homoskedastic and independent. *Alternatively*, because of the danger of a Type II error and because prior evidence and research suggests that gun levels may be subject to endogeneity bias from various sources, treat *guns* as endogenous and go to Step 5.

5. (*guns* is endogenous) Estimate the equation by efficient GMM or by IV. The former is consistent and efficient, but may be more prone to Type I errors; the latter is consistent, but inefficient if errors are not homoskedastic and independent.

Instrument relevance and the weak instruments problem

As the simple IV example above showed, instrument relevance is another type of moment condition required for an IV-type estimator to be consistent. The minimal relevance requirement is just the rank condition for identification of the model; if the rank condition is not satisfied, the model is unidentified. The intuition behind instrument relevance is simply that the excluded instruments must be

correlated with any endogenous regressors. If there are multiple endogenous regressors, measuring instrument relevance is not straightforward because it requires estimation of the rank of the covariance matrix of regressors and instruments.[16] In the case of a single endogenous regressor, however, the recommendation in the literature is straightforward and easy to implement: the statistic for instrument relevance is the F test of the excluded instruments in the "first-stage" regression, i.e.,

$$guns_i = \qquad \theta_1 control_i \qquad\qquad + \theta_3 X1_i \qquad + \theta_4 X2_i \qquad + v_i \qquad\qquad (4)$$

and the researcher examines θ_3 and θ_4 for their consistency with theory and prior evidence, and for their joint significance using a standard F statistic; a large value indicates that the model is identified. If heteroskedasticity or clustering is suspected, a heteroskedastic- or cluster-robust test statistic can be used. The researcher should also examine the individual significance of both θ_3 and θ_4 in the estimation of the first-stage equation (4). Adding instruments does not come without a cost; in particular, the finite sample bias of the IV/GMM estimator is increasing in the number of instruments. If the first-stage regression suggests that one instrument is strongly relevant and the other is irrelevant, the researcher should consider dropping the irrelevant instrument.[17]

An important practical problem arises if the excluded instrumental variables are correlated with the endogenous regressor but only weakly. Recent research (e.g., Bound et al., 1995; Staiger and Stock, 1997; Stock et al. 2002 provide a survey) has shown that when instruments are weak, IV/GMM estimates of parameters will be badly biased (in the same direction as OLS), estimated standard errors will be unreliable, and therefore so will the Wald-based hypothesis tests and confidence intervals that use these estimates; in particular, standard errors will be too small and the null will be rejected too often. This is an area of ongoing research in econometrics, and guidelines for how to detect a "weak instrument" problem are not yet well established. There is, however, a consensus that it is not enough for the first-stage F

---

[16] Two statistics that have been suggested for this purpose are the Cragg-Donald statistic and Anderson's canonical correlations statistic. Both statistics require homoskedastic and independent errors to be valid. See Hall et al. (1996) and Stock and Yogo (2002) for further discussion.

[17] The formal approach would be to test the instrument for "redundancy". An excluded instrument is redundant if the efficiency of the estimation is not improved by its use. See Hall and Peixe (2000) for further discussion.

statistic to be significant at conventional levels of 5% or 1%; higher values are required. Staiger and Stock (1997) recommend an F statistic of at least 10 as a rule of thumb for the standard IV estimator.[18]

If weak instruments are detected or suspected, recent research suggests two possible approaches. The investigator may use an estimator that is relatively robust to weak instruments. These estimators are often non-standard and not widely available in commercial software packages, they may not be robust to heteroskedasticity or clustering, their performance in the presence of weak instruments is still being explored, and so we do not go down this route here. The second approach is to employ a procedure for construction of a confidence interval that is robust to weak instruments. Various such methods have been proposed; see, e.g., Dufour (2003) and Andrews and Stock (2005). One method that is easy to implement for the single-endogenous-regressor case using standard regression software packages is based on the Anderson-Rubin (1949) test.[19]

The method is as follows. Consider a specific hypothesized value $\widetilde{\beta}_0$ for the coefficient on guns in model (1). Subtract the quantity $\widetilde{\beta}_0 guns_i$ from both sides of (1) and then substitute using (4) to obtain

$$(hom_i - \widetilde{\beta}_0 guns_i) = [(\beta_0 - \widetilde{\beta}_0)\theta_1 + \beta_1] control_i + (\beta_0 - \widetilde{\beta}_0)\theta_3 X1_i + (\beta_0 - \widetilde{\beta}_0)\theta_4 X2_i \\ + [(\beta_0 - \widetilde{\beta}_0)v_i + u_i] \tag{5}$$

If the null hypothesis H$_0$: $\beta_0 = \widetilde{\beta}_0$ is correct, both *X1* and *X2* will drop out, and equation (5) simplifies to

$$(hom_i - \beta_0 guns_i) = \beta_1 control_i + u_i \tag{6}$$

The AR test of the null hypothesis is therefore to estimate

$$(hom_i - \widetilde{\beta}_0 guns_i) = \pi_1 control_i + \pi_3 X1_i + \pi_4 X2_i + \eta_i \tag{7}$$

and employ a standard F test of whether the coefficients on the excluded instruments π_3 and π_4 are indeed both equal to zero. The AR test is suitable when a model is just-identified or the degree of

---

[18] Stock and Yogo (2002) provide more detailed advice based on Monte Carlo studies, and show *inter alia* that the Staiger-Stock rule of thumb is a reasonable guideline to follow for the single-endogenous-regressor case when the number of excluded instruments is small.

[19] Not to be confused with the Anderson-Rubin overidentification test, which is the test analogous to the Sargan-Hansen J test for maximum-likelihood estimation.

overidentification is low, and it can easily be made heteroskedastic- or cluster-robust by using the appropriate covariance estimator for (7).[20] Note that the AR test assumes that the orthogonality conditions $E(X1_i \, u_i) = 0$ and $E(X2_i \, u_i) = 0$ are satisfied; if these did not hold, the AR test statistic would be significant, not because $\beta_0 \neq \widetilde{\beta}_0$, but because the excluded instruments were correlated with the composite error term η_i. If either $X1$ or $X2$ is endogenous, the AR test is invalid.

An AR confidence interval is simply the region where the AR test fails to reject the null. For example, an AR 95% confidence interval for the coefficient on guns β_0 would be the range of specific values for $\widetilde{\beta}_0$ such that the AR test statistic is below the 5% critical value for the F distribution. An AR confidence interval is fully robust to weak instruments, and the intuition for this is clear from inspection of equations (5) and (7). Note that $\pi_3 = (\beta_0 - \widetilde{\beta}_0)\theta_3$ and similarly for π_4. If $X1$ and $X2$ are only weakly correlated with *guns*, then θ_3 and θ_4 will be small, so will π_3 and π_4, and a test of H_0: $\pi_3 = \pi_4 = 0$ can fail to reject the null even if the hypothesized $\widetilde{\beta}_0$ is far from the true β_0. Weak instruments will therefore tend to widen the AR confidence interval for the impact of guns β_0, i.e., low instrument relevance will be properly reflected in the imprecision of the estimated impact of gun levels on homicide.[21]

5.   DATA, MODEL AND ESTIMATION STRATEGY

Data

To estimate the impact of gun availability on homicide rates, we use cross-sectional data for all U.S. counties which had a population of 25,000 or greater in 1990, and for which relevant data were available (N=1,462). These counties account for about half of all U.S. counties but over 90% of the U.S. population in that year. The use of 1990 data is dictated by two factors. First, most of the control variables included in the homicide equations to mitigate omitted variable bias are available at the county-level only during census years. A second reason for choosing 1990 is the fact that the firearm crime rate (homicide, robbery, and assault) had reached its highest level in nearly 30 years by 1990. It is reasonably argued that

---

[20] The AR test loses power as the number of excluded instruments goes up. More powerful tests have recently been proposed; see Andrews and Stock (2005) for a survey.

[21] We calculate our AR confidence intervals with a simple grid search over possible values β_0. Note that some care in calculating AR confidence intervals may be needed in practice because the AR test can sometimes generate confidence intervals that are either empty or disjoint. See Zivot et al. (1998).

if gun availability is responsible for higher homicide rates, the high levels of firearm crime in 1990 should provide one of the best opportunities to date for testing the gun availability-homicide relationship.

County-level data were chosen for several reasons. First, the use of counties provides for a diverse sample of ecological units, including urban, suburban, and rural areas. Second, counties are more internally homogenous than nations, states, or metropolitan areas, thereby reducing potential aggregation bias. Third, counties exhibit great between-unit variability in both gun availability and homicide rates, which is precisely what gun availability and homicide research is trying to explain. Fourth, county data provide a much larger sample than previous gun level studies, which have focused mainly on nations, states, or large urban cities (almost all with 50 or fewer cases; see Table 1). The large sample size provides us with greater statistical power to detect more modest effects of gun availability on homicide rates, while still permitting us to enter numerous control variables in the homicide equations to minimize omitted variable bias. A large sample size is particularly important when IV/GMM methods are used, because these estimators are only consistent, i.e., bias approaches zero only as the sample size increases.

Model

We follow the convention for crime policy studies and use a linear model in which most variables are specified in logs. The dependent variables in our model are the gun, nongun, and total homicide rates per 100,000 county population. Homicide data for each county were obtained using special Mortality Detail File computer tapes (not the public use tapes) made available by the National Center for Health Statistics (U.S. NCHS 1997). The data include all intentional homicides in the county with the exception of those due to legal intervention (e.g., police shootings and executions). Homicide rates are averages for the seven years 1987 to 1993, thus bracketing the census year of 1990 for which data on many of the control variables were available. Seven years were covered to reduce the influence of random year-to-year aberrations, e.g., misclassification of homicides as other kinds of deaths such as suicides or unintentional deaths, and to allow the use of rates in a linear model as an approximation for count data (see below).

It is desirable to separately assess rates of homicide with and without guns, to provide sharper tests of the hypothesis that gun levels affect homicide rates. The decision criteria upon which we rely for determining whether gun levels causes a net increase in homicide rates are as follows. If the gun level has a net positive impact on homicide rates, it should have (1) a significant positive association with the gun homicide rate, (2) a significant positive association with the total homicide rate, and (3) be less strongly positively correlated with the nongun homicide rate than with the gun homicide rate. On the other hand, if gun levels are as strongly positively associated with nongun homicides rates as with gun homicide rates (or more so), this suggests that the gun level is merely a correlate of some omitted variable that affects homicide in general, but the gun level has no effect of its own, since there is no strong reason why gun levels should increase the rate of homicides committed without guns. If (1) is true, but (2) is not, it would generally indicate that gun availability merely shifts criminals from nongun weapons to guns, but has no net effect on the number of people murdered. If (2) is true, but (1) is not, it suggests that gun levels are merely associated with some omitted variables that have an effect on total homicide rates but that gun levels themselves have no effect, since they should have their effects by, at minimum, increasing homicides committed with guns.

In our main regressions, the crime rate variables are specified in logs. This poses some minor problems, because even though we are using 7-year averages and excluding the smallest counties, a small number of counties have zero murders: of the 1,462 counties in the sample, 20 had no gun murders (about 1% of the sample), 52 had no nongun murders (about 4%), and 3 had no murders at all. Our approach is to report in detail the results using the logged crime rates and dropping the observations for which the dependent variable is undefined. Histograms of the logged crime rates after dropping zero-crime-rate counties are shown in Figure 2; simple visual inspection suggests no particular skewness or outlier problems after log transformation.

Figure 2: Distributions of Log Crime Rates
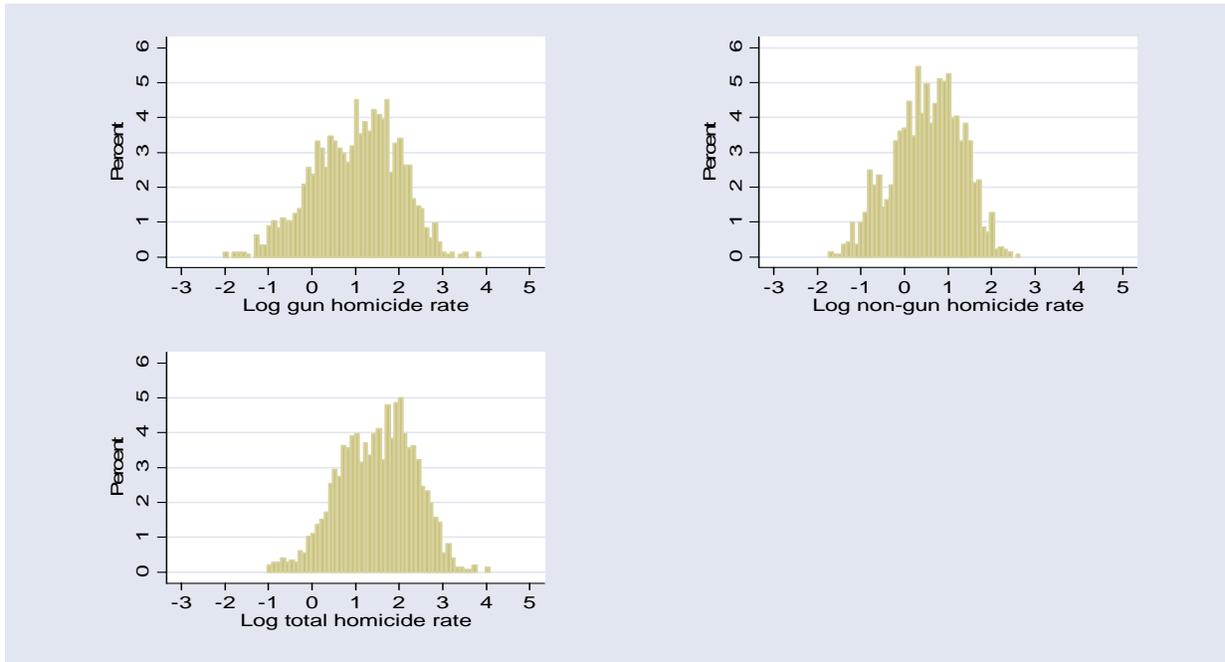
To check the robustness of the results, we also estimated using a variety of transformations of the dependent variable that do not require us to drop observations: (1) logged "add 1" crime rates, where the number of murders for the 7-year period has 1 added to it; (2) "one-sided winsorized"[22] log crime rates, where prior to taking logs the zero crime rate counties are assigned a crime rate equal to the lowest non-zero-crime-rate county; (3) "two-sided winsorized" log crime rates, where the treatment of the zero crime rate counties is as in (2), and the same number of the highest crime counties are, prior to logging, assigned a crime rate equal to that of the next highest (i.e., the highest non-transformed) county; (4) raw crime rates (murders per 1,000 population).

While it might be desirable to measure gun ownership levels directly using survey-based estimates, surveys asking people directly whether they own guns are usually limited to a single large area, such as a nation or state. Instead, we use the best indirect measure of gun availability for cross-sectional research, the percent of suicides committed with guns (PSG). Some of the smaller counties typically have few or no suicides in a given year, and misclassification of a few suicides as homicides or accidents in small counties could produce substantial measurement error in a single year's PSG. Therefore, as was

---

[22] See the help file for the Stata command "winsor" by Nick Cox (2003) for a discussion of "winsorizing" and for references to the relevant statistical literature.

done with homicide rates, PSG was computed for the 7-year period 1987 to 1993, bracketing the decennial census year of 1990. Similar to homicide, data for the percent of suicides committed with guns were obtained using special Part III Mortality Detail File computer tapes made available by the National Center for Health Statistics. Unlike widely available public use versions, the tapes permit the aggregation of death counts for even the smallest counties (U.S. NCHS 1997).

Figure 3 shows histograms of our gun proxy PSG in raw percentage form and after logging. There are no problems of zeros here; rather, the issue is whether the skewness to the left seen in the distribution of log PSG is of any concern. Our approach is to report detailed results using log PSG, and to confirm robustness of the findings with regressions using PSG in raw percentage form.

Figure 3: Distributions of PSG and Log PSG
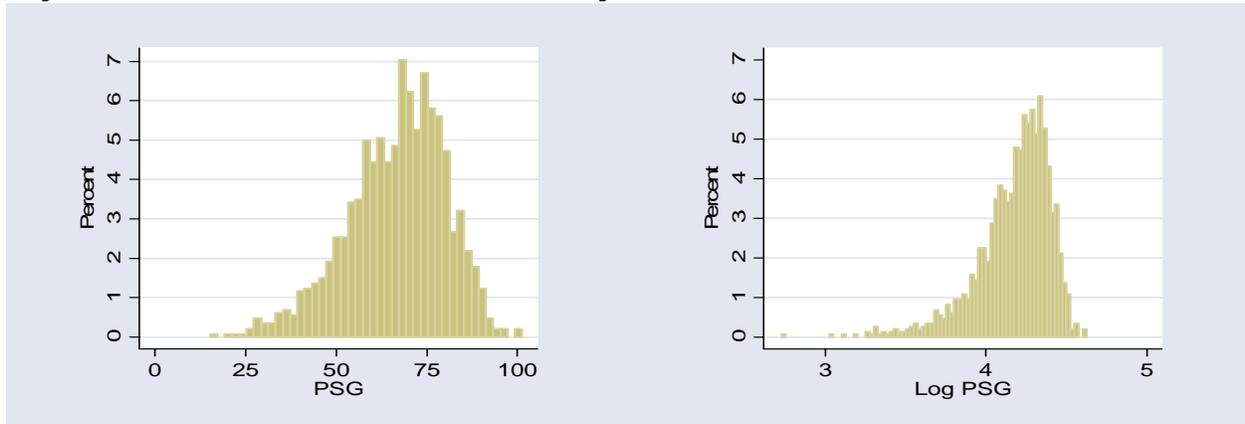
We also need to calibrate our proxy to available survey-based measures of gun levels. The most convenient calibration is to the mean percentage of households with guns (HHG) according to the General Social Surveys (GSS). National gun survey prevalence figures have been available since 1959, though not for every year. The mean HHG for 1959-2003 is 44.2% while the mean PSG for 1959-2002 (the latest year available) is 54.9%. These figures imply a value of 54.9/44.2=1.24 for the calibration factor δ_0 by which we should inflate or deflate the estimated coefficient on PSG b_0 so as to obtain an estimate of β_0, the impact of gun levels on homicide rates (see equation 3). Neither PSG nor HHG varied greatly during this period, and the use of a different reference period would matter little.

Using state-level measures from surveys conducted by the Centers for Disease Control (CDC) in 2002 (Okoro et al., 2005) and PSG data for 1995-2002 taken from CDC's WONDER service, simple OLS

regressions (based on 50 states) of PSG and log PSG and the corresponding HHG and log HHG measures are as follows (standard errors are in parentheses):

```
Log PSG    =    2.31   +   0.481 Log HHG + e
               (0.12)     (0.035)

PSG        =   30.11   +   0.706 HHG + e
               (2.85)     (0.069)
```

The coefficients on log HHG and HHG can be interpreted as estimates of the log-log and level-level calibration δ_0, respectively. The figures suggest that in our main log-log estimations, the coefficient on log PSG should be approximately halved ($\delta_0=0.481$) in order to be interpreted as the HHG-homicide elasticity. In a levels-levels estimation, the calibration should be in the neighborhood of 0.706 (based on the 50 states regression) to 1.24 (based on national means). These are, however, only approximations based on limited data and simple linear calibrations. A more cautious conclusion would be that PSG is, within an order of magnitude, already calibrated to HHG, and that interpretations of the estimated coefficient on gun levels will be most robust in the neighborhood of mean PSG.

In addition to the gun prevalence measure, we included numerous county-level control variables, paying particular attention to those that prior theory and research suggest are important determinants of *both* gun ownership levels and homicide rates. Failing to control for confounders that affect both gun availability and homicide rates would generate an endogeneity bias in the coefficient on PSG, as discussed earlier. Decisions as to which control variables to include in the homicide equations were based on a review of previous macro-level studies linking homicide rates to structural characteristics of ecological units (see Kleck, 1997, Chapter 3; Kovandzic et al., 1998; Land et al., 1990; Sampson, 1986; Vieraitis, 2000 and the studies reviewed therein).

We were particularly concerned to control for variables that had opposite-sign associations with gun levels and homicide rates because such variables could suppress evidence of any positive effect of gun levels on homicide rates. Thus, we controlled for the percent of the population that is rural because rural people are more likely to own guns, but less likely to commit homicide. Likewise, we controlled for the poverty rate, the share of the population in the high-homicide ages of 18 to 24 and 25 to 34, and the

African-American share of the population because people in these groups are less likely to own guns, but more likely to commit homicide, than other people (Kleck, 1997; Cook and Ludwig, 1997; U.S. FBI, 2000). The other controls used were percent Hispanic, population density, average education level, unemployment rate, transient population (born out-of-state), vacant housing units, female-headed households with children, median household income, households earning less than $15,000, and inequality (ratio of households earning more than $75,000 to households earning less than $15,000).

The sets of controls for rurality and age structure are, exceptionally, used in percentage rather than log form. Because the raw percentages sum to 100, including all categories would generate a perfect collinearity problem, and so one category must be omitted. Using raw percentages therefore has the appealing feature that the results are invariant to whichever percentage is the omitted category. We omit the percentage rural and the percentage aged 65+.

The use of county-level data has an additional advantage: by including state fixed effects, i.e., state dummy variables, we are able to control for any unobserved or unmeasured county characteristics that vary at the state level and that could be expected to influence both gun levels and homicide rates. Examples of such confounders would be state laws and judicial practice relating directly or indirectly to homicide and gun ownership, state-level resources devoted to law enforcement, and incarceration rates in state prisons. Testing of the fixed effects in the estimations below strongly supported their inclusion as controls.[23] The disadvantage of this approach is that only variables available at the county level can be used in the estimations, because state-level measures would be perfectly collinear with the fixed effects.[24] We do not include explicit state dummies and instead use an estimation routine that obtains numerically equivalent results by applying the "within" transformation to all variables, i.e., expressing them as deviations from state means. The R^2 reported with the regressions is the "within R^2", and measures how much of the within-state variation in homicide rates our models explain.

The key challenge in using IV methods is finding a source of identifying variation: here, variables that are correlated with gun levels, but that are exogenous with respect to homicide and that a priori

---

[23] F tests of the significance of the state fixed effects vs. a specification using a set of 9 regional dummies.

[24] It also means that the single observation for Washington, D.C. drops out of the fixed-effects regressions.

reasoning and evidence suggest should be excluded from the homicide equation. The excluded instrumental variables used in this paper are (1) an index (RGUNMAG) comprised of subscriptions to each of the three most popular outdoor/sport magazines (*Field and Stream*, *Outdoor Life*, and *Sports Afield*) in 1993, per 100,000 county population (Audit Bureau of Circulations, 1993), and (2) the percent of the county population voting for the Democratic candidate in the 1988 Presidential election (PCTDEM88). The gun magazine index was created using principal components analysis; the analysis yielded a factor that explained 84 percent of the cumulative variance in this latent construct.[25]

Both excluded instruments are theoretically important correlates of gun ownership that are plausibly otherwise unrelated to homicide. RGUNMAG serves as a measure of interest in outdoor sports such as hunting and fishing, or perhaps as a measure of a firearms-related "sporting/outdoor culture" (Bordua and Lizotte, 1979). PCTDEM88 serves as a measure of political liberalism and hence should be negatively correlated with gun ownership. The 1988 election results were chosen in preference to the 1992 results because the date precedes the census year from which most our data are taken (and hence is more plausibly exogenous), and because the choice between the two main candidates in 1988 maps more closely to attitudes towards gun ownership.[26] Prior research suggests that both variables are important predictors of gun ownership (Kleck, 1997, pp. 70-72; Cook and Ludwig, 1997, p. 35).

Table 2 lists and provides a brief description of each variable used along with their means and standard deviations. Data for the control variables were obtained from the U.S. Bureau of the Census, *County and City Data Book, 1994*, except for PCTDEM88, which is from ICPSR (1995), and rurality, which is from U.S. Census Bureau (2000).

Estimation strategy

The estimation and testing procedure we follow is the GMM approach outlined above in Section 4. It is natural to expect the presence of heteroskedasticity in a cross-sectional dataset, and indeed

---

[25] The factor had an eigenvalue of 2.52, well above the conventional threshold of 1.00. The loading scores for each magazine were as follows: SPAFIELD (0.824), LIFE (0.968), and STREAM (0.948).

[26] In the 1992 election, unlike the 1988 election, the politically less conservative candidate (negatively correlated with gun ownership) was also a southerner (positively correlated with gun ownership). The 1992 results are also less easily interpreted because of the significant share of the vote that went to the third-party candidate, Ross Perot.

application of the Pagan-Hall (1983) test for heteroskedasticity in IV estimation to our data suggests it is present here. We therefore use GMM estimation that is efficient in the presence of arbitrary heteroskedasticity.

An oft-neglected issue with cross-sectional data on locations is the potential problem of spatial correlation. It is reasonable to suspect that observations in two physically adjacent counties are more likely to have correlated disturbance terms than two counties at opposite ends of the state or country. If spatial correlation is present, standard errors that assume independence are likely to be underestimated and test statistics will be invalid. A straightforward method of addressing this in the GMM context is the "cluster-robust" approach, where clusters are defined as groups of counties – here, states.[27] The corresponding GMM estimator is efficient in the presence of both arbitrary heteroskedasticity and arbitrary within-state correlation of the disturbance term, and requires only the assumption of independence across states, which is a reasonable assumption given our fixed-effects specification. The main drawback to this approach is that we have only 50 clusters (states), and hence relatively few degrees of freedom. Most of our regressions have 18 exogenous regressors and 2 excluded instruments, leaving us with effectively only 50-18-2=30 observations on clusters for calculating cluster-robust standard errors. We therefore use the fixed-effects GMM estimator that is efficient in the presence heteroskedasticity for our benchmark results, and report the results using the cluster-robust fixed-effects GMM estimator as a check on the sensitivity of the findings to the presence of spatial correlation.

Our initial specification treats PSG as exogenous; it is the two-step efficient GMM estimator with exogenous regressors that allows for arbitrary heteroskedasticity (Cragg's HOLS estimator). The corresponding J statistic is a test of the full set of orthogonality conditions, i.e., the exogeneity of PSG, RGUNMAG and PCTDEM88 (plus the other covariates). Our second specification is the two-step efficient GMM estimator that treats PSG as endogenous, and the J statistic is a test of the reduced set of orthogonality conditions, i.e., the exogeneity of RGUNMAG and PCTDEM88 (again plus the other covariates). The C statistic reported with the initial specification is a test of the endogeneity of PSG and is

---

[27] See Wooldridge (2003) for an overview and discussion of the case where the number of groups is small.

based on the difference of the two J statistics.[28] We test for instrument relevance using a heteroskedastic or cluster-robust F test of the joint significance of the excluded instruments RGUNMAG and PCTDEM88 in an OLS estimation of the first-stage equation of the gun proxy (PSG or log PSG), and we also examine the significance of RGUNMAG and PCTDEM separately using conventional t-statistics. Because these instruments are, in some specifications, bordering on weak, we also report an Anderson-Rubin confidence interval for the coefficient on PSG, and, for comparison, the usual Wald-based confidence interval employing the GMM-estimated standard error. The versions of these tests that we use are all heteroskedastic- or cluster-robust; non-robust test statistics would tend to be biased upwards and lead to Type I errors. Lastly, as a check on the sensitivity of the results, we also estimate the main equations using the corresponding inefficient GMM estimators, OLS and IV. With the exception of the F test-based AR confidence intervals, we report large-sample standard errors and significance tests; i.e., we do not make a finite-sample adjustment for the number of explicit regressors, and we use the normal or χ^2 distributions for significance tests, p-values and confidence intervals.[29]

The statistical package Stata was used for all estimations. The main IV/GMM estimation programs, *ivreg2* and *xtivreg2*, were co-authored by one of us (Schaffer), and can be freely downloaded via the software database of RePEc.[30] For further discussion of how the estimators and tests are implemented, see Baum, Schaffer and Stillman (2003) and (2005), and the references therein.

6.   **RESULTS**

Estimation results for the benchmark regressions using the logged gun homicide rate as the dependent variable are in Table 3. Columns 1 and 2 report the results of 2-step GMM estimations that are efficient in the presence of arbitrary heteroskedasticity and that treat PSG as exogenous and endogenous, respectively, along with heteroskedastic-robust standard errors and test statistics. Columns 3 and 4 report

---

[28] The difference in reported J statistics is not exactly equal to the reported C statistic, because we use a version of the latter that is guaranteed to be non-negative in finite samples. See Hayashi (2000) and Baum et al. (2003).
[29] The heteroskedastic-robust variances and statistics have a degrees-of-freedom adjustment for the number of fixed effects (50); the adjustment is not required for the cluster-robust variances. See Wooldridge (2002), pp. 271-75.
[30] http://ideas.repec.org/SoftwareSeries.html. *ivreg2* is a general-purpose IV/GMM estimation routine for linear models; *xtivreg2* supports fixed-effects panel data models.

the corresponding estimates that are efficient in the presence of heteroskedasticity and within-state clustering, plus heteroskedastic- and cluster-robust standard errors and statistics.

We consider first the heteroskedatic-robust results in columns 1 and 2. Most of the parameter estimates for the 18 control variables are significant, and the significant coefficients have the expected sign in both specifications. High gun murder rates are associated with high population density, lower education levels, and the various poverty, low-income and inequality measures. The percentage of the population that is black is associated with high gun crime rates, as is the percentage Hispanic, but in the latter case only in the first specification, when PSG is treated as exogenous. The only modest surprise is provided by the age structure controls; contrary to expectations, the 18-24 age group was associated with relatively low gun homicide rates,[31] though this is a common finding (Marvell and Moody 1991). The overall fit of the regressions is quite good, with the PSG-exogenous specification explaining 44% of the within-state variation in county-level log gun homicide rates, and the PSG-endogenous specification explaining 33%.

The key results concern the coefficient on and the exogeneity of PSG. Column 1 shows that in the efficient GMM estimation where log PSG is treated as exogenous, the variable has a coefficient of 0.290 and is statistically significant at the 5% level. This confirms the oft-reported result that, when endogeneity issues are ignored, gun levels are associated with higher gun crime rates. When log PSG is treated as endogenous and instrumented with RGUNMAG and PCTDEM88, the picture changes dramatically. Column 2 shows that log PSG has a *negative* coefficient of -1.50 that is statistically significant at the 1% level.

We now apply the GMM-based procedure outlined in Section 4 for testing the exogeneity of PSG. [Step 1] The J statistic in column 1 is 13.3. This is very large for a $\chi^2(2)$ statistic; the p-value is only 0.001. We therefore reject the null hypothesis that the orthogonality conditions in the PSG-exogenous estimation are satisfied, and take this as strong evidence that one or more variables – log PSG, RGUNMAG, PCTDEM88, and/or the control variables – are endogenous. [Step 2] The J statistic for the

---

[31] The effect is comparable to that of the omitted category of 65+ and lower than those for the remaining categories.

PSG-endogenous estimation in column 2 is 0.25, which small for a $\chi^2(1)$ statistic – the corresponding p-value is 0.61. We therefore cannot reject the null that RGUNMAG, PCTDEM88, and the control variables are exogenous; in other words, the evidence suggests that our instruments are valid. [Step 3] We now test explicitly whether log PSG is endogenous using a C test based on the J statistics for the PSG-exogenous and PSG-endogenous estimations. The C statistic reported in column 1 is 13.0.[32] This is very large for a $\chi^2(1)$ statistic; the p-value is only 0.0003. We therefore have strong evidence that log PSG is endogenous, and that some form of IV/GMM estimation is required. Before we turn to this last estimation [Step 5], however, we must also consider whether our excluded instruments RGUNMAG and PCTDEM88 are relevant as well as valid.

The first-stage heteroskedastic-robust F statistic reported in column 2 is 26.2, well above the Staiger-Stock rule-of-thumb level of 10. Both the excluded instruments are correlated with log PSG in the expected directions and at the 0.1% significance level: counties voting Democrat in 1988 tend to have fewer guns as proxied by log PSG, and subscriptions to outdoor magazines are associated with higher gun levels. We conclude that our instruments in this estimation are relevant and not weak. Nevertheless, as a check we also report a heteroskedastic-robust AR 95% confidence interval for log PSG. The AR confidence interval is [–3.05 , –0.34] only slightly wider than the conventional Wald-based confidence interval of [–2.52 , –0.48].

Having shown that the GMM estimation in column 2 satisfies the requirements of both validity and relevance, we turn to the issue of calibration. The calibration exercise above suggests we should roughly halve coefficient estimate of –1.50 to obtain the elasticity of gun homicide with respect to HHG, i.e., about -0.75. The estimate is therefore not only statistically significant, it is potentially also practically significant. We hasten to add, however, that this is a very approximate estimate: after calibration to HHG and allowing for uncertainty in the calibration itself, both the conventional Wald and weak-instrument-robust AR confidence intervals would include low (practically insignificant albeit statistically significant) HHG-gun homicide elasticities. In sum, we have strong evidence that gun levels are endogenous, and

---

[32] The C statistic differs slightly from the difference between the relevant J statistics because we use a version of the C test that guarantees a positive test statistic. See Hayashi (2000) or Baum et al. (2003) for details.

when this is accounted for in the estimation, the positive association of gun levels with gun homicide completely disappears; if anything, gun levels are associated with lower, not higher, gun homicide rates.

These findings are essentially unchanged if we allow for intra-state spatial correlation as well as heteroskedasticity. When we treat log PSG as endogenous (column 4), the two-step efficient GMM estimate of the coefficient on log PSG is virtually the same as before: –1.53, again significant at the 1% level. The large J statistic in the PSG-exogenous GMM estimation in column 3 again strongly suggests that we should reject the hypothesis that PSG, RGUNMAG, PCTDEM88 and the controls are all exogenous; the small J statistic in the PSG-endogenous GMM estimation in column 4 means we fail to the reject the hypothesis that RGUNMAG, PCTDEM88 and the controls are all exogenous; and the large C statistic in column 3 strongly implies we should reject the hypothesis that PSG in particular is exogenous. Again, some form of IV/GMM estimation that treats PSG as endogenous is called for. With respect to the requirement of instrument relevance, column 4 shows that both of the excluded instruments are still significant in the first-stage regression. A potential cause for concern, however, is the first-stage F statistic, which at 12.2 is somewhat low and indicative of a possible weak instruments problem. We therefore refer to the Anderson-Rubin 95% confidence interval for log PSG. This is [–3.35 , –0.06], vs. the Wald-based interval of [–2.52 , –0.48]. The reduced relevance of the excluded instruments when allowing for spatial correlation noticeably widens the AR confidence interval, but we can still reject the hypothesis that our estimated coefficient on log PSG is zero, albeit at a reduced level of statistical significance. Our key finding is therefore clearly robust to within-state spatial correlation. We again conclude that we must treat gun levels as endogenous, and when this is done we find that they are, if anything, associated with lower, not higher, rates of gun homicide across counties.

Table 4 presents the corresponding results for the log nongun homicide rate. The patterns in the coefficients on the covariates are similar to those for the gun homicide equations; the main differences are that rurality is now significantly associated with lower nongun homicide rates, and the economic controls are less significant. The main difference vis-à-vis the gun homicide coefficients is in the estimated impact of gun levels: across all estimations for the nongun homicide rate, the coefficient on log PSG is insignificantly different from zero. The tests of orthogonality conditions suggest, moreover, that guns

may be treated as exogenous in the nongun homicide equation: the J statistics are low for both the specifications that treat guns as exogenous (columns 1 and 3) and the specifications that treat guns as endogenous (columns 2 and 4), and the C statistics for the endogeneity of log PSG are also low. The implication is that the estimates for nongun homicide that treat guns as exogenous are to be preferred on efficiency grounds, and the estimates that treat guns as endogenous are preferable if we wish to avoid a Type II error in concluding they are exogenous; but the choice matters little because both sets of estimations suggest no impact of guns on nongun homicide rates. The tests of instrument relevance are (naturally) very similar to those for the gun homicide equations. Again the low first-stage F statistic when allowing for intra-state spatial correlation is a concern, and again the AR confidence intervals for log PSG support the Wald test results: gun levels have no impact on nongun homicide rates.

If we take these estimates of gun effects seriously, they suggest that gun levels in the general public may, on net, have a deterrent effect on gun homicide rates, but no such effect on nongun homicides. Deterrent effects would be stronger for gun homicides if their perpetrators were more likely to plan the killings (or crimes leading up to the attacks, such as robbery or a drug deal) than those who use less lethal weapons. The fact that an aggressor chose a lethal weapon, better suited to lethal purposes, rather than merely making use of whatever weapons happened to be available at the scene, may itself be an indication of premeditation. Thus, people who kill with guns may be more easily deterred by the prospect of confronting a gun-armed victim than those who kill with other weapons, because the former are more likely to think about the potential costs of their actions.

The results in Table 5 are for the estimations using the log total homicide rate (gun+nongun) as the dependent variable. The results are similar to, but slightly weaker than, those for the gun homicide rate. In the specifications that treat log PSG as exogenous, its coefficient is positive and either insignificant (column 1) or significant at the 5% level (column 3). When gun levels are treated as endogenous, the coefficient on log PSG becomes negative (about −1.0) and significant at the 5% level, and the AR and Wald confidence intervals are again similar. The J statistics for the PSG-exogenous estimations are large and suggest that one or more orthogonality conditions is violated. When the orthogonality condition for PSG is dropped the J statistics become very small, suggesting that the

remaining orthogonality conditions are valid, and the C test strongly rejects the hypothesis that gun levels are exogenous. The first-stage statistics again indicate a possible weak instrument problem for the cluster-robust specification, and in this case the weak-instrument-robust AR 95% confidence interval in column 4, though mostly negative, includes zero, i.e., an AR test of the coefficient on log PSG fails to reject a zero impact of gun levels on homicide at the 5% significance level.[33] Our key finding remains: gun levels have at least no impact, and possibly a negative impact, on total homicide rates.

We tested the robustness of the coefficient estimates by using the corresponding inefficient but consistent GMM estimators, OLS and IV, and heteroskedastic- and cluster-robust standard errors; the results were essentially identical.[34] Finally, we checked robustness to the functional form by comparing the results using the dependent variables in the four possible log measures (drop zero-crime-rate counties as used in the results above, "add 1", "one-sided winsorized", and "two-sided winsorized") plus a fifth, unlogged crime rates, and by using both log PSG and unlogged PSG as our gun proxies. This gave us 5*2=10 possible combinations of specifications, and estimating with and without cluster-efficient GMM gave us 10*2=20 estimations in total for each of the three categories of homicide rates. The qualitative results were almost entirely unaffected by these variations in functional form with respect both to the estimate of impact of gun levels and to the findings on endogeneity, instrument validity and instrument relevance. When both crime rates and PSG were logged, the quantitative results were also essentially unchanged, with estimated coefficients very similar to those reported above. The specifications that used the unlogged nongun homicide rate (CRNGMUR) provided the sole exception to this pattern, and, if anything, strengthen our general conclusions: both PSG and log PSG were found to be endogenous, and when instrumented, the estimated coefficients were negative and statistically significant, suggesting that higher gun levels are associated with lower nongun homicide rates.

With respect to calibration, in the level-level specifications using the heteroskedastic-robust efficient GMM estimator, the coefficients on PSG in the gun, nongun, and total homicide equations were

---

[33] The AR confidence interval is negative and excludes zero if the level of confidence is reduced from 95% to 89%.
[34] The first-stage and J and C statistics were identical by construction, the latter two because efficient GMM was needed to produce heteroskedastic- and cluster-robust J and C statistics.

–0.227, –0.066, and –0.294, all statistically significant with p-levels about 0.01 (the cluster-robust estimates were very similar). Using the calibration to HHG suggested above ($0.706 < \delta_0 < 1.24$) implies that an increase in 10 percentage points in the number of households with guns in a county would reduce gun homicides by about 2 persons per 100,000 population, nongun homicides by about 0.7 persons, and total homicides by about 3 persons. The implied elasticities of homicide to gun levels are somewhat higher those estimated directly with the log-log specifications in Tables 3-5.[35] Again, however, we hasten to add that these are very approximate estimates, and confidence intervals would include low, i.e., practically unimportant, impacts. Our main conclusions, after from all these robustness checks, remain that the positive correlation between gun levels and homicide rates is driven by endogeneity bias, and when the endogeneity of gun levels is properly addressed in the estimation, any positive correlation vanishes.

7.   CONCLUSIONS

Most studies of the gun-crime relationship have ignored the endogeneity problem, and the few that have tried to address it with IV methods have failed to perform the tests needed to tell whether their estimation procedures were adequate. We have presented a formal analysis of the three main problems facing researchers – reverse causality, mismeasurement of gun levels, and omitted/confounding variables – and discussed how to address these problems using estimation and specification testing procedures in a GMM framework. We applied these procedures to U.S. county level data, and found strong evidence of the existence of endogeneity problems. When the problem is ignored, gun levels are associated with higher rates of gun homicide; when the problem is addressed, this association disappears or reverses (though the reversal, suggesting that higher gun levels lead to lower gun crime rates, should be treated with caution).

Our findings provide no support for the more guns, more homicide thesis. The appearance of such an effect in past research appears to be the product of methodological flaws, especially the failure to

---

[35] E.g., mean county CRGMUR in the sample is 4.1 homicides per 100,000 (Table 2); average HHG in the U.S. is 44% (Section 5). With a 1:1 calibration of PSG to HHG, the implied elasticity is therefore $(-0.227)*44/4.1 \approx -2.4$, vs. about -1.5 in Table 3.

properly account for the possible effect of crime rates on gun ownership levels. Higher gun prevalence may have a net violence-elevating effect, but one that is confined to criminals or perhaps other high-risk subsets of the population. To be consistent with our generally null findings regarding the effects of gun levels, however, if there were such a violence-increasing effect of guns among criminals, it would have to be counterbalanced by violence-reducing effects among noncriminals. Guns among criminals may increase homicide while guns among noncriminals decrease it, with the two opposite-sign effects canceling each other out. The most straightforward policy implication of such a combination of effects would be that gun control measures should focus on reducing gun prevalence among criminals while avoiding reducing it among noncriminals.

It might be argued that we failed to find support for the more guns, more homicide thesis because PSG serves primarily as an indicator of gun prevalence among noncriminals, especially in suburban/rural counties. Even if this were true, however, one would still expect criminal gun prevalence to be positively correlated with noncriminal gun prevalence, if for no other reason than that most criminals acquire guns as a direct or indirect result of thefts from noncriminals (Wright and Rossi 1986, p. 196). Thus, PSG would still measure criminal gun prevalence, but less strongly than more direct measures. Therefore, a more precise variant of this speculation would be that PSG might be a weaker proxy for criminal gun prevalence than it is for noncriminal gun prevalence, or that the excluded instruments that we use (RGUNMAG and PCTDEM88) are more weakly correlated with criminal than with noncriminal gun prevalence; either would lead to weaker associations between gun prevalence and violence rates than would be obtained if we could more specifically measure criminal gun prevalence. It bears repeating, however, that at this point this idea is nothing more than a plausible but empirically unsupported speculation. Therefore, identifying proxies that can separately measure criminal and noncriminal gun prevalence should be a top priority for future research.

Table 1. Macro-Level Studies of the Impact of Gun Levels on Crime Rates[a]

Study	Sample	Gun Measure[b]	Crime Rates[c]	Results[d]	Simul?[e]
Brearley (1932)	42 states	PGH	THR	Yes	No
Krug (1967)	50 states	HLR	ICR	No	No
Newton and Zimring (1969)	4 years, Detroit	NPP	THR,TRR,AAR, GHR	Yes	No
Seitz (1972)	50 states	GHR,FGA,AAR	THR	Yes	No
Murray (1975)	50 states	SGR,SHR	GHR,AAR,TRR	No	No
Fisher (1976)	9 years, Detroit	NPP,GRR,PGH	THR	Yes	(No)
Phillips et al. (1976)	18 years, U.S.	PROD	THR	Yes	No
Brill (1977)	11 cities	PGC	ICR	No	No
			THR	Yes	
			TRR	No	
Kleck (1979)	27 years, U.S.	PROD	THR	Yes	(No)
Cook (1979)	50 cities	PGH,PSG	TRR	No	No
			RMR	Yes	
Kleck (1984)	32 years, U.S.	PROD	THR	No	(No)
			TRR	Yes	
Maggadino and Medoff (1984)	31 years, U.S.	PROD	THR	No	(No)
Lester (1985)	37 cities	PCS	VCR	No	No
Bordua (1986)	102 counties 9 regions	GLR,SIR	HAR,THR,GHR	No	No
McDowall (1986)	48 cities, 2 years	PGH,PSG	TRR	No	(No)
Lester (1988b)	9 regions	SGR	THR	Yes	No
McDowall (1991)	36 years, Detroit	PSG,PGR	THR	Yes	(No)
Killias (1993)	16 nations	SGR	THR,GHR	Yes	No
Kleck and Patterson (1993)	170 cities	5-item factor incl. PSG[f]	THR,GHR,TRR, GRR,AAR,GAR	No	Yes
Lester (1996)	12 nations	PGH,PSG	THR,GHR	Yes	No
Southwick (1997)	48 years, U.S.	PROD	THR,TPR,TRR, AAR	No	Yes
Southwick (1999)	34 years, U.S.	HGS	THR,TRR,AAR, TPR,VCR,BUR	No	No
Hemenway and Miller (2000)	26 nations	PGH,PSG	THR	Yes	No
Lott (2000)	15 states, 2 years[g]	SGR	THR,TPR,TRR, AAR, 3 others	No	No
Stolzenberg and D'Alessio (2000)	4 years, 46 counties	CCW, GUNSTOL	VCR	Yes	No
Duggan (2001)	19 years, 50 states	GMR	THR,TPR,TRR, AAR	Yes	No
Hoskin (2001)	36 nations	PSG	THR	Yes	(No)
Killias et al. (2001)	21 nations	SGR	THR,TRR,TAR, GHR,GRR,GAR	No	No
Sorenson and Berk (2001)	22 years, California	HGS	THR	Yes	(No)
Cook and Ludwig (2002)	22 years, 50 states	PSG	BUR	Yes	(No)
Miller et al. (2002)	10 years, 50 states	PSG,PHG	THR	Yes	No
	10 years, 9 regions	SGR	THR	No	No
Ruddel and Mays (2005)	50 states	3-item factor incl. PSG[h]	THR	Yes	No
Moody and Marvell (2005)	50 states, 22 years	PSG, SHR	THR,TPR,TRR, AAR,BUR	No	Yes

Notes to Table 1:

a. Table covers only studies and findings where the dependent variable was a crime rate, as opposed to the fraction of crimes committed with guns, and where gun ownership levels were actually measured, rather than assumed. Studies that examined only gun violence rates (e.g., only gun homicides) are excluded.

b. Measures of Gun Level: CCW = concealed carry permits rate; FGA = Fatal gun accident rate; GLR = Gun owners license rate; GMR = Gun magazine subscription rates; GRR = Gun registrations rate; GUNSTOL = % of $ value of stolen property due to guns; HGS = handgun sales (retail); HLR = Hunting license rate; NPP = Number of handgun purchase permits; PGA = % aggravated assaults committed with guns; PGC = % homicides, aggravated assaults and robberies (combined together) committed with guns; PCS = same as PGC, but with suicides lumped in as well; PGH = % homicides committed with guns; PGR = % robberies committed with guns; PSG = % suicides committed with guns; PROD= Guns produced minus exports plus imports, U.S.; SGR = Survey measure, % households with gun(s); SHR = Survey measure, % households with handgun(s); SIR = Survey measure, % individuals with gun(s)

c. Crime Rates: AAR = Aggravated assault rate; BUR = burglary rate; GAR = Gun aggravated assault rate; GHR = Gun homicide rate; GRR = Gun robbery rate; HAR = Homicide, assault and robbery index (factor score); ICR = Index crime rate; RMR = Robbery murder rate; THR = Total homicide rate; TPR = Total rape rate; TRR = Total robbery rate; VCR = Violent crime rate

d. Yes=Study found significant positive association between gun levels and violence; No=Study did not find such a link.

e. Did research address possible simultaneous relationship between gun levels and crime rates with properly identified model? (No) means researchers tried to address the issue, but model was still underidentified.

f. Five-item factor composed of PSG, PGH, PGR, PGA, and the percent of dollar value of stolen property due to stolen guns.

g. Panel design, two waves.

h. Three-item factor composed of PSG, the gun theft rate, and the fatal gun accident rate.

Table 2a: Descriptive Statistics		Mean	Std. Dev.	Min	Max
Homicide variables, 1987-93 average					
CRGMUR	Gun homicides per 100,000 population	4.13	4.18	0.00	46.08
Log CRGMUR	" logged	1.01	0.96	-2.02	3.83
CRNGMUR	Nongun homicides per 100,000 pop.	2.17	1.73	0.00	13.39
Log CRNGMUR	" logged	0.53	0.78	-1.74	2.59
CRMUR	Total homicides per 100,000 pop.	6.30	5.62	0.00	57.67
Log CRMUR	" logged	1.50	0.86	-1.01	4.05
Gun availability, 1987-93 average					
PSG	% suicides with guns	66.67	13.46	15.28	100.00
Log PSG	" logged	4.18	0.23	2.73	4.61
Excluded instruments					
PCTDEM88	% presidential vote Democrat, 1988	42.49	9.87	17.70	84.74
Log PCTDEM88	" logged	3.72	0.24	2.87	4.44
RGUNMAG	Principal components measure of 3 top outdoor/sport magazine subscriptions	0.00	1.00	-4.79	2.39
Controls					
DENSITY	Persons per square mile	418	2075	2	53126
Log DENSITY	" logged	4.78	1.27	0.47	10.88
PCTURBAN	% urban (inside urbanized area)	28.32	36.95	0.00	100.00
PCTSUBURBAN	% suburban (outside urbanized area)	25.52	22.22	0.00	100.00
PCTRURAL	% rural (farm+nonfarm)	46.16	26.14	0.00	100.00
PCT0T17	% aged 17 and under	26.34	3.24	15.10	41.70
PCT18T24	% aged 18-24	10.70	3.72	5.10	37.10
PCT25T44	% aged 25-44	30.94	3.02	20.30	45.30
PCT45T64	% aged 45-64	18.95	2.13	8.40	27.10
PCT65PLUS	% aged 65 and over	13.08	3.59	3.00	33.80
PCTBLK	% African-American	9.24	12.67	0.01	72.13
Log PCTBLK	" logged	1.05	1.82	-4.36	4.28
PCTHISP	% Hispanic	4.43	10.25	0.14	97.22
Log PCTHISP	" logged	0.41	1.30	-1.97	4.58
PCTFEM18	% female-headed HHs w/children < 18	58.54	7.11	33.40	84.10
Log PCTFEM18	" logged	4.06	0.12	3.51	4.43
PCTEDUC	% aged 25+ with a BA degree or higher	16.06	7.36	4.60	52.30
Log PCTEDUC	" logged	2.68	0.42	1.53	3.96
PCTTRANS	% born out of state	31.22	15.85	5.09	86.54
Log PCTTRANS	" logged	3.32	0.51	1.63	4.46
Log MEDHHINC	Log median household income, 1989	10.18	0.24	9.23	10.99
PCTINCLT15K	% households with income < $15,000	27.86	8.88	5.00	65.20
Log PCTINCLT15K	" logged	3.27	0.36	1.61	4.18
INEQUALITY	% HHs w/income <$15k / % income >$75k	0.32	0.50	0.02	6.74
Log INEQUALITY	" logged	-1.64	0.89	-4.08	1.91
PCTPOOR	% persons below poverty line, 1989	14.26	6.89	2.20	60.00
Log PCTPOOR	" logged	2.55	0.49	0.79	4.09
PCTUNEMP	% persons unemployed	6.65	2.47	1.50	23.60
Log PCTUNEMP	" logged	1.83	0.36	0.41	3.16
PCTVACANT	% housing units vacant	11.03	7.65	2.70	66.20
Log PCTVACANT	" logged	2.24	0.54	0.99	4.19

Table 2b: Alternative measures of logged homicide variables		Obs	Mean	Std. Dev.	Min	Max
Log CRGMUR	Logged, dropping zero-rate counties	1442	1.01	0.96	-2.02	3.83
	"Add 1", logged	1462	1.11	0.88	-1.33	3.83
	Winsorized lower tail, logged	1462	0.97	1.02	-2.02	3.83
	Two-sided winsorized, logged	1462	0.96	1.01	-2.02	2.84
Log CRNGMUR	Logged, dropping zero-rate counties	1410	0.53	0.78	-1.74	2.59
	"Add 1", logged	1462	0.65	0.70	-1.52	2.60
	Winsorized lower tail, logged	1462	0.45	0.87	-1.74	2.59
	Two-sided winsorized, logged	1462	0.44	0.86	-1.74	1.77
Log CRMUR	Logged, dropping zero-rate counties	1459	1.50	0.86	-1.01	4.05
	"Add 1", logged	1462	1.58	0.79	-0.83	4.06
	Winsorized lower tail, logged	1462	1.49	0.87	-1.01	4.05
	Two-sided winsorized, logged	1462	1.49	0.87	-1.01	3.73

Sources for Tables 2a and 2b: U.S. Bureau of the Census, *County and City Data Book* (1994), except for (a) homicide rates and PSG, from U.S. NCHS (1997); (b) PCTDEM88, from ICPSR (1995); (c) magazine subscription rates used to construct RGUNMAG, from Audit Bureau of Circulations (1993); (d) rurality measures, from U.S. Bureau of the Census (2000).

Table 3: Log gun homicide equation
Dependent variable: Log CRGMUR

	2-step Efficient GMM (heteroskedastic-efficient)		2-step Efficient GMM (heteroskedastic- and cluster-efficient)	
	PSG-exogenous	PSG-endogenous	PSG-exogenous	PSG-endogenous
Log PSG	0.290*	-1.500**	0.389***	-1.530**
	(0.115)	(0.521)	(0.112)	(0.528)
Log DENSITY	0.241***	0.145***	0.226***	0.143***
	(0.027)	(0.039)	(0.030)	(0.042)
PCTSUBURBAN	-0.004*	-0.005***	-0.003*	-0.005***
	(0.001)	(0.002)	(0.001)	(0.001)
PCTURBAN	-0.001	-0.002*	-0.001	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
PCT0T17	0.032**	0.043***	0.030*	0.043**
	(0.010)	(0.012)	(0.013)	(0.015)
PCT18T24	-0.013	-0.003	-0.013	-0.003
	(0.009)	(0.010)	(0.008)	(0.009)
PCT25T44	0.026**	0.025**	0.024**	0.025**
	(0.009)	(0.010)	(0.008)	(0.008)
PCT45T64	0.053**	0.074***	0.055***	0.075***
	(0.017)	(0.020)	(0.017)	(0.016)
Log PCTBLK	0.149***	0.161***	0.150***	0.161***
	(0.014)	(0.016)	(0.016)	(0.016)
Log PCTHISP	0.061*	0.016	0.041	0.015
	(0.024)	(0.030)	(0.032)	(0.033)
Log PCTFEM18	-0.118	0.393	-0.153	0.406
	(0.192)	(0.251)	(0.262)	(0.339)
Log PCTEDUC	-0.375***	-0.468***	-0.333**	-0.483***
	(0.097)	(0.108)	(0.116)	(0.124)
Log PCTTRANS	-0.023	0.008	0.012	0.009
	(0.055)	(0.058)	(0.065)	(0.077)
Log MEDHHINC	-0.294	-0.100	-0.297	-0.104
	(0.310)	(0.338)	(0.411)	(0.448)
Log PCTINCLT15K	0.508	0.900**	0.282	0.896*
	(0.265)	(0.302)	(0.340)	(0.402)
Log INEQUALITY	0.392***	0.422***	0.354**	0.433***
	(0.093)	(0.102)	(0.112)	(0.115)
Log PCTPOOR	0.726***	0.519**	0.799***	0.532**
	(0.142)	(0.162)	(0.206)	(0.206)
Log PCTUNEMP	-0.020	0.097	0.060	0.090
	(0.083)	(0.096)	(0.092)	(0.092)
Log PCTVACANT	0.144***	0.153***	0.158***	0.153***
	(0.041)	(0.045)	(0.040)	(0.045)
R^2	0.442	0.333	0.439	0.329
N	1441	1441	1441	1441

Table 3: Log gun homicide equation (continued)
Dependent variable: Log CRGMUR

	2-step Efficient GMM (heteroskedastic-efficient)		2-step Efficient GMM (heteroskedastic- and cluster-efficient)	
	PSG-exogenous	PSG-endogenous	PSG-exogenous	PSG-endogenous
J statistic	$\chi^2(2)=13.29$**	$\chi^2(1)=0.254$	$\chi^2(2)=8.18$*	$\chi^2(1)=0.162$
p-value	0.0013	0.6145	0.0168	0.6874
C statistic	$\chi^2(1)=13.01$***		$\chi^2(1)=8.01$**	
p-value	0.0003		0.0047	
95% confidence interval for log PSG				
Wald		[-2.52, -0.48]		[-2.56, -0.50]
Anderson-Rubin		[-3.05 , -0.34]		[-3.35 , -0.06]
First-stage regression:				
F statistic		26.2		12.2
Log PCTDEM88		-0.075***		-0.075**
		(0.022)		(0.027)
RGUNMAG		0.046***		0.046***
		(0.008)		(0.013)

Notes: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.
Standard errors in parentheses.
Excluded instruments are log PCTDEM88 and RGUNMAG.
R^2 is the within-R^2 (see text).
All equations are estimated with 50 state fixed effects.

Table 4: Log nongun homicide equation
Dependent variable: Log CRNGMUR

	2-step Efficient GMM (heteroskedastic-efficient)		2-step Efficient GMM (heteroskedastic- and cluster-efficient)	
	PSG-exogenous	PSG-endogenous	PSG-exogenous	PSG-endogenous
Log PSG	-0.009	-0.490	0.013	-0.559
	(0.107)	(0.478)	(0.111)	(0.552)
Log DENSITY	0.135***	0.110***	0.134***	0.107**
	(0.022)	(0.033)	(0.021)	(0.034)
PCTSUBURBAN	0.003*	0.002	0.003*	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
PCTURBAN	0.003**	0.002*	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
PCT0T17	0.017	0.020*	0.017	0.019*
	(0.009)	(0.010)	(0.009)	(0.009)
PCT18T24	-0.034***	-0.032***	-0.031***	-0.033***
	(0.008)	(0.009)	(0.007)	(0.008)
PCT25T44	0.004	0.003	0.007	0.005
	(0.008)	(0.008)	(0.008)	(0.009)
PCT45T64	-0.012	-0.007	-0.009	-0.010
	(0.015)	(0.016)	(0.014)	(0.014)
Log PCTBLK	0.152***	0.154***	0.150***	0.154***
	(0.013)	(0.013)	(0.010)	(0.010)
Log PCTHISP	0.062**	0.046	0.061*	0.045
	(0.023)	(0.027)	(0.030)	(0.032)
Log PCTFEM18	0.073	0.229	0.042	0.235
	(0.181)	(0.238)	(0.148)	(0.256)
Log PCTEDUC	-0.196*	-0.219*	-0.232**	-0.230**
	(0.085)	(0.089)	(0.080)	(0.084)
Log PCTTRANS	0.057	0.060	0.079	0.071
	(0.048)	(0.048)	(0.050)	(0.050)
Log MEDHHINC	0.324	0.395	0.334	0.377
	(0.274)	(0.283)	(0.308)	(0.303)
Log PCTINCLT15K	0.916***	1.011***	1.020***	1.099***
	(0.231)	(0.251)	(0.259)	(0.270)
Log INEQUALITY	0.214*	0.220*	0.242*	0.228*
	(0.092)	(0.093)	(0.101)	(0.103)
Log PCTPOOR	0.355**	0.309*	0.320*	0.240
	(0.124)	(0.133)	(0.138)	(0.155)
Log PCTUNEMP	0.116	0.151	0.130	0.178
	(0.083)	(0.091)	(0.095)	(0.104)
Log PCTVACANT	0.045	0.044	0.047	0.039
	(0.038)	(0.038)	(0.047)	(0.046)
R^2	0.435	0.425	0.434	0.422
N	1409	1409	1409	1409

Table 4: Log nongun homicide equation (continued)
Dependent variable: Log CRNGMUR

	2-step Efficient GMM (heteroskedastic-efficient)		2-step Efficient GMM (heteroskedastic- and cluster-efficient)	
	PSG-exogenous	**PSG-endogenous**	**PSG-exogenous**	**PSG-endogenous**
J statistic	$\chi^2(2)=1.99$	$\chi^2(1)=0.898$	$\chi^2(2)=2.43$	$\chi^2(1)=1.37$
p-value	0.3703	0.3432	0.2964	0.2413
C statistic	$\chi^2(1)=1.07$		$\chi^2(1)=0.977$	
p-value	0.3005		0.3230	
95% confidence interval for log PSG				
Wald		[-1.43, 0.45]		[-1.64, 0.52]
Anderson-Rubin		[-1.71, 0.62]		[-2.10, 1.06]
First-stage regression:				
F statistic		24.2		10.9
Log PCTDEM88		-0.071**		-0.071*
		(0.022)		(0.027)
RGUNMAG		0.045***		0.045**
		(0.008)		(0.013)

Notes: * p<0.05, ** p<0.01, *** p<0.001.
Standard errors in parentheses.
Excluded instruments are log PCTDEM88 and RGUNMAG.
R^2 is the within-R^2 (see text).
All equations are estimated with 50 state fixed effects.

Table 5: Log total homicide equation
Dependent variable: Log CRMUR

	2-step Efficient GMM (heteroskedastic-efficient)		2-step Efficient GMM (heteroskedastic- and cluster-efficient)	
	PSG-exogenous	PSG-endogenous	PSG-exogenous	PSG-endogenous
Log PSG	0.160	-1.023*	0.225*	-1.023*
	(0.095)	(0.437)	(0.094)	(0.450)
Log DENSITY	0.212***	0.147***	0.200***	0.147***
	(0.021)	(0.033)	(0.023)	(0.032)
PCTSUBURBAN	-0.001	-0.003*	-0.001	-0.003*
	(0.001)	(0.001)	(0.001)	(0.001)
PCTURBAN	-0.000	-0.001	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
PCT0T17	0.024**	0.030**	0.022**	0.030**
	(0.009)	(0.010)	(0.008)	(0.010)
PCT18T24	-0.021**	-0.015	-0.020**	-0.015
	(0.007)	(0.008)	(0.007)	(0.008)
PCT25T44	0.020**	0.020*	0.023**	0.020*
	(0.008)	(0.008)	(0.007)	(0.008)
PCT45T64	0.034*	0.048**	0.037**	0.048***
	(0.014)	(0.016)	(0.012)	(0.013)
Log PCTBLK	0.154***	0.161***	0.153***	0.161***
	(0.012)	(0.013)	(0.013)	(0.013)
Log PCTHISP	0.061**	0.027	0.041	0.027
	(0.020)	(0.025)	(0.021)	(0.024)
Log PCTFEM18	0.070	0.427	0.008	0.426
	(0.166)	(0.218)	(0.185)	(0.255)
Log PCTEDUC	-0.273***	-0.317***	-0.290**	-0.317***
	(0.081)	(0.087)	(0.088)	(0.092)
Log PCTTRANS	-0.004	0.008	0.021	0.008
	(0.047)	(0.049)	(0.058)	(0.064)
Log MEDHHINC	-0.084	0.032	-0.126	0.032
	(0.261)	(0.276)	(0.326)	(0.343)
Log PCTINCLT15K	0.584**	0.848***	0.502	0.848**
	(0.223)	(0.256)	(0.279)	(0.317)
Log INEQUALITY	0.309***	0.325***	0.340***	0.325***
	(0.080)	(0.085)	(0.088)	(0.090)
Log PCTPOOR	0.599***	0.441**	0.674***	0.441*
	(0.117)	(0.137)	(0.156)	(0.175)
Log PCTUNEMP	0.115	0.207*	0.162*	0.208*
	(0.073)	(0.086)	(0.081)	(0.085)
Log PCTVACANT	0.080*	0.074	0.099**	0.074*
	(0.036)	(0.038)	(0.031)	(0.030)
R^2	0.527	0.463	0.525	0.463
N	1458	1458	1458	1458

Table 5: Log total homicide equation (continued)
Dependent variable: Log CRMUR

	2-step Efficient GMM (heteroskedastic-efficient)		2-step Efficient GMM (heteroskedastic- and cluster-efficient)	
	PSG-exogenous	PSG-endogenous	PSG-exogenous	PSG-endogenous
J statistic	$\chi^2(2)=8.51$	$\chi^2(1)<0.001$	$\chi^2(2)=6.01$	$\chi^2(1)<0.001$
p-value	0.0142	0.9907	0.0497	0.9922
C statistic	$\chi^2(1)=8.51$		$\chi^2(1)=6.01$	
p-value	0.0035		0.0143	
95% confidence interval for log PSG				
Wald		[-1.88, -0.17]		[-1.90, -0.14]
Anderson-Rubin		[-2.35, -0.02]		[-2.64, 0.21]
First-stage regression:				
F statistic		24.8		11.3
Log PCTDEM88		-0.072***		-0.072*
		(0.022)		(0.027)
RGUNMAG		0.045***		0.045***
		(0.008)		(0.013)

Notes: * p<0.05, ** p<0.01, *** p<0.001.
Standard errors in parentheses.
Excluded instruments are log PCTDEM88 and RGUNMAG.
R^2 is the within-R^2 (see text).
All equations are estimated with 50 state fixed effects.

REFERENCES

Anderson, Thomas W., and Herman Rubin. 1949. "Estimation of the parameters of a single equation in a complete system of stochastic equations." *Annals of Mathematical Statistics* 91: 46-63.

Andrews, Donald W.K., and James H. Stock. 2005. "Inference with weak instruments." NBER Technical Working Paper 313, August. http://www.nber.org/papers/T0313.

Audit Bureau of Circulations. 1993. Supplementary Data Report, covering county paid circulation for gun and related sports magazines. Schaumburg, IL: Audit Bureau of Circulations.

Azrael, Deborah, Philip J. Cook and Matthew Miller. 2004. "State and local prevalence firearms ownership: measurement, structure, and trends." *Journal of Quantitative Criminology* (forthcoming).

Basmann, Robert L. 1960. "On finite sample distributions of generalized classical linear identifiability test statistics." *Journal of the American Statistical Association* 55:650-659.

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2003. "Instrumental variables and GMM: Estimation and testing." *The Stata Journal* 3:1-31.

Baum, Christopher, Mark E. Schaffer and Steven Stillman. 2005. "IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation". http://ideas.repec.org/c/boc/bocode/s425401.html.

Bordua, David J. 1986. "Firearms ownership and violent crime: a comparison of Illinois counties." Pp. 156-88 in *The Social Ecology of Crime*, edited by James M. Byrne and Robert J. Sampson. N.Y.: Springer-Verlag.

Bordua, David J., and Alan J. Lizotte. 1979. "Patterns of legal firearms ownership: a cultural and situational analysis of Illinois counties." *Law and Policy Quarterly* 1:147-75.

Bound, John, David A. Jaeger, and Regina Baker. 1995. "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak." *Journal of the American Statistical Association* 90:443-450.

Brearley, H.C. 1932. *Homicide in the United States*. Chapel Hill: University of North Carolina Press.

Brill, Steven. 1977. *Firearm Abuse: A Research and Policy Report*. Washington, D.C.: Police Foundation.

Clarke, Ronald V., and Pat Mayhew. 1988. "The British gas suicide story and its criminological implications." Pp 79-116 in *Crime and Justice*, Vol. 10, edited by Michael Tonry and Norval Morris. Chicago: University of Chicago Press.

Clotfelter, Charles T. 1981. "Crime, disorders, and the demand for handguns." *Law & Policy Quarterly* 3:425-446.

Cook, Philip J. 1976. "A strategic choice analysis of robbery." Pp. 173-87 in *Sample Surveys of the Victims of Crime*, edited by Wesley Skogan. Cambridge: Ballinger.

____. 1982. "The role of firearms in violent crime." Pp.236-91 in *Criminal Violence*, edited by Marvin E. Wolfgang and Neil Alan Weiner. Beverly Hills: Sage.

Cook, Philip J., and Jens Ludwig. 1997. *Guns in America*. Washington, D.C.: Police Foundation.

____. 2000. *Gun Violence: The Real Costs*. N.Y.: Oxford.

____. 2003. "Guns and burglary." Pp. 74-118 in *Evaluating Gun Policy*, edited by Jens Ludwig and Philip J. Cook. Washington, D.C.: Brookings Institution Press.

Cox, Nick. 2003. "WINSOR: Stata module to Winsorize a variable". http://ideas.repec.org/c/boc/bocode/s361402.html.

Cragg, J. 1983. "More efficient estimation in the presence of heteroskedasticity of unknown form." *Econometrica* 51:751-763.

Dufour, J.M. 2003. "Identification, weak instruments and statistical inference in econometrics." CIRANO Working Paper 2003s-49. http://www.cirano.qc.ca/pdf/publication/2003s-49.pdf.

Duggan, Mark. 2001. "More guns, more crime." *Journal of Political Economy* 109:1086-1114.

Duwe, Grant. 2000. "Body-count journalism." *Homicide Studies* 4:364-399.

Fisher, Joseph C. 1976. "Homicide in Detroit: The role of firearms." *Criminology* 14:387-400.

Hahn, Jinyong and Jerry Hausman. 2003. "Weak instruments: Diagnoses and cures in empirical econometrics." *American Economic Review* 93(2):118-125.

Hall, A.R. and F.P.M Peixe. 2000. "A consistent method for the selection of relevant instruments." Econometric Society World Congress 2000 Contributed Papers. http://econpapers.repec.org/paper/ecmwc2000/0790.htm

Hansen, Lars. 1982. "Large sample properties of generalized method of moments estimators." *Econometrica* 50: 1029-1054.

Hayashi, Fumio. 2000. *Econometrics*. Princeton: Princeton University Press.

Hemenway, David, and Matthew Miller. 2000. "Firearm availability and homicide rates across 26 high-income countries." *Journal of Trauma* 49:985-988.

Hoskin, Anthony W. 2001. "Armed Americans." *Justice Quarterly* 18:569-592.

Inter-university Consortium for Political and Social Research (ICPSR). 1995. *General Election Data for the United States, 1950-1990*. [Computer file]. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1995.

____. 2000. Uniform Crime Reporting Program Data [United States]: Offenses known and Clearances by Arrest, 1989 [1990, 1991] [Computer file]. Compiled by U.S Department of Justice, Federal Bureau of Investigation. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 2000.

____. 2001. Uniform Crime Reporting Program Data [United States]: Supplementary Homicide Reports, 1976-1999 [Computer file]. Compiled by U.S Department of Justice, Federal Bureau of Investigation.

ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 2001.

Killias, Martin. 1993. "Gun ownership, suicide, and homicide: an international perspective." Pp. 289-303 in *Understanding Crime: Experiences of Crime and Crime Control*, edited by Anna del Frate, Uglijesa Zvekic, and Jan J. M. van Dijk. Rome: UNICRI.

Killias, Martin, John van Kesteren, and Martin Rindlisbacher. 2001. "Guns, violent crime, and suicide in 21 countries." *Canadian Journal of Criminology* 43:429-448.

Kleck, Gary. 1979. "Capital punishment, gun ownership, and homicide." *American Journal of Sociology* 84:882-910.

____. 1984. "The relationship between gun ownership levels and rates of violence in the United States." Pp. 99- 135 in *Firearms and Violence: Issues of Public Policy*, edited by Don B. Kates, Jr. Cambridge, Mass.: Ballinger.

____. 1997. *Targeting Guns: Firearms and their Control*. N.Y.: Aldine.

____. 2001. "Modes of news media distortion of gun issues." Pp. 173-212 in *Armed: New Perspectives on Gun Control*, edited by Gary Kleck and Don B. Kates, Jr. Amherst, NY: Prometheus.

____. 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, Gary, Chester L. Britt, and David J. Bordua. 2000. "The emperor has no clothes: using interrupted time series designs to evaluate social policy impact." *Journal on Firearms and Public Policy* 12:197-247.

Kleck, Gary, and Miriam DeLone. 1993. "Victim resistance and offender weapon effects in robbery." *Journal of Quantitative Criminology* 9:55-82.

Kleck, Gary, and Don B. Kates. 2001. *Armed: New Perspectives on Gun Control*. Amherst, NY: Prometheus.

Kleck, Gary, and Tomislav Kovandzic. 2001. "The impact of gun laws and gun levels on crime rates." Paper presented at the annual meetings of the American Society of Criminology, Atlanta, GA.

Kleck, Gary, and Karen McElrath. 1991. "The effects of weaponry on human violence." *Social Forces* 69:669-92.

Kleck, Gary, and E. Britt Patterson. 1993. "The impact of gun control and gun ownership levels on violence rates." Journal *of Quantitative Criminology* 9:249-288.

Lester, David. 1985. "The use of firearms in violent crime." *Crime & Justice* 8:115-20.

____. 1988b. "Firearm availability and the incidence of suicide and homicide." *Acta Psychiatrica Belgium* 88:387-393.

____. 1996. "Gun ownership and rates of homicide and suicide." *European Journal of Psychiatry* 10:83-85.

Lott, John R., Jr. 2000. *More Guns, Less Crime*. 2nd edition. Chicago: University of Chicago Press.

Marvell, Thomas B., and Carlisle E. Moody, Jr. 1991. "Age structure and crime rates: the conflicting evidence." *Journal of Quantitative Criminology* 7(3):237-273.

McDowall, David. 1986. "Gun availability and robbery rates: a panel study of large U.S. cities, 1974-1978." *Law & Policy* 8:135-48.

____. 1991. "Firearm availability and homicide rates in Detroit, 1951-1986." *Social Forces* 69:1085-1099.

McDowall, David, and Colin Loftin. 1983. "Collective security and the demand for handguns." *American Journal of Sociology* 88:1146-1161.

Miller, Matthew, Deborah Azrael, and David Hemenway. 2002. "Firearm availability and unintentional firearm deaths, suicide, and homicide among 5-14 year olds." J*ournal of Trauma Injury, Infection, and Critical Care* 52:267-275.

Moody, Carlisle E. 2001. "Testing for the effects of concealed weapons laws: Specification errors and robustness." *Journal of Law and Economics* 44:799-813.

Moody, Carlisle E., and Thomas B. Marvell. 2003. "Pitfalls of using proxy variables in studies of guns and crime". SSRN working paper. http://ssrn.com/abstract=473661.

Moody, Carlisle E., and Thomas B. Marvell. 2005. "Guns and crime." *Southern Economic Journal* 71:720-736.

Newton, George D., and Franklin Zimring. 1969. *Firearms and Violence in American Life. A Staff Report to the National Commission on the Causes and Prevention of Violence*. Washington, D.C.: U.S. Government Printing Office.

Okoro, Catherine A., David E. Nelson, James A. Mercy, Lina S. Balluz, Alex E. Crosby, and Ali H. Mokdad. 2005. "Prevalence of household firearms and firearm-storage practices in the 50 states and the District of Columbia." *Pediatrics* 116:e370-e376.

Pagan, A.R. and D. Hall. 1983. "Diagnostic tests as residual analysis." *Econometric Reviews* 2(2):159-218.

Phillips, Llad, Harold L. Votey, and John Howell. 1976. "Handguns and homicide." *Journal of Legal Studies* 5:463-78.

Rice, Douglas C., and David D. Hemley. 2002. "The market for new handguns." *Journal of Law and Economic*s 45:251-265.

Ruddell, Rick, and G. Larry Mays. 2005. "State background checks and firearms homicides." *Journal of Criminal Justice* 33:127-136.

Sargan, D. 1958. "The estimation of econometric relationships using instrumental variables." *Econometrica* 26:393-415.

Sommers, Paul M. 1984. "Letter to the Editor." *New England Journal of Medicine* 310:47-8.

Sorenson, Susan B, and Richard A. Berk. 2001. "Handgun sales, beer sales, and youth homicide, California, 1972-1993." *Journal of Public Health Policy* 22:183-197.

Southwick, Lawrence, Jr. 1997. "Do guns cause crime? Does crime cause guns?: a Granger test." *Atlantic Economic Journal* 25:256-273.

____. 1999. "Guns and justifiable homicide: deterrence and defense." *Saint Louis University Public Law Review* 18:217-246.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental variables regression with weak instruments." *Econometrica* 65:557-586.

Stock, James H., Jonathan H. Wright and Motohiro Yogo, 2002. "A survey of weak instruments and weak identification in generalized method of moments." *Journal of Business and Economic Statistics* 20(4):518:529.

Stock, J.H., and M. Yogo. 2002. "Testing for weak instruments in linear IV regression." NBER Technical Working Paper 284. http://www.nber.org/papers/T0284.

Stolzenberg, Lisa, and Stewart J. D'Alessio. 2000. "Gun availability and violent crime." *Social Forces* 78:1461-1482.

Tark, Jongyeon, and Gary Kleck. 2004. "Resisting crime: the effects of victim action on the outcomes of crimes." *Criminology* 42(4):861-909.

U.S. Bureau of the Census. 1994. County and City Data Book, 1994. Washington, D.C.: U.S. Government Printing Office.

U.S. Bureau of the Census. 1990. "Census 1990 Summary File 3 (SF3) – Sample Data, Table P006 Urban and Rural". Retrieved 7 February 2005 from U.S. Census http://factfinder.census.gov.

U.S. Bureau of Justice Statistics. 2001. Criminal Victimization in the United States - Statistical Tables. Tables available online at http://www.ojp.usdoj.gov/bjs/abstrat/cvusst.htm. Accessed 12-6-01.

U.S. Federal Bureau of Investigation (FBI). 1990-2000a. Crime in The United States 1989 [-1999] Uniform Crime Reports. Washington, D.C.: U.S. Government Printing Office.

U.S. National Center for Health Statistics. 1997. Special versions of Mortality Detail Files, 1987-1993, with location detail, supplied to third author.

Vieraitis, Lynne M. 2000. "Income inequality, poverty, and violent crime: A review of the empirical evidence." *Social Pathology* 6(1):24-45.

Wooldridge, Jeffrey M. 1995. "Score diagnostics for linear models estimated by two stage least squares." In *Advances in Econometrics and Quantitative Economics: Essays in honor of Professor C. R. Rao*, eds. G. S. Maddala, P. C. B. Phillips, and T. N. Srinivasan. Pp. 66–87. Cambridge, MA: Blackwell Publishers.

____. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

____. 2003. "Cluster-sample methods in applied econometrics." *American Economic Review* 93(2):133-138.

Wright, James D., and Peter H. Rossi. 1986. Armed *and Considered Dangerous*. New York: Aldine.

Wright, James D., Peter H. Rossi, and Kathleen Daly. 1983. *Under the Gun: Weapons, Crime and Violence*. New York: Aldine.

Zimring, Franklin E., and Gordon Hawkins. 1997. *Crime is Not the Problem*. N.Y.: Oxford.

Zivot, Eric, Richard Startz, and Charles R. Nelson. 1998. "Valid confidence intervals and inference in the presence of weak instruments". *International Economic Review* 39(4):1119-1144.

BIOGRAPHICAL SKETCHES

Tomislav Kovandzic is Assistant Professor of Criminal Justice in Department of Justice Sciences at the University of Alabama at Birmingham. His current research interests include criminal justice policy and gun-related violence. His most recent articles have appeared in *Criminology & Public Policy*, *Criminology*, and *Homicide Studies*. He received his Ph.D. in Criminology from Florida State University in 1999.

Mark E. Schaffer is Professor of Economics at Heriot-Watt University, Edinburgh, U.K. His research interests include economic reform in transition and developing economies, and the implementation of econometric estimators. He is a member of the Executive Committee of the Association for Comparative Economic Studies (ACES) and is an Associate Editor of the Stata Journal.

Gary Kleck is Professor of Criminology and Criminal Justice at Florida State University. His research focuses on the links between guns and violence and the deterrent effects of punishment. He is the author of four books, including *Point Blank*, which won the 1993 Hindelang Award, and, most recently, *Armed* (Prometheus, 2001).