J Quant Criminol
DOI 10.1007/s10940-012-9185-7

ORIGINAL PAPER

Estimating the Causal Effect of Gun Prevalence on Homicide Rates: A Local Average Treatment Effect Approach

Tomislav Kovandzic · Mark E. Schaffer · Gary Kleck

Abstract
Objective This paper uses a "local average treatment effect" (LATE) framework in an attempt to disentangle the separate effects of criminal and noncriminal gun prevalence on violence rates. We first show that a number of previous studies have failed to properly address the problems of endogeneity, proxy validity, and heterogeneity in criminality. We demonstrate that the time series proxy problem is severe; previous panel data studies have used proxies that are essentially uncorrelated in time series with direct measures of gun relevance.
Methods We adopt instead a cross-section approach: we use US county-level data for 1990, and we proxy gun prevalence levels by the percent of suicides committed with guns, which recent research indicates is the best measure of gun levels for crosssectional research. We instrument gun levels with three plausibly exogenous instruments: subscriptions to outdoor sports magazines, voting preferences in the 1988 Presidential election, and numbers of military veterans. In our LATE framework, the estimated

T. Kovandzic (✉)
Program in Criminology, University of Texas at Dallas, P.O Box 830688, GR 31, Richardson, TX 75083-0688, USA
e-mail: tkovan@utdallas.edu

M. E. Schaffer
Heriot-Watt University, Edinburgh EH14 4AS, UK
e-mail: m.e.schaffer@hw.ac.uk

M. E. Schaffer
CEPR, London, UK

M. E. Schaffer
IZA, Bonn, Germany

G. Kleck
College of Criminology and Criminal Justice, Florida State University, Tallahassee, FL 32306-1127, USA
e-mail: gkleck@fsu.edu

Published online: 11 October 2012

impact of gun prevalence is a weighted average of a possibly negative impact of noncriminal gun prevalence on homicide and a presumed positive impact of criminal gun prevalence.
Results We find evidence of a significant negative impact, and interpret it as primarily "local to noncriminals", i.e., primarily determined by a negative deterrent effect of noncriminal gun prevalence. We also demonstrate that an ATE for gun prevalence that is positive, negative, or approximately zero are all entirely plausible and consistent with our estimates of a significant negative impact of noncriminal gun prevalence.
Conclusions The policy implications of our findings are perhaps best understood in the context of two hypothetical gun ban scenarios, the first more optimistic, the second more pessimistic and realistic. First, gun prohibition might reduce gun ownership equiproportionately among criminals and noncriminals, and the traditional ATE interpretation therefore applies. Our results above suggest that plausible estimates of the causal impact of an average reduction in gun prevalence include positive, nil, and negative effects on gun homicide rates, and hence no strong evidence in favor of or against such a measure. But it is highly unlikely that criminals would comply with gun prohibition to the same extent as noncriminals; indeed, it is virtually a tautology that criminals would violate a gun ban at a higher rate than noncriminals. Thus, under the more likely scenario that gun bans reduced gun levels more among noncriminals than criminals, the LATE interpretation of our results moves the range of possible impacts towards an increase in gun homicide rates because the decline in gun levels would primarily occur among those whose gun possession has predominantly negative effects on homicide.

Keywords Crime · Homicide · Gun levels · Endogeneity

Introduction

Guns are heavily involved in violence in America, especially homicide. In 2005, 68 % of homicides were committed by criminals armed with guns (US Federal Bureau of Investigation 2000). Probably an additional 100,000–150,000 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, 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 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.[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; Duggan 2001).

While gun possession among aggressors in violent incidents may serve to increase the probability of a victim's death, gun possession and defensive use among victims may

[1] See, e.g., Sloan et al. (1990), Killias (1993), and Zimring and Hawkins (1997). 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 et al. 2001, pp. 436, 440).

reduce their chances of injury or death. Individual-level research (e.g., Kleck and McElrath 1991; 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. Criminals usually cannot visually distinguish people carrying concealed weapons from other people, or residences with gun-owning occupants from other residences, and so deterrent effects would not be limited to gun owners (Kleck 1988; Kleck and Kates 2001, pp. 153–154; Lott 2000). Because the protective effects of gun ownership may spill over to nonowners, 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. 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. The result is that empirical researchers face a classic problem of endogeneity bias: unless it is successfully addressed, what is asserted to be the impact of gun levels on crime rates will also include the impact of crime rates on gun levels, and estimates of the former will therefore typically be biased upwards. Second, measurement of gun levels is subject to well-documented problems (Kleck 2004). Given the shortcomings in the available direct measures of gun levels, researchers have commonly used proxy measures instead. But proxies are also problematic. At the most basic level, a proxy for gun prevalence should be demonstrated to be "valid": correlated with direct measures of gun levels. Panel data studies face a particular problem here, since the time-series variation in the chosen proxy must be demonstrably correlated with the time-series variation in gun levels, a much tougher requirement than cross-sectional correlation.

A third problem that has hitherto been ignored in the empirical literature in this area is heterogeneity in criminality: the effect of gun prevalence depends on who holds guns, criminals or noncriminals, and this may vary across localities. This gap in the literature is particularly surprising given the policy debate. Empirical researchers have implicitly focused on estimating an "average treatment effect" (ATE) of gun prevalence on crime rates, i.e., the impact of a change in gun prevalence that is randomly distributed across the population. Policy interventions (e.g., gun control laws), however, will typically have different impacts on the prevalence of guns among criminals and noncriminals. Even if the first two problems can be addressed and an ATE estimated consistently, it would therefore be of limited use in helping to assess the net impact of a change in gun policy on crime rates. Separate estimates of the criminal and noncriminal effects of gun prevalence on crime would be more useful, since they could be combined with information on how a new policy would affect criminal versus noncriminal gun prevalence.

This study uses a "local average treatment effect" (LATE) framework to try to disentangle the separate effects of criminal and noncriminal gun prevalence on violence rates. We first show that previous studies have not properly addressed the problems of endogeneity, proxy validity, or heterogeneity in criminality. We demonstrate that the time series proxy problem is severe—previous panel data studies have used proxies that are essentially uncorrelated in time series with direct measures of gun relevance. We adopt instead a cross-section approach: we use US county-level data for 1990, and we proxy gun prevalence levels by the percent of suicides committed with guns, which recent research indicates is the best measure of gun levels for cross-sectional research (Kleck 2004). We instrument gun levels with three plausibly exogenous instruments: subscriptions to outdoor sports magazines, voting preferences in the 1988 Presidential election, and numbers of military veterans. In our LATE framework, the estimated impact of gun prevalence is a

weighted average of a possibly negative impact of noncriminal gun prevalence on homicide and a presumed positive impact of criminal gun prevalence. We find evidence of a significant negative impact, and interpret it as primarily "local to noncriminals", i.e., primarily determined by a negative deterrent effect of noncriminal gun prevalence.

The paper is organized as follows. "Prior Research" critically reviews previous research on the gun-homicide relationship. "Valid and Invalid Proxies" considers in detail the problem of finding an adequate proxy for gun prevalence, and demonstrates that previous studies using longitudinal data have failed to employ valid proxies. In "Modeling Criminal/Noncriminal Heterogeneity and LATE Estimation" we set out how we model heterogeneity of criminality and discuss estimation and specification testing in an IV/GMM/LATE framework. "Data and Model Specification" describes the data and model specification that we use. Estimation results are presented in "Estimation Results", and "Conclusions" concludes.

Prior Research

Modeling Framework

The basic model throughout this literature is one in which crime rates—here homicide, h_i—in a locality i are a function of the level of gun prevalence g_i, $h_i = h(g_i)$. The key question is the sign and magnitude of the derivative $\partial h_i/\partial g_i$, i.e., the sign and size of the impact of gun prevalence on crime.

We distinguish between various channels through which gun prevalence could influence homicide rates:

(a) Criminality: guns are an offensive technology, and can be used by criminals to facilitate criminal activity. Even if gun prevalence had no effect on the frequency of criminal activity, it could still increase the fraction of crimes that resulted in death, due to the greater lethality of guns relative to other weapons. Increased general gun prevalence means, ceteris paribus, more criminals have more access to this technology, and hence commit more homicides. $\partial h_i/\partial g_i > 0$.

(b) Deterrence: guns are also a defensive technology, and can be used by noncriminals or criminals to deter crime. The more likely the potential victim is armed, the less likely a criminal is to attack. Increased general gun prevalence also means, ceteris paribus, more potential victims have access to this technology, and hence deter more homicides. $\partial h_i/\partial g_i < 0$.

(c) Self-defense: the more likely the potential victim is armed, the more likely the victim is to attempt to use a gun to disrupt an attack. Gun use by potential victims reduces the likelihood of a crime victim being injured and, ceteris paribus, the number of homicides. $\partial h_i/\partial g_i < 0$.

These effects can also interact: for example, the stronger the positive self-defense effect (c) because victims shoot their attackers, the stronger can be the negative deterrent effect (b). Whether the net effect of increased gun prevalence on homicide is positive, negative, or null, is therefore essentially an empirical question.

An estimating equation is obtained by linearizing the model and adding an error term, e.g., in a cross-section of localities, the data generating process (DGP) is typically modeled as

$$h_i = \beta_0 + \beta_1 g_i + u_i \quad (1)$$

Researchers also include a range of control variables X_i as exogenous regressors, which we omit here for simplicity of exposition. The parameter of interest is β_1, the impact of gun levels on the homicide rate.

Macro-level studies of the impact of gun prevalence on crime rates number in the dozens.[2] The conclusions of these studies have been contradictory, with some (e.g., Duggan 2001, Cook and Ludwig 2004, 2006) finding a significant positive association between crime or violence rates and some measure of gun ownership, and others (e.g., Kleck and Patterson 1993, Moody and Marvell 2005) finding a null or negative relationship. All these studies, however, can be criticized on the methodological grounds mentioned earlier: endogeneity, proxy validity, and heterogeneity in criminality.

The Endogeneity Problem

Numerous macro-level studies have found effects of crime rates on gun levels (e.g., Kleck 1979, 1984; McDowall and Loftin 1983; Kleck and Patterson 1993; Duggan 2001; Rice and Hemley 2002), and individual-level survey evidence directly indicates that people buy guns in response to higher crime rates (Kleck 1997, pp. 74–79). We can represent this in the simple model above by writing gun prevalence as an increasing function of homicide rates, $g_i = g(h_i)$, $\partial h_i/\partial g_i > 0$. This reverse causality creates a potential endogeneity bias: OLS estimation of Eq. (1) would generate an estimate of β_1 that is biased upwards. A similar problem arises when an unmeasured characteristic of locales (e.g., low "social capital" or a "violent culture") is associated with both high crime rates and high gun prevalence: if sc_i denotes social capital, then $g_i = g(sc_i)$, $\partial g_i/\partial sc_i < 0$ and $h_i = h(sc_i)$, $\partial h_i/\partial sc_i < 0$. The effect is again to cause OLS estimates of β_1 to suffer from an upward endogeneity bias.

The guns-crime studies that have tried to address the endogeneity problem have used one of two approaches: instrumental variables (IV) techniques on cross-section data, and panel data methods.

Critics of the IV approach applied to cross-section data have argued that it is difficult or impossible to find plausible instruments that are both correlated with the gun level measure and uncorrelated with the error term. The most-often cited such study, Kleck and Patterson (1993), used a sample of 170 cities and a proxy for gun levels, instrumented with the rate of subscriptions to gun-related magazines and the state hunting license rate. Duggan (2001, p. 1095, n. 10) and Cook and Ludwig (2004, p. 10, n. 6) question the assumption that these instruments are exogenous, speculating that they may be correlated with unmeasured city-level correlates of violent crime. Kleck and Patterson did not report a test of instrument exogeneity in their study, and did not account for econometric problems such as heteroskedasticity that are common in cross-sectional analysis.

The second and more recent approach to dealing with the endogeneity problem—e.g., Duggan (2001), Moody and Marvell (2005), and Cook and Ludwig (2004, 2006)—has been to use panel data to control for all time-invariant unobserved heterogeneity through the use of fixed effects or first differences. Thus first-differencing Eq. (1),

$$\Delta h_{it} = b_1 \Delta g_{it} + \Delta u_{it} \quad (2)$$

[2] See our 2005 CEPR discussion paper for a discussion and list of 30 such studies, and Kleck (1997), Chapter 7 for a more extensive review of the pre-1997 research.

eliminates from the error term u_{it} any time-invariant omitted characteristics of locales; because these omitted variables are purged from the error term, they no longer contribute to endogeneity bias in OLS estimates of Eq. (2). The panel approach can be combined with the notion of Granger causality to try to establish whether changes in *past* gun levels help predict changes in current crime rates by replacing the current period change in gun prevalence Δg_{it} with one or more lags. Moody and Marvell (2005) estimate using both first differences and fixed effects on state level panel data; they find no effect of gun prevalence on homicide rates. Duggan (2001) uses first differences on county-level panel data; Cook and Ludwig (2004, 2006) use fixed effects on county-level panel data. Both Duggan and Cook-Ludwig find a positive impact of gun prevalence on homicide rates.

All of the aforementioned studies have relied heavily or exclusively on proxies for gun prevalence. It turns out this poses severe problems for the panel data approach; indeed, we show below that the results reported by these studies are largely uninterpretable. By contrast, the proxy problem is not nearly so serious for the cross-sectional approach.

The Proxy Validity Problem

Like the endogeneity problem, the proxy problem is straightforward to state. Typically, data on gun levels are sparse, noisy, or simply unavailable. Researchers have responded by using a diverse set of proxy measures (Kleck, 2004). In cross-section, this gives a feasible estimating equation

$$h_i = \mathrm{b}_0 + \mathrm{b}_1 p_i + e_i \tag{3}$$

where p_i is the proxy for gun prevalence g_i. The relationship between g_i and p_i is given by

$$p_i = \delta_0 + \delta_1 g_i + v_i \tag{4}$$

where δ_1 is some positive parameter and v_i is measurement error. Since a consistent estimate of b_1 will be a consistent estimate of the quantity (β_1/δ_1), estimates of b_1 can be used to make inferences about the sign of β_1, and after calibration (estimation of, or evidence on, the magnitude of δ_1), of the magnitude of β_1.

The proxy problem is that the variable p_i must be demonstrably valid, i.e., highly correlated with the true measure of gun prevalence g_i. Kleck (2004) shows, however, that most of the measures used in prior studies have poor validity. The main exception is the percent of suicides committed with guns (PSG), which correlates strongly in cross-section with direct survey measures across cities, states, and nations. Other authors (e.g., Azrael et al. 2004; Moody and Marvell 2005; Cook and Ludwig 2004, 2006) have also concluded that PSG is currently the best available proxy. Our approach in this paper is cross-section estimation with gun levels proxied by PSG.

Whereas the proxy problem is solvable in cross-section studies, it is essentially fatal for longitudinal studies. The problem is that panel data estimation relies on *changes* in gun prevalence to identify *changes* in homicide rates. In a panel data study, therefore, demonstrating the validity of a proxy for gun levels means showing that *changes* in g_i are correlated with *changes* in p_i. In a recent survey, Kleck (2004) shows that all ten proxies of aggregate measures of gun ownership in the US used in various studies show no significant intertemporal correlation with direct survey measures. We show in Sect. 3 that the three aforementioned panel studies between them used two gun level proxies, PSG and GAR (subscription rates to *Guns & Ammo* magazine), that show little or no time series correlation with directly available survey measures of gun prevalence.

The "Heterogeneity in Criminality" Problem

The "heterogeneity in criminality" problem follows from some widely accepted stylized facts about crime in general and homicide in particular. A relatively large fraction of homicides in the US is accounted for by a small portion of the population, namely criminals. Criminals also account for a disproportionately large share of homicide victims. Unlike the endogeneity and proxy problems, the complications posed by heterogeneity in the degree of criminality have been largely ignored in the empirical literature on gun prevalence and crime. In particular, all the studies cited above have implicitly attempted to estimate what is known in the program evaluation literature as the "average treatment effect" (ATE) or "average causal effect". In Sect. 4 we use a simple model to show that estimation of Eq. (1) or Eq. (3) by least squares, as in Duggan (2001) and Cook and Ludwig (2004, 2006), does not in general produce a consistent estimate of the ATE even in the absence of endogeneity or proxy problems. Instrumental variables methods face a different but related problem: estimation of Eq. (1) or Eq. (3) by IV, as in Kleck and Patterson (1993) or our own previous work (Kovandzic et al. 2012), produces a consistent estimate, not of the ATE, but of the "local average treatment effect" or LATE (Imbens and Angrist 1994; Angrist and Imbens 1995).

More fundamentally, we believe that the implicit focus by previous studies on the ATE is misplaced. The ATE is of limited use in the context of gun policy, since policy interventions in this area—e.g., gun control measures—aim to have, by design, differential impacts on criminal and noncriminal gun prevalence. More useful would be results about the signs and magnitudes of the separate criminal and noncriminal effects of gun prevalence on homicide, because this information could be combined with information about the likely impacts of a policy on criminal and noncriminal gun prevalence to forecast the overall impact of the policy on crime rates.

The approach we take in this paper is to return to the cross-section setting, where the basic validity of the proxy used—PSG—has been well established, and the problem of endogeneity bias is well understood. We expand the standard modeling framework in the literature to accommodate criminal/noncriminal heterogeneity, and adopt a LATE approach to estimation and interpretation of our results. Our results provide evidence in particular about the sign and magnitude of the noncriminal effect of gun prevalence on homicide.

Valid and Invalid Proxies

A proxy must be strongly correlated with the proxied variable—in our application, gun prevalence g—in the same dimension as the analysis using the proxy. These correlations should be checked using direct measures of g. In this section, we show that PSG is a valid cross-sectional proxy: it is strongly correlated in a cross-section of US states with direct measures of gun prevalence. We also show that neither PSG nor GAR are valid longitudinal proxies: changes in both PSG and GAR have little or no correlation with changes in direct measures of gun prevalence. We then briefly consider the Moody-Marvell, Duggan and Cook-Ludwig studies, and demonstrate how their claims of proxy validity are mistaken.

By "strongly correlated" we mean literally a high correlation coefficient or R^2; a large proportion of the variance of p must be shared with g. By "in the same dimension" we mean that if the model is cross-sectional, the proxy must be also be correlated in cross-

section with gun prevalence; and if the model is longitudinal, changes over time in the proxy must be correlated with changes over time in gun prevalence. In other words, the same approach to estimation (cross-section or longitudinal) should be used for both the model and the proxy check.

It is important to note that a correlation that is statistically significantly different from zero is *not* adequate evidence that a proxy is valid. A proxy p that had an R^2 of 1 % in a regression of p on g would be an invalid proxy, even if we could reject the null hypothesis of zero correlation at some high level of statistical significance. This is for the obvious reason that a 1 % R^2 means that 99 % of the variation in g has nothing to do with p. The variable g and its proxy must share a large fraction of their variances.

For cross-sectional and first-differences model, where there are no other covariates, R^2 has the standard definition of the squared correlation coefficient between g and p. For the fixed-effects model, the correct R^2 to use is the "within R^2", which is the R^2 from the regression after g and p have been transformed into mean-deviation form. Some authors have included year dummies in their longitudinal proxy checks. In this case, the correct R^2 with which to assess proxy validity is the "partial R^2" or squared partial correlation coefficient, where the year dummies have been partialed out.[3] The R^2 for the entire estimation would, of course, be inappropriate as a measure of proxy validity since it would include the contribution of the year dummies. The partial correlation coefficient can also be obtained from the t-statistic for g in the regression including the year dummies using the following expression: partial $R^2 = t^2 / (t^2 + df)$, where df is the residual degrees of freedom in the equation (see, e.g., Greene 2008, p. 30). This will be useful below, because some authors do not report partial R^2s but do report t-tests in their proxy validity checks.

Cross-Sectional Proxy Checks

The first dataset we use for our validation checks is that compiled by Moody and Marvell (2005) and consists of state-level observations for the period 1977–98 (with gaps) on PSG, GAR and the percentage of households with a gun (HHG) from the General Social Survey (GSS).[4]

We check the PSG and GAR proxies for validity in cross-section by estimating a "between" or group averages equation:

$$\bar{p}_i = \delta_0 + \delta_1 \bar{g}_i + \eta_i \tag{5}$$

where g is HHG, the GSS measure of gun prevalence at the state level, and a bar indicates time averages for a panel unit. Gaps in the PSG and HHG data mean we can estimate a between equation covering 13 years in the period 1980–98; the GAR between equation covers 14 years over 1977–98. In the estimation of our own model later in the paper we use 1987–93 averages of our proxy p, and so we also report a check of PSG against HHG for this subperiod. The results are reported in Table 1 below.

Table 1, column 1 reports the between regression of 1980–98 state average PSG on 1980–98 state average HHG for 44 states plus Washington, DC. The R^2 is 69 %, which is very high. The corresponding scatterplot is shown in Fig. 1a, and the strong correlation is obvious. These results justify our choice of PSG as a proxy for gun prevalence in our cross-sectional

[3] Greene (2008), pp. 29–30 has a short and clear discussion of how to interpret the partial R^2 for an OLS regression. Note that the "within R^2" for the fixed effects model can be interpreted as the partial R^2 after the panel dummies have been partialed out.

[4] The original dataset is kindly provided on Moody's website, http://cemood.people.wm.edu/research.html.

Table 1 Cross-section proxy checks using GSS data

Proxy	(1) PSG	(2) PSG	(3) GAR
Estimator/period	Between (group averages), 1980–98	Between (group averages), 1987–93	Between (group averages), 1977–98
Coefficients/R^2			
GSS HHG (SE)	0.698*** (0.071)	0.660*** (0.070)	0.261*** (0.078)
R^2	0.691	0.694	0.206
N	45	41	45

A constant is included in all specifications. * 10 %, ** 5 %, *** 1 % significance

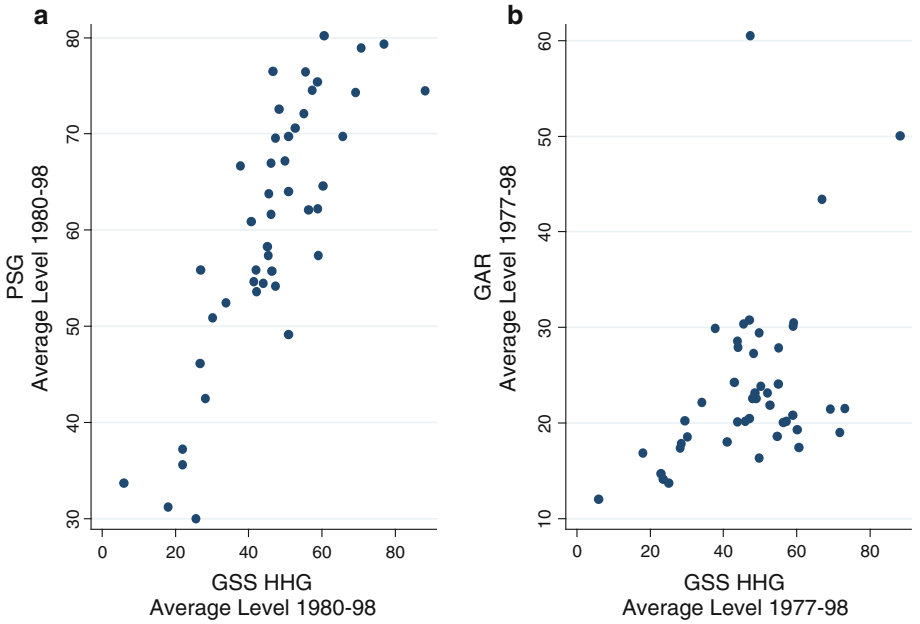

Fig. 1 Cross-section proxy checks. **a** PSG versus HHG, cross-section, **b** GAR versus HHG, cross-section

model. The results for the 1987–93 sample (column 2) are almost identical: the R^2 is again 69 %. Log specifications generate similarly high R^2s.

Column 3 of Table 1 reports the between regression for Duggan's (2001) alternative proxy, GAR, for the period 1977–98. The R^2 is 21 %, much lower than that for PSG. The corresponding scatterplot is shown in Fig. 1b. The weaker correlation is very apparent.

We conclude from the above that PSG is a satisfactory proxy for our cross-sectional analysis and much superior to GAR.

Longitudinal Proxy Checks

Longitudinal proxy validity can be checked using either a first-differences estimation or a fixed-effects estimation:

Table 2 Longitudinal proxy checks using GSS data, PSG

Proxy	(1) PSG	(2) PSG	(3) PSG	(4) PSG
Estimator/period	Fixed effects, 1980–98	Fixed effects, 1980–98	First-differences, 1985–94	First-differences, 1985–94
Coefficients/R^2				
GSS HHG (SE)	0.028*** (0.011)	0.020 (0.011)	0.012 (0.014)	0.007 (0.014)
Time dummies	No	Yes	No	Yes
R^2/Partial R^2	0.015	0.008 (partial)	0.004	0.001 (partial)
N	498	498	234	234

A constant is included in all specifications. * 10 %, ** 5 %, *** 1 % significance

Table 3 Longitudinal proxy checks using GSS data, GAR

Proxy	(1) GAR	(2) GAR	(3) GAR	(4) GAR
Estimator/period	Fixed effects, 1977–98	Fixed effects, 1977–98	First-differences, 1985–94	First-differences, 1985–94
Coefficients/R^2				
GSS HHG (SE)	−0.016** (0.008)	−0.002 (0.007)	−0.002 (0.007)	0.000 (0.004)
Time dummies	No	Yes	No	Yes
R^2/Partial R^2	0.009	0.000 (partial)	0.000	0.000 (partial)
N	531	531	234	234

A constant is included in all specifications. * 10 %, ** 5 %, *** 1 % significance

$$\Delta p_i = \delta_1 \Delta g_i + \eta_i \tag{6a}$$

$$(\boldsymbol{p}_{it} - \bar{\boldsymbol{p}}_i) = \delta_1 (\boldsymbol{g}_{it} - \bar{\boldsymbol{g}}_i) + \eta_{it} \tag{6b}$$

where we have written the fixed-effects estimation in mean-deviation form.

We report proxy checks using first differences (FD) and fixed effects (FE) for both PSG and GAR. Because some authors have included year dummies in their own longitudinal proxy checks, we also report specifications with and without year dummies. When the year dummies are included, we report the partial R^2. The checks of the longitudinal validity of PSG and GAR are reported in Tables 2 and 3. Note that the first-differences specifications have a smaller number of observations because of gaps in the GSS HHG series.

The longitudinal validity results for PSG are abysmal. In both the FE and FD specifications, the R^2 is less than 2 %. The longitudinal validity of GAR is even worse: not only is the R^2 also 1 % or less, but the correlation with HHG is sometimes negative. Log specifications (not reported here) generate the same results: near-zero R^2s for both PSG and GAR in both FE and FD specifications. In Figs. 2 and 3 we show the corresponding scatterplots of PSG and GAR versus HHG in mean-deviation form and in first differences. The absence of correlation is obvious.

Note that in the FE proxy check for PSG without time dummies, the coefficient on HHG is significant at the 1 % level. But this is *not* evidence that PSG is a good proxy for HHG; the tiny R^2 of 1.5 % means that over 98 % of the variation in PSG has nothing to do with

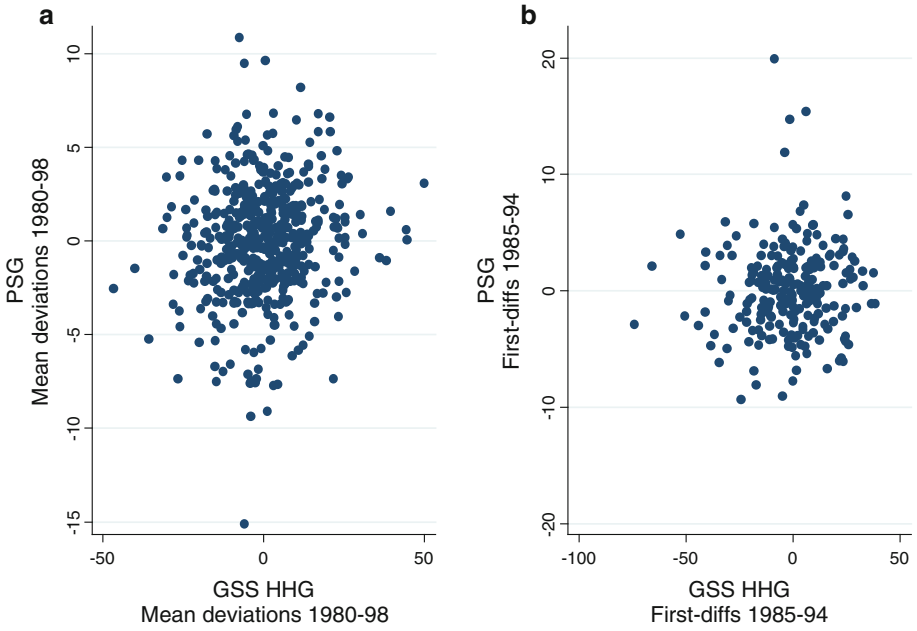

Fig. 2 Longitudinal proxy checks, PSG. **a** PSG versus HHG, mean dev, **b** PSG versus HHG, first-diff

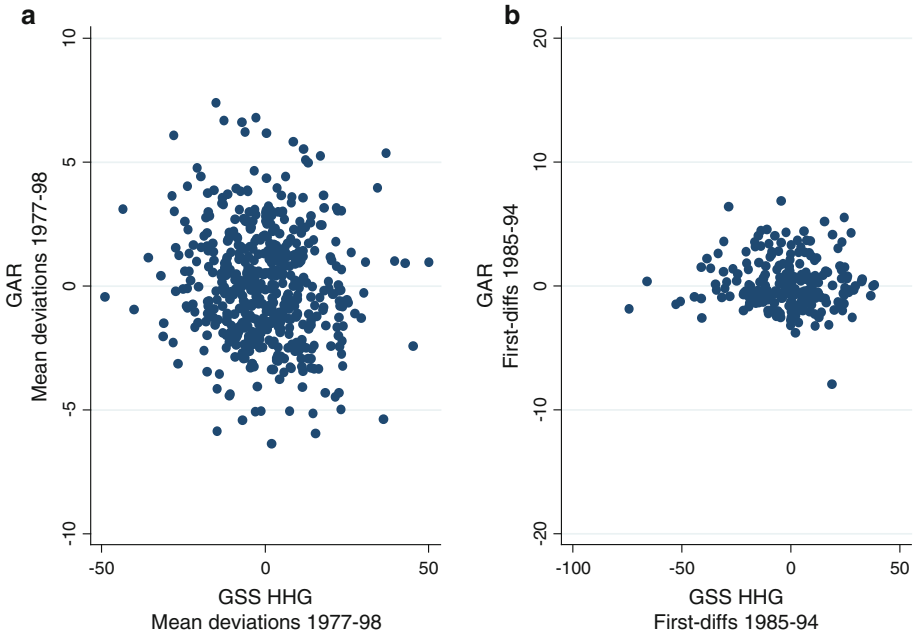

Fig. 3 Longitudinal proxy checks, GAR. **a** GAR versus HHG, mean dev, **b** GAR versus HHG, first-diff

HHG. This "significance trap" is one that a number of authors working in this area have fallen into, as we show below.

We conclude, as does Kleck (2004), that based on the GSS survey data neither PSG nor GAR has any demonstrable validity as a proxy in a longitudinal analysis. We now reconsider the three aforementioned panel studies and their respective claims of proxy validity.

Duggan (2001)

Duggan (2001) estimates a model in first differences using county-level data with log (GAR) as the proxy for gun prevalence. His check of GAR versus GSS gun ownership state-level data is reported in his Table 3, p. 1093. Column 3 in this table is a between (cross-section) estimation of log(HHG) on log(GAR), with an R^2 of 0.384 (somewhat higher than what we report, probably because Duggan weights by the number of GSS interviewees per state). But this cross-sectional correlation is irrelevant for justifying his longitudinal model. Column 4 is a FE specification, again using weighted state-level; Duggan also includes time dummies. He reports a total R^2 for the equation of 71.2 %, but this includes the effects of the dummies and hence is irrelevant.

Duggan points to the significant coefficient on log(GAR) in his FE specification as evidence of the time series validity of the proxy, but this claim is simply mistaken.[5] As we note above, a significant non-zero correlation says *nothing* about how much of the variance in changes in log(GAR) is shared by changes in log(HHG). In fact, we can deduce how much variance is shared from Duggan's reported results: the implied partial R^2 for his weighted FE estimation is only 2 %.[6] In other words, based on both Duggan's own reported results and our own checks in Table 3, log(GAR) is a hopelessly inadequate proxy for a longitudinal study.

Cook and Ludwig (2004)

Cook and Ludwig (2004, 2006) estimate a fixed-effects model using county-level data for the 200 largest counties in the US for the period 1980–99, with log(PSG) as their proxy for gun prevalence. In the published version of their paper, Cook and Ludwig (2006, p. 380) justify the use of PSG (FSS in their terminology) as a proxy as follows: "We ran panel regressions of GSS-based estimates of gun prevalence against [FSS] … The estimated coefficients of our GSS measures on FSS are in every case significantly positive…." Cook and Ludwig thus fall prey to the same trap as Duggan—they mistakenly cite statistical significance as evidence of proxy validity.

In Table 2 of the 2004 working paper version of their paper, they report the regression coefficients and standard errors for regressions of various GSS measures of gun prevalence on PSG. These are panel data regressions with 9 Census Division fixed effects covering the

[5] "The significantly positive estimate of 0.354 suggests that this magazine's sales are a valid measure both of the level and of the change in gun ownership within an area." (Duggan 2001, p. 1093) We note that his claim that this regression is evidence that GAR is correlated with the *level* of g is also mistaken; a FE specification cannot justify cross-sectional validity. But the mistake of confusing statistical significance with proxy validity is the key error.

[6] Duggan reports a coefficient of 0.354 with a standard error of 0.114, giving a t-statistic of 3.105. The estimation has 488 observations with 45 state dummies and 13 estimated parameters (log(HHG) and 12 year dummies), leaving 430 residual degrees of freedom. Using the formula provided in the previous section, the implied partial R^2 is $t^2/(t^2 + df) = 3.1052/(3.1052 + 430) = 2$ %.

period 1980–98 (14 years of data, N = 126), with and without year dummies. They note the level of statistical significance on PSG but do not report or discuss the relevant R^2 measures. We can easily deduce these using the formula above: the partial R^2s corresponding to their reported proxy validation regressions using the GSS measure of HHG with and without year dummies are 6 % and 5 %, respectively.[7] The R^2s for PSG in these fixed effects estimations are therefore almost as abysmal as those we reported earlier. Cook and Ludwig's own proxy checks, like the checks we report in Table 2, provide evidence that PSG is a very poor longitudinal proxy and in no way can be used to justify their fixed effect model estimations.

Moody and Marvell (2005)

Moody and Marvell's (2005) results include estimates of a model in first differences using state-level panel data and a hybrid explanatory variable: GSS handgun prevalence where available, and imputed values based on PSG to fill the gaps in the GSS series. Half of the total observations of this hybrid variable are imputations. The imputed values come from the predicted values from an OLS regression of GSS gun prevalence on PSG. Since their estimating equation is in first differences, this procedure is valid only if the proxy is valid in first differences. We have shown above (Table 2) that GSS total gun prevalence is uncorrelated with PSG in first differences. The results with GSS handgun prevalence are identical (in logs and first differences, the R^2 is a negligible 0.1 %).

Moody and Marvell (1991, p. 723 and Table 1) justify their use of PSG to impute missing values in their GSS gun data by referring to the high correlation of PSG the GSS series. Their mistake is to report a correlation coefficient for the pooled dataset—in other words, the R^2 from a pooled OLS regression. This R^2 is large entirely because of the cross-sectional element; the longitudinal correlation, as we have seen, is nil. Because of this flaw in their imputation procedure, half of the observations in these regressions are spurious.

Evidence from the BRFSS

The GSS has been the most commonly used direct measure for attempts to validate proxies of gun prevalence, but it has major shortcomings. Most importantly for the present study, the intertemporal variation in GSS data disaggregated to the state level is extremely noisy, mostly because of the small number of observations: since only about 1,400 people are asked the gun questions in a typical year, GSS samples for any one state in any one year average only about 30 persons. Thus the standard deviation for the first-difference of HHG using the 1980–98 dataset is 18.1, versus just 4.4 for PSG. This will have been a contributing factor to the near zero correlation in the longitudinal dimension between GSS measures and gun prevalence proxies.

An alternative direct measure of gun prevalence is available from the CDC's Behavioral Risk Factor Surveillance System (BRFSS). This is a large telephone survey which for 3 years—2001, 2002 and 2004—gathered data on firearms ownership from between 200,000 and 300,000 thousand households across the US. The survey enables precise state-level measures of gun prevalence to be calculated for these years. We have calculated

[7] The t-statistics for the significance of PSG in these two regressions are 1.108/0.417 = 2.657 and 0.905/0.355 = 2.549, respectively. The corresponding partial R^2s are therefore $2.657^2/(2.657^2 + 103) = 6$ % and $2.549^2/(2.549^2 + 117) = 5$ %.

Table 4 Cross-section and longitudinal proxy checks using BRFSS data, PSG

Proxy	(1) PSG	(2) PSG	(3) PSG	(4) PSG
Estimator/period	Between (group averages), 2001–04	Fixed effects, 2001–04	Fixed effects, 2001–04	First-differences, 2001–02
Coefficients/R^2				
BRFSS HHG (SE)	0.714*** (0.066)	−0.195 (0.281)	0.025 (0.330)	0.424 (0.376)
Time dummies	No	No	Yes	No
R^2/Partial R^2	0.705	0.005	0.000 (partial)	0.026
N	51	151	151	50

A constant is included in all specifications. * 10 %, ** 5 %, *** 1 % significance

HHG using the BRFSS.[8] The standard deviation of the 2001–02 change in HHG is 2.0, much smaller than the GSS-based measure; the standard deviation of the 2001–02 change in PSG is 4.5, essentially the same as for the earlier period.

In Table 4 below, we report the three proxy validation regressions of PSG versus BRFSS HHG: a between or group averages regression, a fixed effects regression, and a first-differences regression for 2001–02. The R^2 for the between estimation is 70 %, essentially the same as when the GSS HHG measure is used, and confirms again that PSG is a valid proxy in cross-section. The R^2s for the fixed effects and first-differences estimations are, once again, abysmally low, at 1 and 3 % respectively. This is so in spite of the fact that we are now using a much less noisy direct measure of gun prevalence.

This is further evidence that PSG is an invalid proxy for gun prevalence in the time-series dimension. Reducing the noise in the direct gun prevalence measure used in validation has revealed no correlation whatsoever over time. The validity of PSG in cross-section, however, is once again supported.

These results from the BRFSS have implications beyond that of proxy validation. They show that the annual variation in gun prevalence at the state level is uncorrelated with suicide method. This is so even though the gun prevalence measure is based on many thousands of observations. If the correlation between suicide method and gun prevalence is this weak in the time-series dimension, it raises serious questions about whether the panel-type strategy employed by Duggan et al. and Cook and Ludwig could work even if direct and precise measures of gun prevalence were available. The time series variation in gun prevalence may generally be so low that causal effects cannot be captured by this kind of regression analysis (Kleck 2004).

In short, of the two approaches available to addressing the endogeneity problem—IV methods applied to cross-section data, and panel data methods—only the first is feasible; the latter is hopeless. In the next section we show how IV methods can be interpreted in the context of the criminal/noncriminal heterogeneity.

[8] We use the answer to the BRFSS survey question, "Are any firearms now kept in or around your home?" We exclude respondents who answered "don't know/not sure" or who refused to answer. Non-respondents account for less than 5 % of the total sample in any year. We used the survey weights provided with the BRFSS; the unweighted measures generate very similar results.

Modeling Criminal/Noncriminal Heterogeneity and LATE Estimation

Estimating Strategy

Our strategy is to use an IV/GMM/LATE framework to obtain and interpret point estimates of, and bounds on, the impact of gun prevalence on homicide rates. We use county data in levels with state fixed effects and we proxy gun prevalence by the percentage of suicides committed with a gun, PSG. In this section we first show that with an underlying population composed of a mixture of criminals and noncriminals and a continuous treatment variable, gun prevalence, OLS estimation does not generate a consistent estimate of the ATE or indeed an estimate that is readily interpretable, even in the ideal circumstances where the investigator can directly observe gun prevalence and there is no endogeneity problem. By contrast, IV estimation with a valid (exogenous) instrument has a straightforward interpretation: it identifies a LATE parameter that is a weighted average of the criminal and noncriminal impacts of gun prevalence on homicide, where the weights depend on the strength of the correlations between the instrument, on the one hand, and criminal and noncriminal gun prevalence, on the other.

Because we have more than one instrument for our gun level proxy, a test of overidentifying restrictions is available. This specification test has both a LATE interpretation (due to Angrist and Imbens 1995) as well as a traditional IV/GMM interpretation. These are, in fact, simply different interpretations: as we discuss below, in both cases the test can be seen as a "vector-of-contrasts" test. The null hypothesis is that the instruments are all identifying the same parameter. Rejection of the null suggests either that the instruments are identifying different LATE parameters, and/or some or all of the instruments are not exogenous and are generating different (biased) parameter estimates. The vector-of-contrasts interpretation shows when the test will have power to detect specification problems, and how to interpret tests of the exogeneity of specific instruments.

We also present two possible "upper bound" arguments that may be available for interpreting our results. First, we discuss how to make use of priors about the possible failure of instrument exogeneity in making inferences about the impact of gun prevalence. If an instrument is positively correlated with the error term, the estimated coefficient will be biased upwards, making it an upper bound on the actual impact. Second, we consider the possibility that the identified LATE parameter could be driven primarily by noncriminal gun prevalence effects. We show that this gives us an upper bound for the noncriminal gun prevalence effect on homicide.

The LATE framework we develop has more general applications, and may be used where the researcher has grouped data, groups are composed of two types of individuals that respond differently to a treatment or policy, and the researcher observes variation across groups in the treatment and outcomes but not the composition of the groups. For example, a researcher may want to estimate the impact on educational outcomes of an intervention at the school level (say, use of a teaching method or technology, or a category of spending), where data are available at the school level, students at a school are naturally categorized into two groups (students who do/don't have learning difficulties or whose mother tongue is/isn't English), but the student composition by school is not observable to the researcher. Another example would be analysis of local labor markets where it is natural to dichotomize workers (e.g., skilled vs. unskilled or high school dropouts vs. others) but the composition of the local labor force is not directly observed.

Modeling Criminal/Noncriminal Heterogeneity

In modeling heterogeneity of criminality in gun homicide, we need to accommodate several stylized facts about gun homicide in the US. These are:

- Empirical studies show that most murders are committed by a relatively small number of "extremely aberrant individuals" (Kates and Mauser 2007, p. 666, who provide a useful recent survey of the evidence). Criminal homicides by previously noncriminal persons are very rare.
- Criminals acquire guns to facilitate criminal activity. In particular, they may intend to use their guns to commit crimes, or to protect themselves. Thus in the interview study of convicted felons by Wright and Rossi (1986), self-protection was the single most important reason for owning a gun, cited by 58 % as a "very important" reason for acquiring their most recent handgun; 28 % cited the need to use a gun in committing crimes as "very important".
- Most homicide victims are "criminals" in the sense of having criminal records or being regularly engaged in illegal activity. For example, in a review of the 112 homicide cases that took place in St. Louis in 2002, the St. Louis Police Department stated that 90 % of suspects and 79 % of victims had a felony criminal history (Decker et al. 2005, pp. 88–89).
- A substantial minority of homicide victims are nevertheless not criminals, at least as far as one can tell from arrest records.
- Homicides by noncriminals of other noncriminals are relatively rare.[9] Most homicides of noncriminals are by criminals.
- Gun use by potential victims reduces the likelihood of the offender injuring the victim (Kleck and DeLone 1993; Southwick 2000; Tark and Kleck 2004). Victim gun use that results in the killing of the attacker, however, occurs so rarely as to have a negligible impact on the total homicide rate.[10]

We begin our modeling by assuming that the population consists of two categories of people: criminals (C) and noncriminals (NC). Our unit of observation is a county, and we specify our model in shares or fractions of the population of the locality: h_i^C and h_i^{NC} are numbers of (gun) homicides of criminals and noncriminals, respectively, per member of the population of county i; g_i^C and g_i^{NC} are the numbers of criminals and noncriminals holding guns, respectively, per member of county i. We specify two DGPs, one for homicides of criminals (7a) and one for homicides of noncriminals (7b):

$$h_i^C = \beta_0^C + \beta_1^{C-C} g_i^C + \beta_1^{NC-C} g_i^{NC} + u_i^C \tag{7a}$$

$$h_i^{NC} = \beta_0^{NC} + \beta_1^{C-NC} g_i^C + \beta_1^{NC-NC} g_i^{NC} + u_i^{NC} \tag{7b}$$

In both equations, changes in criminal and noncriminal gun prevalence have separate impacts on homicide rates. Thus in Eq. (7a), the causal impact of criminal gun prevalence

[9] We are distinguishing here between (a) "criminals" as a type of person, and (b) "criminality" or "criminal behavior". The terms 'criminal' and 'noncriminal' are, of course simplifications, and can be regarded as shorthand for "persons who commit serious crimes" and "persons who do not commit serious crimes.

[10] For example, in 2000 there were 16,765 total homicides counted in the vital statistics data for the US, but only 137 civilian justifiable homicides with firearms (vital statistics homicide counts include justifiable homicides by civilians but exclude those by police), or 8/10 of one percent of total homicides (US Federal Bureau of Investigation 2000, p. 24).

g_i^C on homicides of other criminals h_i^C is given by β_1^{C-C}, and β_1^{NC-C} is the causal impact of noncriminal gun prevalence g_i^{NC} on homicides of criminals. In Eq. (7b), β_1^{C-NC} and β_1^{NC-NC} are the impacts on homicides of noncriminals by criminal and noncriminal gun prevalence g_i^C and g_i^{NC}, respectively.

The channels identified above through which gun prevalence influences homicide rates—criminality, deterrence and self-defense—can operate in both equations. The stylized facts just cited suggest that β_1^{C-C}will be large and positive: criminals use guns on each other to commit crimes and to defend themselves aggressively against other criminals if attacked, and these outweigh any possible negative deterrent effects of criminal gun prevalence. We expect β_1^{C-NC} to be positive, though not as large: criminals use guns to kill noncriminals, but not as often as they use them to kill each other. The parameter β_1^{NC-C} should be positive but, given the rarity with which victim gun use results in the death of the criminal attacker, very small.

The stylized facts tell us less about what to expect for β_1^{NC-NC}. The parameter β_1^{NC-NC} includes the negative effect of deterrence of attacks by criminals; it also includes gun homicides of NCs by NCs, a kind of "criminality" effect (a), though perpetrated by what we are designating as "noncriminals". The stylized facts suggest that this last effect should be positive but quite small. As discussed above, however, there has been considerable debate and no clear consensus in the literature on the existence or size of the deterrent effects. Thus β_1^{NC-NC} may be positive (gun homicide "crimes of passion" dominate), negative (deterrence of criminal attacks dominates), or somewhere in between.

We do not separately observe rates of homicide of criminals and noncriminals; like preceding studies, we investigate instead the impact of gun prevalence on the total homicide rate. Aggregating (7a) and (7b), we obtain

$$h_i = \beta_0 + \beta_1^C g_i^C + \beta_1^{NC} g_i^{NC} + u_i \tag{7c}$$

where $h_i \equiv (h_i^C + h_i^{NC})$, $u_i \equiv (u_i^C + u_i^{NC})$, $\beta_0 \equiv (\beta_0^C + \beta_0^{NC})$, $\beta_1^C \equiv (\beta_1^{C-C} + \beta_1^{C-NC})$, $\beta_1^{NC} \equiv (\beta_1^{NC-C} + \beta_1^{NC-NC})$. We expect the overall impact of criminal gun prevalence on total homicide to be positive, i.e., $\beta_1^C > 0$, because of the positive criminality impacts of criminal guns on both criminal and noncriminal victims.[11] The impact of noncriminal gun prevalence β_1^{NC} is ambiguous, and depends on the relative magnitudes of negative deterrent effects of g_i^{NC} on homicides of noncriminals and criminals versus the positive effects on homicides by noncriminals of criminals (self-defense) and the positive effects on homicides by noncriminals of noncriminals (gun crimes of passion).

The Average Treatment Effect (ATE) and OLS Estimation

The average treatment effect or ATE is, at first glance, a natural way to summarize the overall impact of gun prevalence. When, because localities are heterogeneous, the impact of the treatment (gun prevalence) varies by locality, the ATE β_1^{ATE} is the average impact $E(\beta_{1i})$. In our model of heterogeneity in criminality, the ATE takes the following form (see Appendix 2):

[11] The use of guns by criminals for self-defense against other criminals is likely to have a negative effect, but is probably insufficient to offset the positive criminality effects.

$$\beta_1^{\mathrm{ATE}} = \left(\frac{\mu^{\mathrm{C}} - \mathrm{cov}\left(g_i, \frac{g_i^{\mathrm{C}}}{g_i}\right)}{\mu^{\mathrm{C}} + \mu^{\mathrm{NC}}} \right) \beta_1^{\mathrm{C}} + \left(\frac{\mu^{\mathrm{NC}} + \mathrm{cov}\left(g_i, \frac{g_i^{\mathrm{C}}}{g_i}\right)}{\mu^{\mathrm{C}} + \mu^{\mathrm{NC}}} \right) \beta_1^{\mathrm{NC}} \tag{8}$$

where μ^{C} and μ^{NC} are the county means of criminal and noncriminal gun prevalence, $\mathrm{E}(g_i^{\mathrm{C}})$ and $\mathrm{E}(g_i^{\mathrm{NC}})$, respectively. In the special case that total gun prevalence g_i is uncorrelated with its composition, the covariance terms in Eq. (8) are zero and the ATE takes the following simple form:

$$\beta_1^{\mathrm{ATE}} = \frac{\mu^{\mathrm{C}}}{\mu^{\mathrm{C}} + \mu^{\mathrm{NC}}} \beta_1^{\mathrm{C}} + \frac{\mu^{\mathrm{NC}}}{\mu^{\mathrm{C}} + \mu^{\mathrm{NC}}} \beta_1^{\mathrm{NC}} \tag{9}$$

Equations (8) and (9) show that the treatment effect for a randomly selected locality is a weighted average of the criminal and noncriminal effects. If total gun prevalence is uncorrelated with its composition, we have the intuitive result in Eq. (9) where the weights on β_1^{C} and β_1^{NC} are given by the shares of average criminal and noncriminal gun prevalence in total gun prevalence, respectively. The more complicated general form in Eq. (8) accounts for the fact that the mean of a ratio of two random variables is not in general equal to the ratio of the two means.

Unfortunately, separate gun prevalence measures for criminals and noncriminals are not available. Empirical researchers instead must make use of a measure of total gun prevalence $g_i \equiv g_i^{\mathrm{C}} + g_i^{\mathrm{NC}}$ or a proxy for it. Consider the case where we have data for total gun prevalence only, but otherwise the circumstances are the best possible: we have direct observations on gun prevalence rather than a proxy, and gun prevalence is exogenous. The model to be estimated is Eq. (1), repeated here for convenience:

$$h_i = \beta_0 + \beta_1 g_i + u_i \tag{1}$$

Previous researchers have interpreted OLS and IV estimates of β_1 in Eq. (1) as, in effect, estimates of the ATE. Once we explicitly model the heterogeneity in criminality, however, we see that in fact neither method generates a consistent estimate of β_1^{ATE}. We consider first OLS. In Appendix 2 we show that

$$\begin{aligned} \hat{\beta}_1^{\mathrm{OLS}} \xrightarrow{\mathrm{P}} & \frac{\left\{\mathrm{var}(g_i^C) + \mathrm{cov}(g_i^C, g_i^{NC})\right\}}{\mathrm{var}(g_i^C) + \mathrm{var}(g_i^{NC}) + 2\mathrm{cov}(g_i^C, g_i^{NC})} \beta_1^C \\ & + \frac{\left\{\mathrm{var}(g_i^{NC}) + \mathrm{cov}(g_i^C, g_i^{NC})\right\}}{\mathrm{var}(g_i^C) + \mathrm{var}(g_i^{NC}) + 2\mathrm{cov}(g_i^C, g_i^{NC})} \beta_1^{NC} \end{aligned} \tag{10}$$

OLS estimation does not generate an estimate of the ATE even in these most favorable circumstances. The OLS estimator is, like the ATE, a weighted average of the criminal and noncriminal effects, but the weights are determined by the variability, not the levels, of gun prevalence. High variability generates a larger weight (via the variance terms), and a large positive correlation between criminal and noncriminal guns tends to equalize the weights (via the covariance terms). We are not aware of evidence about the cross-sectional or time-series variation in criminal versus noncriminal gun prevalence. There is some evidence to suggest that criminal and noncriminal gun prevalence will be correlated, because most criminals acquire guns as a direct or indirect result of thefts from noncriminals (Wright and Rossi 1986, p. 196). The covariances in Eq. (10) are therefore likely to be positive, but of unknown magnitude. The magnitude of the variances in (10) are also unknown. In short, OLS estimates of β_1 cannot be readily interpreted in terms of the average effect of gun

prevalence. This argument applies to estimates of b_1 using the proxy p, to both the panel and cross-section studies mentioned above, and more broadly to studies in which the dependent variable is some measure of crime other than homicide, e.g., burglary (Cook and Ludwig 2003a).

Although the ATE appears to be a natural summary measure of the impact of gun prevalence, it is actually of limited use in the context of gun policy. Policy interventions in this area—e.g., gun control measures—aim to have, by design, differential impacts on criminal and noncriminal gun prevalence. That is, the measures typically aim to reduce gun prevalence only among criminals, or more among criminals than among noncriminals. More useful would be results about the sign and magnitude of β_1^C or β_1^{NC}, because this information could be combined with information about the likely impacts of a policy on criminal and noncriminal gun prevalence to forecast the overall impact of the policy on crime rates. Such information can be obtained using a LATE framework, to which we now turn.

Local Average Treatment Effect (LATE) Estimation

We now consider IV estimation of Eq. (1), when gun levels are directly observable. Say we have a single instrument Z_i that is correlated with both criminal and noncriminal gun prevalence, but the strength of the correlations may differ:

$$g_i^C = \pi_0^C + \pi_1^C Z_i + \eta_i^C \tag{11}$$

$$g_i^{NC} = \pi_0^{NC} + \pi_1^{NC} Z_i + \eta_i^{NC} \tag{12}$$

We assume that $\pi_1^C, \pi_1^{NC} \geq 0$ and at least one is strictly greater than zero. If total gun prevalence g_i were observable, then Eq. (1) could be estimated by IV. We show in Appendix 2 that IV estimation of Eq. (1) generates an estimator of β_1 that converges in probability to a weighted average of β_1^C and β_1^{NC}:

$$\hat{\beta}_1^{IV} \xrightarrow{P} \frac{\pi_1^C}{\pi_1^C + \pi_1^{NC}} \beta_1^C + \frac{\pi_1^{NC}}{\pi_1^C + \pi_1^{NC}} \beta_1^{NC} \tag{13}$$

The IV estimate of β_1 is an estimate of the "local average treatment effect" or LATE. It is a weighted average of β_1^C and β_1^{NC}, with weights given by the relative strength of the correlation of the instruments with criminal and noncriminal gun prevalence. Note that Eq. (13) illustrates a feature of a LATE estimator, namely that the *definition* of $\hat{\beta}_1^{IV}$ is dependent on the instrument used (Heckman 1997); a different instrument, say Z^A, would, if the weights were different, converge to a different weighted average.

In the polar case of π_1^C=0, when the instrument Z is correlated only with noncriminal gun prevalence, the IV estimator $\hat{\beta}_1^{IV}$ is a consistent estimate of β_1^{NC}, the impact of gun prevalence on homicide "local to noncriminals". Similarly, if π_1^{NC}=0, $\hat{\beta}_1^{IV}$ consistently estimates β_1^C.

Gun prevalence g_i is not observable; instead, we observe only the proxy p_i. The relationship between the gun proxy and unobserved criminal and noncriminal gun prevalence is given by

$$p_i = \delta_0 + \delta_1^C g_i^C + \delta_1^{NC} g_i^{NC} + v_i \tag{14}$$

where we assume that $\delta_1^C, \delta_1^{NC} \geq 0$ and at least one is strictly greater than zero. The estimating equation, again repeated here for convenience, is

$$h_i = \mathrm{b}_0 + \mathrm{b}_1 p_i + e_i \tag{3}$$

and we use IV methods to obtain an estimate of b_1. In Appendix 2 we show that the IV estimator of b_1 also converges in probability to a weighted average of β_1^{C} and β_1^{NC}:

$$\begin{aligned} \hat{\mathrm{b}}_1^{\mathrm{IV}} &\xrightarrow{\mathrm{P}} \frac{\pi_1^{\mathrm{C}}}{\delta_1^{\mathrm{C}}\pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}}\pi_1^{\mathrm{NC}}}\beta_1^{\mathrm{C}} + \frac{\pi_1^{\mathrm{NC}}}{\delta_1^{\mathrm{C}}\pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}}\pi_1^{\mathrm{NC}}}\beta_1^{\mathrm{NC}} \\ &= \frac{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}}{\delta_1^{\mathrm{C}}\pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}}\pi_1^{\mathrm{NC}}}\left\{\frac{\pi_1^{\mathrm{C}}}{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}}\beta_1^{\mathrm{C}} + \frac{\pi_1^{\mathrm{NC}}}{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}}\beta_1^{\mathrm{NC}}\right\} \end{aligned} \tag{15}$$

Equation (15) shows that $\hat{\beta}_1^{\mathrm{IV}}$ and $\hat{\mathrm{b}}_1^{\mathrm{IV}}$ converge to quantities that differ only by a positive scaling factor. This is a useful result: it means we can interpret results for the sign of $\hat{\mathrm{b}}_1^{\mathrm{IV}}$ as results for the sign of the infeasible LATE estimator $\hat{\beta}_1^{\mathrm{IV}}$.

Specification Testing in the LATE Framework

We have more than one instrument available for gun prevalence, and therefore our estimating equation is overidentified. The standard IV/GMM test of overidentifying restrictions is the Sargan-Hansen J test (see, e.g., Hayashi 2000), distributed as χ^2 with degrees of freedom equal to the number of overidentifying restrictions. In our LATE application, the test can reject the null for two reasons: either the different instruments are identifying different LATE parameters $\mathrm{b}_1^{\mathrm{IV}}$, or the instruments are correlated with the error u.

Recall that a feature of the LATE estimator with a single instrument is that the weights on the criminal and noncriminal effects are given by the strength of the correlations between the instrument on the one hand, and criminal and noncriminal gun prevalence on the other Eq. (13). Different instruments can therefore define different LATE parameters $\mathrm{b}_1^{\mathrm{IV}}$, depending on whether the strengths of the correlations with gun prevalence also differ. As Angrist and Imbens (1995, p. 437) point out, the J statistic provides a test of whether the different instruments are identifying the same local impact; in our application, whether they are identifying the same weighted average of criminal and noncriminal gun prevalence. Thus if the instruments identify different LATE parameters, the test will tend to reject the null. Note this will be so even if the instruments are orthogonal to the error, i.e., $\mathrm{E}(Z_i u_i) = 0$. The second reason the null may be rejected is precisely because the orthogonality assumption $\mathrm{E}(Z_i u_i) = 0$ may fail, i.e., because some or all instruments may fail to be exogenous.

Although these two interpretations of the J test appear quite different, they are in fact closely related. The LATE interpretation of the J statistic is as a "vector of contrasts" test: the test rejects if the instruments are identifying different parameters. The second interpretation of J test also has a "vector of contrasts" interpretation.[12] Intuitively, a failure of an instrument to be exogenous generates a bias in the estimate $\hat{\mathrm{b}}_1^{\mathrm{IV}}$ of the parameter $\mathrm{b}_1^{\mathrm{IV}}$, but other instruments that are uncorrelated with the error will generate consistent estimates of $\mathrm{b}_1^{\mathrm{IV}}$. The invalid and valid instruments will effectively identify different parameters, and

[12] This interpretation occasionally appears in textbooks and expositions of IV/GMM. Deaton (1997, p. 112) provides a good example: "the OID [overidentifying restrictions] test tells us whether we would get (significantly) different answers if we used different instruments or different combinations of instruments in the regression. … If we have only k instruments and k regressors, the model is exactly identified, … there is only one way of using the instruments, and no alternative estimates to compare."

the contrast will be detected by the J test. Similarly, even if all the instruments are invalid in the sense that they are correlated with the error, the test can detect this failure if the induced biases in the estimates of b_1^{IV} differ across instruments.

The vector-of-contrasts interpretation also makes clear when the J test will lack the power to reject the null hypothesis of a single LATE parameter and valid instruments when it is false. The J statistic will be small when the null hypothesis is correct; but it will also be small if the biases induced in $\widehat{b}_1^{IV}$ by invalid instruments coincide (i.e., the instruments all identify the same wrong parameter), or if the biases combine with different LATE parameters to generate a small contrast.

The interpretation of our exogeneity and overidentification tests is central to our results. It is important to stress, therefore, that the J test may still have power to detect these problems.[13] This point is easily illustrated by making use of the well-known relationship between the J and GMM distance statistics on the one hand and Hausman vector-of-contrasts tests on the other (Hayashi 2000, pp. 233–234; Newey 1985). In the case when there are only 2 instruments, the J test statistic is numerically identical to a Hausman test statistic that contrasts the estimator using both instruments with an estimator using just one instrument. The intuition is again straightforward: a Hausman test will reject the null hypothesis that the two estimators being contrasted are both consistent so long as the estimators converge to different values. It is not a requirement for one of the two estimators to be consistent for the Hausman test (and therefore the J test) to have power to reject the null.[14]

Note that it follows from this argument that the more unrelated the instruments are to each other, the more credible is a failure to reject the null that the instruments are exogenous, since a failure to reject would require that two unrelated instruments generate the same bias in $\hat{b}_1$.

Choice of Instruments and Signing Bias

We use four instruments for PSG in our estimations: (1) combined subscriptions per 100,000 county population to three of the most popular outdoor/sport magazines (*Field and Stream*, *Outdoor Life*, and *Sports Afield*) in 1993 (OMAG); (2) the percent of the county population voting for the Republican candidate in the 1988 Presidential election (PCTREP88); (3) military veterans per 100,000 county population (VETS); (4) subscriptions per 100,000 county population to *Guns & Ammo* magazine (G&A). We use all in log form. OMAG, PCTREP88 and VETS are theoretically important correlates of gun ownership that are plausibly exogenous and hence suitable instruments; prior research suggests that all three variables are important predictors of gun ownership (Kleck 1997, pp. 70–72; Cook and Ludwig 1997, p. 35). G&A has been shown by Duggan (2001) to be correlated with gun prevalence, but in our view it is less likely to be exogenous, for reasons we discuss below. Failures of exogeneity can, however, still be informative. In particular, if we

[13] Occasionally one finds in the literature the claim is that a test of overidentifying restrictions has power only if there is a subset of instruments that are all valid and identify the model. The claim is incorrect; the correct statement is that the test will have power if there is such a subset, and *might*, or might not, lack power, if there are not enough valid instruments to identify the model.

[14] White (1994) is very clear on this point, for example. "Nor is there any necessity for either estimator to retain consistency in the presence of misspecification. Power is achieved because the estimators chosen have differing probability limits under misspecified alternatives. These alternatives necessarily go beyond those that allow one of the estimators to retain consistency for a certain parameter value." White (1994), p. 274.

can sign the direction of the induced bias in $\widehat{b}_1^{IV}$, we may be able to sign or bound the impact of gun prevalence, as we now discuss before discussing our instruments in detail.

In the case of a single instrument Z, the asymptotic bias in the IV estimator of $\widehat{b}_1^{IV}$, in both the simple case of no heterogeneity in criminality and in the extension to LATE, is $\text{cov}(Z_i, u_i)/\text{cov}(Z_i, p_i)$ (see Appendix 2). If Z is an exogenous instrument and hence uncorrelated with u, the numerator and hence the asymptotic bias are zero. If Z is not exogenous, but we have priors about the likely correlation with the error, then we can sign the bias. For all the instruments used in this paper, $\text{cov}(Z_i, p_i) > 0$, and so the sign of the bias will be given by $\text{sign}\{\text{cov}(Z_i, u_i)\}$. Evidence, for example, that the instrument will be either uncorrelated or positively correlated with the error means that the IV estimator using the instrument can be interpreted as providing an upper bound on a LATE estimate of b_1 (and vice versa for a negative correlation/lower bound).[15] Note that a J test of instrument exogeneity could indicate that one or more instruments are not exogenous, but we may still be able to make inferences about b_1 if our priors allow us to sign the bias.

As noted above, different instruments may be uncorrelated with the error u, but if they identify different LATE parameters b_1, the J test may indicate a failure of exogeneity. Again, we may be able to sign the effect on the estimated coefficient. In this case, we can use the reasonable assumption that $\beta_1^C > \beta_1^{NC}$ and prior beliefs about whether one instrument is identifying an effect more local to criminals when another is identifying one more local to noncriminals. Use of an instrument that is relatively more correlated with g^C will tend to raise (make more positive) the estimated $\widehat{b}_1^{IV}$, and vice versa for an instrument that is relatively less correlated with g^C and more correlated with g^{NC}.

We now consider our instruments in detail. First, OMAG 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). We are not aware of any evidence that pursuit of outdoor sports has a direct impact on homicide rates, nor that it is correlated with unobservables that affect homicide given that we will also condition on county measures of rurality and on population density. Similarly, we do not know of evidence that interest in outdoor sports is correlated more or less strongly with criminal or noncriminal gun levels. The best we can do is rely on a priori reasoning. The most plausible source of bias in the case of OMAG, in our view, is reverse causality: high crime rates may increase interest in gun-related activities. If so, the bias from using OMAG as an instrument will be upwards and the estimate of b_1 will be an upper bound on the LATE parameter. This suggestion is, however, admittedly speculative and we do not regard this bias as very likely, given that OMAG does not include subscriptions to any "gun magazines" per se.

PCTREP88 serves as a measure of political conservatism and hence should be positively correlated with gun ownership.[16] Again, we are not aware of any evidence that voting Republican should either be directly related to homicide rates or correlated more or

[15] For another example of signing bias in a criminological application, see the Cook and Ludwig (2003a, b) study of burglary, Appendix 2.

[16] 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: 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.

less strongly with noncriminal or criminal gun prevalence. We note that Ayres and Donohue (2003, p. 1256) question the exogeneity of voting Republican as an instrument for gun prevalence, based on the argument that it is positively correlated with wealth/income and hence negatively correlated with crime. This is not a powerful objection in our application, however. It is an "omitted variables bias" argument, but we will control for economic conditions using a wide range of economic variables (county median income, inequality, unemployment, poverty rate and vacant housing). Reverse causality again seems to us to be the most plausible source of bias in IV estimation: high crime rates may make residents more likely to vote for "law and order," i.e., Republican candidates. This would again generate an upward bias and make the estimate of b_1 an upper bound on the LATE parameter. Again, however, this is a speculative suggestion and we do not regard this potential bias to be very likely or very large.[17]

Veterans are a subgroup of the population who are relatively more likely to have experience with or to own guns. There is, however, a modest literature indicating that military veterans are not only more likely to own guns in civilian life, but also are more likely to be violent after leaving the military.[18] VETS might therefore correlate more strongly with criminal gun prevalence than would either OMAG or PCTREP88 and hence identify a larger (more positive) LATE estimate of b_1. Thus if VETS is used as an instrument, it is likely to generate an estimate of b_1 that is either unbiased or that provides an upper bound on the LATE impact of gun levels.

The credibility of these three instruments for gun prevalence is enhanced by the fact that they come from different sources and are each capturing different aspects of a population's willingness to hold guns: outdoor sports culture, political conservatism, and prior experience with arms. As argued above, this strengthens the power of the J test to detect any failure of the assumption that one or more is exogenous.

Our fourth instrument, G&A, is less likely either to be exogenous or to identify the same local effect as the other three instruments. The distinct content of *Guns & Ammo* magazine suggests that it appeals to a segment of the population that is more interested in the application of guns in self-defense. This makes G&A a strong candidate for generating an upward reverse-causality bias in the estimated IV coefficient: the same mechanism that leads people to acquire guns in response to high local crime rates also leads to interest in the magazine. More speculatively, it may be correlated with the underlying "violence proneness" of the county population—subscribers may include people who have an interest in violence more generally—and hence relatively more strongly correlated with criminal gun prevalence than the either OMAG or PCTREP88, and possibly also VETS. This would cause G&A to identify a larger LATE estimate of b_1 than the other instruments. We include G&A in our analysis for three reasons. First, it can provide an upper bound for the estimated LATE effect. Second, it can provide a useful demonstration of whether our exogeneity tests have any power to reject the null when our prior beliefs suggest the null—in this case, the validity of G&A as an instrument—is implausible. Lastly, given the widely

[17] An anonymous reviewer also questioned our use of the voting Republican instrument because of reverse causality. The extant evidence, however, only supports an effect of *publicity* about crime (usually concerning especially notorious crimes) on political preference, not changes in actual crime rates. Since studies have repeatedly found publicity (news coverage) about crime to be unrelated to actual crime rates, the evidence about the effects of publicized crime has no bearing on the effects of actual crime rate on voting preferences. We are not aware of any evidence indicating that actual crime rates affect voting preferences.

[18] See, e.g., the case control study by Kleck and Hogan (1999, p. 285), who found that veterans are 2.8 times more likely than nonveterans to commit murder.

cited papers by Duggan (2001, 2003) that make extensive use of G&A as a proxy for gun prevalence, the variable is of interest in its own right.

An "Upper Bound" Argument for the Impact of Noncriminal Gun Prevalence

A small J statistic will suggest that the different instruments are identifying the same weighted average of β_1^C and β_1^{NC}, but will not indicate what these weights are. The only direct evidence we will have will be from the standard first-stage regression used to establish the relevance of instruments with the gun level proxy,

$$p_i = \theta_0 + \theta_1 Z_i + err_i \tag{16}$$

Because noncriminals greatly outnumber criminals in the general population, the strong correlation of PSG vis-à-vis the GSS survey data cited above is evidence that PSG is at least a good proxy for g^{NC}. This says little, however, about how well PSG correlates with g^C. Given the preponderance of noncriminals in the general population, a cautious interpretation of θ_1 would be that it is driven primarily by π_1^{NC}. This applies to all our instruments, even VETS & G&A; although we have some prior suspicions that they might be relatively more correlated with g^C than OMAG and PCTREP88, most veterans and most subscribers to *Guns and Ammo* are, of course, noncriminals.

Since we cannot rule out the possibility that π_1^C is small or zero, it is therefore possible that our estimate of $\widehat{b}_1^{IV}$ could be local to the noncriminal population, i.e., driven primarily by β_1^{NC}. In the polar case that the weight π_1^C on the criminal effect in the LATE estimate is zero, Eq. (15) reduces to

$$\hat{b}_1^{IV} \xrightarrow{P} \frac{\pi_1^{NC}}{\delta^{NC}\pi_1^{NC}} \left\{ \frac{\pi_1^{NC}}{\pi_1^{NC}} \beta_1^{NC} \right\} = \frac{1}{\delta^{NC}} \beta_1^{NC} \tag{17}$$

We can apply an upper bound argument here, should we find that $\widehat{b}_1^{IV} < 0$. Note that the weight on β_1^{NC} in Eq. (15) can be rewritten as:

$$\frac{1}{\delta^{NC}} * \frac{\pi_1^{NC}}{\frac{\delta^C}{\delta^{NC}}\pi_1^C + \pi_1^{NC}} \tag{18}$$

The weight on β_1^{NC} is greatest when $\pi_1^C = 0$; in this case, the weight on β_1^C is zero, the weight on β_1^{NC} is $1/\delta_1^{NC}$, and $\widehat{b}_1^{IV}$ is an estimate of $(\beta_1^{NC} * 1/\delta_1^{NC})$ as in (17). As π_1^C increases, (18) shows that the weight on β_1^{NC} falls, and from (15) we see that the weight on β_1^C increases. Since we expect a priori that $\beta_1^C > 0$, $\pi_1^C > 0$ means that $\widehat{b}_1^{IV}$ will now exceed $(\beta_1^{NC} * 1/\delta_1^{NC})$, i.e., $\widehat{b}_1^{IV}$ scaled by δ_1^{NC} provides an upper bound for β_1^{NC}.

Data and Model Specification

Data

The dependent variables in our model are the gun and nongun homicide rates per 100,000 county population. Homicide rates are averages for the 7 years 1987–1993, thus bracketing the census year of 1990 for which data on many of the control variables were available. Averages of 1987–93 values are more precise measures than values from any single year;

using a 7-year average reduces measurement error (e.g., misclassification of homicides as other kinds of deaths such as suicides), and, because more information is being incorporated, leads to more efficient estimates. We separately assess rates of homicide with and without guns, to provide sharper tests of the hypothesis that gun levels affect homicide rates. The estimations use cross-sectional data for US counties which had a population of 25,000 or greater in 1990, and for which relevant data were available (N = 1,456). These counties account for about half of all US counties but over 90 % of the US population in that year. County-level data were chosen for several reasons. The use of counties provides for a diverse sample of ecological units, including urban, suburban, and rural areas. Counties are more internally homogenous than nations, states, or metropolitan areas, thereby reducing potential aggregation bias. Counties also 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. Finally, county data provide a much larger sample than was used in previous gun level studies that focused on nations, states, or large urban cities.

Although the simple model outlined in the previous section specifies untransformed homicide rates as the dependent variable, in our main estimations we specify homicide rates in logs. Figure 4 illustrates why: the distribution of the total homicide rate is skewed to the right with a number of outlier counties with high homicide rates, whereas the distribution of the log homicide rate is roughly normal and with fewer outliers. The distributions for gun and nongun homicide are similar. The use of logs poses some minor problems, however. Logs mean that the interpretation of the estimating Eq. (7c) as a simple additive aggregation of (7a) and (7b) is not available. There is also a "log of zero" problem, because even though we are using 7-year averages and excluding the smallest

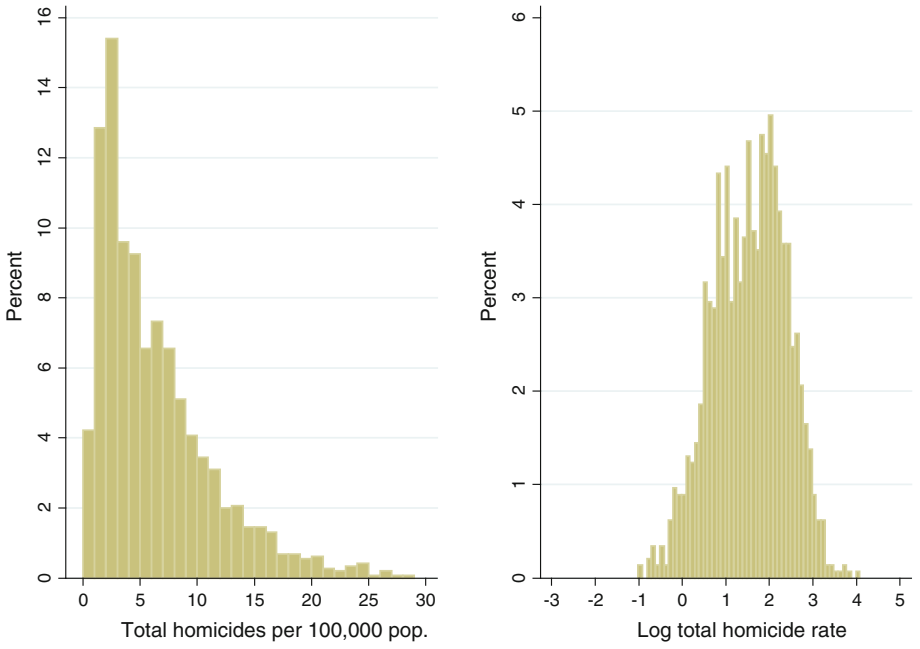

Fig. 4 Distributions of total homicide rate and log total homicide rate. 8 counties with homicide rates greater than 30 per 100,000 are omitted from the *left-hand panel*

counties, a small number of counties have zero murders: of the 1,456 counties in the sample, 20 had no gun murders (about 1 % of the sample) and 39 had no nongun murders (about 3 %); 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. To check the robustness of the results, we also estimate using untransformed homicide rates as the dependent variable.

Our proxy for gun availability in our main regressions is PSG in levels. Suicides in our sample of counties are roughly as prevalent as homicides (the mean county suicide rate is about 13 suicides per 100,000 persons, vs. a mean county homicide rate of about 6.5 per 100,000 persons), and so we again use the average over 1987–93 to calculate PSG. We use PSG in levels primarily because the treatment of criminal/noncriminal heterogeneity is very straightforward in our framework if we maintain the simplifying assumption that gun prevalence is measured in levels. Figure 5 shows histograms of our gun proxy PSG in levels and after logging. There are no problems of skewness or outliers with PSG in levels, but some moderate skewness to the left in the distribution of log PSG. Our approach is to report detailed results using PSG in levels, and to confirm robustness of the findings with regressions using log PSG.

In order to be able to say anything about the practical significance of any nonzero impact of guns on homicide, we 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 GSS. Inspection of the Moody-Marvell data in Fig. 1 suggests that PSG $\approx$ HHG, and closer inspection of the available data confirms this. National gun survey prevalence figures have been available since 1959. The mean HHG for 1959–2003 is 44.2 % while the mean PSG for 1959–2002 (the latest year available) is 54.9 %. These

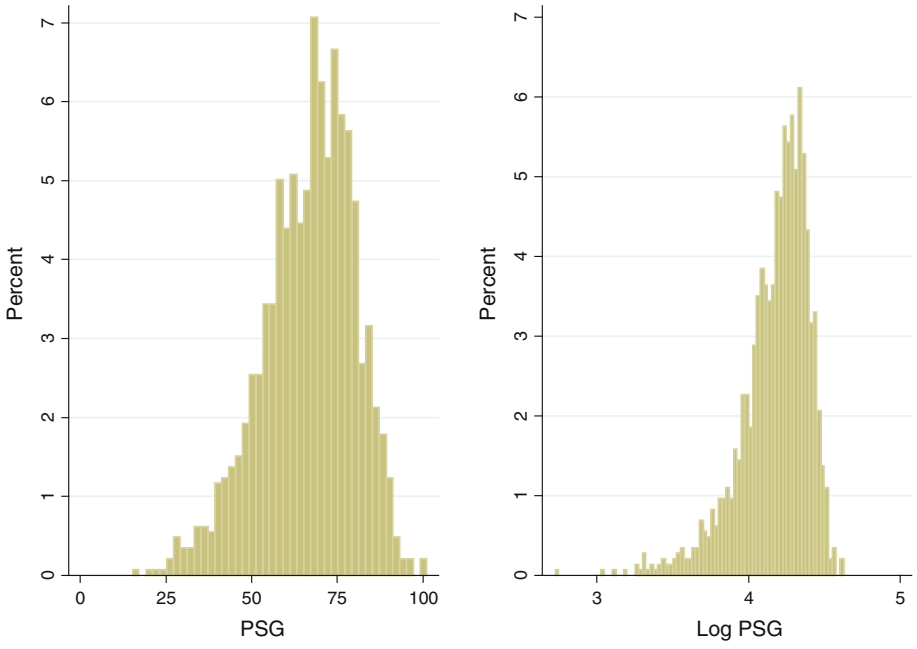

Fig. 5 Distributions of total homicide rate and log total homicide rate

figures imply a value of 54.9/44.2 = 1.24 for the calibration factor δ_1. Neither PSG nor HHG varied greatly during this period, and the use of a different reference period would matter little. A second calibration is available using state-level cross-sectional data. We use survey data from the Centers for Disease Control (CDC) in 2002 (Okoro et al. 2005) and PSG data for 1995–2002 taken from CDC's WONDER service to calibrate PSG to HHG with a simple OLS regression, which yields $\delta_1 = 0.706$ (SE = 0.07, $R^2 = 0.79$, N = 50). Thus in our main estimations using the level of PSG as our proxy, the calibration factor δ_1 by which we should inflate or deflate the estimated coefficient on PSG b_1 so as to obtain an estimate of β_1 should be in the neighborhood of 0.706–1.24. These are, however, only approximations based on limited data and simple linear calibrations. A cautious conclusion would be that PSG is already roughly calibrated to HHG.

We report results using PSG divided by 100 so that it is a proportion rather than a percentage, for reporting convenience and in order to ease interpretation with a logged dependent variable. Thus an estimated coefficient of $\hat{b}_1$=0.5 on PSG/100 and treating PSG as already calibrated to HHG implies that an increase in 10 percentage points in PSG (a bit less than one standard deviation, and an increase of about 15 % at the mean) will increase the log homicide rate by 0.10*0.5 = 0.05, i.e., the homicide rate would go up by about 5 %.

Our homicide equation includes numerous county-level control variables. We paid particular attention to those that prior theory and research suggest are important determinants of *both* gun ownership levels and homicide rates. 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 18–24 and 25–34, 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; US 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 used in percentage rather than log form. Because the raw percentages sum to 100, using them instead of logs 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 + .

Table 5 lists and provides a brief description of each variable used along with their means and standard deviations. Further details on the data and sources are discussed in Appendix 1.

Econometric Framework

We estimate Eq. (3) using county-level data with state fixed effects. Fixed effects are used 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. 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.

While it is common in the macro-level crime literature to weight observations by resident population to correct for possible heteroskedasticity, this will be the efficient feasible GLS (generalized least squares) procedure only if the heteroskedasticity takes a particular form, i.e., variance proportional to the square of the population. In the present study, the unweighted results are qualitatively consistent with the weighted results, although they differ slightly quantitatively. As described later, however, the weighted results are sometimes sensitive to the inclusion/exclusion of a relatively small number of large counties. For this reason, we present the unweighted results as the main results and the weighted results as part of our numerous robustness checks.

Our main results use two-step feasible efficient GMM estimator (see, e.g., Hayashi 2000). Since it is reasonable to suspect that observations on two counties in the same state are more likely to have correlated disturbance terms than two counties in different states, we use a "cluster-robust" GMM estimator that is both robust to, and efficient in the presence of, arbitrary heteroskedasticity and within-state correlation of the error. This robustness carries over to specification testing, including GMM tests of exogeneity.

Estimation Results

Results Treating Gun Prevalence as Exogenous

We focus on the impact of gun prevalence on *gun* homicide rates because if higher gun levels really do cause higher total homicide rates, it must surely operate through an effect on rates of homicides committed with guns. Estimation results for the benchmark regressions using the logged gun homicide rate as the dependent variable are in Panel A of Table 6. Column 1 reports the results of unweighted 2-step GMM estimations that are efficient in the presence of arbitrary heteroskedasticity and clustering, treating PSG/100 as exogenous. The estimator used is also known as "heteroskedastic OLS" (HOLS); it is the GMM estimator that uses the additional orthogonality conditions provided by the excluded instruments to improve efficiency. Coefficient estimates using OLS with robust standard errors were essentially identical, and we report HOLS results because we make use of the corresponding GMM test statistics.

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, the various poverty, low-income and inequality measures, and the percentage of the population that is black. The overall fit of the regressions is quite good, with the unweighted PSG-exogenous specification explaining 47 % of the within-state variation in county-level log gun homicide rates.[19]

[19] *Contra* Cook and Ludwig (2003b, p. 12): "the usual approach [to addressing heterogeneity in cross-sectional gun/crime studies] … has been to statistically control for the handful of local characteristics that are readily available in standard data sources, such as population density, poverty, and the age and racial composition of the population. But these variables never [sic] explain very much of the cross-sectional variation in crime rates, suggesting that the list of control variables is inadequate to the task." Our results suggest this view is too pessimistic about the feasibility of cross-sectional studies.

Table 5 Descriptive statistics

		Mean	SD	Min	Max
Homicide variables and gun prevalence, 1987–93 average					
CRMUR	Total homicides per 100,000 population	6.48	5.57	0.00	55.89
CRGMUR	Gun homicides per 100,000 population	4.11	4.16	0.00	46.08
Log CRGMUR	CRGMUR logged (N = 1,436)	1.01	0.96	−2.02	3.83
CRNGMUR	Nongun homicides per 100,000 pop.	2.37	1.82	0.00	11.83
Log CRNGMUR	CRNGMUR logged (N = 1,417)	0.62	0.77	−1.66	2.47
SRATE	Total suicides per 100,000 population	12.95	3.62	3.34	31.82
PSG	% suicides with guns	66.67	13.44	15.28	100.00
Log PSG	PSG logged	4.18	0.23	2.73	4.61
Excluded instruments					
PCTREP88	% pres. vote Republican, 1988 (N = 1,455)	56.55	9.90	14.83	81.40
Log PCTREP88	PCTREP88 logged (N = 1,455)	4.02	0.19	2.70	4.40
OMAG	Subscriptions to 3 top outdoor/sport magazines per 100,000 pop. (N = 1,450)	2,259	928	220	6,296
Log OMAG	OMAG logged (N = 1,450)	7.63	0.45	5.39	8.75
VETS	Veterans per 100,000 population	11,448	2,072	2,745	20,429
Log VETS	VETS logged	9.33	0.19	7.92	9.92
G&A	Subscriptions to *Guns & Ammo* per 100,000 pop. (N = 1,450)	256.1	90.0	48.0	1,313
Log G&A	G&A logged	5.49	0.32	3.87	7.18
Controls					
DENSITY	Persons per square mile	413	2,064	2	53,126
Log DENSITY	DENSITY logged	4.78	1.26	0.67	10.88
PCTRURAL	% rural (farm + nonfarm)	46.23	26.10	0.00	100.00
PCTSUBURBAN	% suburban (outside urbanized area)	25.48	22.15	0.00	100.00
PCTURBAN	% urban (inside urbanized area)	28.30	36.91	0.00	100.00
PCT0T17	% aged 17 and under	26.33	3.22	15.10	41.70
PCT18T24	% aged 18–24	10.70	3.73	5.10	37.10
PCT25T44	% aged 25–44	30.91	2.98	20.30	45.30
PCT45T64	% aged 45–64	18.96	2.12	8.40	27.10
PCT65PLUS	% aged 65 and over	13.11	3.56	3.00	33.80
PCTBLK	% African–American	9.22	12.60	0.01	72.13
Log PCTBLK	PCTBLK logged	1.05	1.82	−4.36	4.28
PCTHISP	% Hispanic	4.43	10.27	0.14	97.22
Log PCTHISP	PCTHISP logged	0.40	1.30	−1.97	4.58
PCTFEM18	% female-headed HHs w/children < 18	10.22	3.16	3.38	28.61
Log PCTFEM18	PCTFEM18 logged	2.28	0.28	1.22	3.35
PCTEDUC	% aged 25 + with a BA degree or higher	16.03	7.34	4.60	52.30
Log PCTEDUC	PCTEDUC logged	2.68	0.42	1.53	3.96
PCTTRANS	% born out of state	31.07	15.69	5.09	86.54
Log PCTTRANS	PCTTRANS logged	3.31	0.50	1.63	4.46
Log MEDHHINC	Log median household income, 1989	10.17	0.24	9.23	10.99
PCTINCLT15 K	% households with income < $15,000	27.91	8.86	5.00	65.20
Log PCTINCLT15 K	PCTINCLT15 K logged	3.27	0.36	1.61	4.18

Table 5 continued

		Mean	SD	Min	Max
INEQUALITY	% HHs w/income < $15 k/% income > $75 k	0.32	0.49	0.02	6.74
Log INEQUALITY	INEQUALTY logged	−1.65	0.88	−4.08	1.91
PCTPOOR	% persons below poverty line, 1989	14.28	6.90	2.20	60.00
Log PCTPOOR	PCTPOOR logged	2.55	0.48	0.79	4.09
PCTUNEMP	% persons unemployed	6.64	2.46	1.50	23.60
Log PCTUNEMP	PCTUNEMP logged	1.83	0.36	0.41	3.16
PCTVACANT	% housing units vacant	11.00	7.63	2.70	66.20
Log PCTVACANT	PCTVACANT logged	2.23	0.54	0.99	4.19

N = 1,456 except where noted

The main result of interest is the coefficient on PSG/100. Column 1 of Table 6 shows that when PSG/100 is treated as exogenous, the estimated coefficient is 0.696 and is statistically significantly different from zero at the 1 % level. The calibration exercise above implies this would be a significant effect in practical as well as statistical terms: an increase of 10 percentage points (0.10) in gun prevalence would imply an increase in gun homicide of about 6.9 %. This confirms the oft-reported result that, when endogeneity issues are ignored, gun levels are associated with higher gun crime rates. As we shall see, however, the picture changes considerably when we treat gun prevalence as endogenous.

Results Treating Gun Prevalence as Endogenous

We start with results using the three instruments PCTREP88, OMAG and VETS. We begin with the first requirement for an instrument, relevance. To test for the presence of weak instruments in our two-step GMM estimations, we use a heteroskedastic- and cluster-robust F statistic and the critical values compiled by Stock and Yogo (2005) for the IV estimator.[20] The Stock-Yogo test we report is the one for maximal size distortion. This test is based on the performance of a Wald test for the significance of $\hat{b}_1$ at the 5 % level. If instruments are weak, a Wald test rejects too often. The critical values correspond to the rejection rate r (10, 15, 20 %) that the researcher is willing to accept when the true rejection rate is 5 %. The null hypothesis of the Stock-Yogo test is that the instruments are weak and r is unacceptably high; rejection of the null hypothesis means instruments are not weak in the sense that r is acceptable to the researcher.

The first-stage regressions corresponding to the log gun homicide equation are reported in Table 7. These are fixed effects OLS regressions with PSG as the dependent variable and all controls plus the specified excluded instruments as regressors; for brevity, only coefficients on the instruments are reported. Used together, two of the three instruments, log PCTREP and log OMAG, have coefficients that are, as expected, positive and statistically highly significant, whereas the third instrument, log VETS, has a positive coefficient but a large standard error and is insignificant at conventional levels. Used one at a time, however, all three instruments are positive and significant in the first-stage regressions.

[20] Stock and Yogo do not tabulate critical values for the non-homoskedastic case. We use the IV critical values on the grounds that the IV estimator is a special case of two-step GMM for homoskedastic and independent errors. Using the simple alternative of the Staiger and Stock (1997) rule of thumb that the F statistic should be at least 10 to avoid weak instrument problems leads to similar conclusions as those we report below.

Table 6 Log homicide estimations

Dependent variables: Log homicides per 100,000 population)
Estimation method: 2-step Efficient GMM

	A. Gun homicide		B. Nongun homicide	
	(1) PSG-exogenous	(2) PSG-endogenous	(3) PSG-exogenous	(4) PSG-endogenous
PSG/100	0.696***	−2.407**	−0.160	−0.312
	(0.172)	(1.058)	(0.186)	(0.697)
Log DENSITY	0.184***	0.102**	0.116***	0.112***
	(0.031)	(0.045)	(0.020)	(0.029)
PCTSUBURBAN	−0.005***	−0.008***	0.002*	0.002
	(0.001)	(0.002)	(0.001)	(0.001)
PCTURBAN	−0.004***	−0.005***	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
PCT0T17	0.005	0.015	0.004	0.004
	(0.011)	(0.015)	(0.008)	(0.008)
PCT18T24	−0.001	0.010	−0.028***	−0.029***
	(0.008)	(0.010)	(0.007)	(0.007)
PCT25T44	0.017**	0.018**	−0.004	−0.004
	(0.007)	(0.007)	(0.007)	(0.007)
PCT45T64	0.047***	0.067***	−0.016	−0.018
	(0.014)	(0.016)	(0.013)	(0.014)
Log PCTBLK	0.089***	0.093***	0.115***	0.115***
	(0.016)	(0.018)	(0.013)	(0.014)
Log PCTHISP	0.042	0.018	0.051*	0.049*
	(0.029)	(0.030)	(0.028)	(0.029)
Log PCTFEM18	1.050***	1.179***	0.599***	0.610***
	(0.130)	(0.167)	(0.125)	(0.133)
Log PCTEDUC	−0.401***	−0.537***	−0.214***	−0.207***
	(0.103)	(0.110)	(0.077)	(0.078)
Log PCTTRANS	0.072	0.127*	0.124**	0.123**
	(0.057)	(0.065)	(0.050)	(0.050)
Log MEDHHINC	−0.194	0.247	0.762***	0.777***
	(0.402)	(0.438)	(0.245)	(0.256)
Log PCTINCLT15 K	0.280	0.840**	1.148***	1.168***
	(0.314)	(0.396)	(0.247)	(0.266)
Log INEQUALITY	0.357***	0.381***	0.166*	0.165*
	(0.104)	(0.107)	(0.090)	(0.090)
Log PCTPOOR	0.455**	0.254	0.070	0.062
	(0.203)	(0.213)	(0.124)	(0.128)
Log PCTUNEMP	−0.063	−0.060	0.059	0.070
	(0.094)	(0.096)	(0.106)	(0.107)
Log PCTVACANT	0.188***	0.176***	0.067	0.064
	(0.037)	(0.043)	(0.041)	(0.042)

Table 6 continued

Dependent variables: Log homicides per 100,000 population) Estimation method: 2-step Efficient GMM				
	A. Gun homicide		B. Nongun homicide	
	(1) PSG-exogenous	(2) PSG-endogenous	(3) PSG-exogenous	(4) PSG-endogenous
J statistic p value	$\chi^2(3) = 7.29$ 0.063	$\chi^2(2) = 1.04$ 0.595	$\chi^2(3) = 2.33$ 0.507	$\chi^2(2) = 2.27$ 0.322
Test of exogeneity of PSG p value		$\chi^2(1) = 6.25$ 0.013		$\chi^2(1) = 0.05$ 0.817
Within-R^2	0.467	n.a.	0.447	n.a.
N	1,429	1,429	1,410	1,410
Number of clusters/fixed effects	49	49	49	49

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Excluded instruments are log REP88, log OMAG and log VETS. All test statistics are robust to heteroskedasticity and clustering on state

Comparing the F statistics to the Stock-Yogo IV critical values suggests that we do not have weak instrument problems.

The estimation results for gun homicide when PSG is treated as endogenous are shown in column 2 of Table 6. We saw earlier that when PSG is treated as exogenous, the estimated impact on gun homicide, $\hat{b}_1$, was positive and statistically highly significant. When PSG is treated as endogenous and instrumented with log OMAG, log PCTREP88 and log VETS, however, the picture changes dramatically. Column 2 shows that PSG/100 has a *negative* coefficient of −2.41 that is significant at the 5 % level.[21]

We now apply the procedures outlined above for testing whether our gun proxy is endogenous and whether our instruments are exogenous. The J statistic in Table 6, column 1, when PSG is treated as exogenous, is 7.3 with a p-level of 6 %. We therefore reject the null hypothesis that the orthogonality conditions in the PSG-exogenous estimation are satisfied, and take this as evidence that one or more variables—log PSG, log OMAG, log PCTREP88, and/or log VETS—are endogenous. When PSG is treated as endogenous in column 2, however, the J statistic drops to 1.04, with a corresponding p value of 0.60. We therefore cannot reject the null that OMAG, PCTREP88 and VETS are all exogenous. Lastly, we test explicitly whether PSG is endogenous using a GMM distance test based on the difference between the J statistics for the PSG-exogenous and PSG-endogenous estimations. The χ^2 test statistic reported in column 2 is 6.25, significant at the 5 % level.[22] In short, we have evidence that PSG is endogenous, and that OMAG, PCTREP88 and VETS are all exogenous and are identifying the same LATE parameter b_1.

We checked the robustness of these results by varying the specification in a number of ways. First, we estimated as above but weighted by county population in 1990. Because the results using population weights were sometimes sensitive to the inclusion/exclusion of a small number of large counties, we estimated using both the full sample of counties with populations in excess of 25,000 as in the main results, and using a subsample that excludes

[21] The results are similar when using the traditional but inefficient IV estimator.

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

Table 7 Tests of instrument relevance (first-stage regression results)

Dependent variable: PSG/100 Estimation method: OLS						
	(1)	(2)	(3)	(4)	(5)	(6)
Log PCTREP88	0.083***	0.099***			0.083***	
	(0.018)	(0.018)			(0.018)	
Log OMAG	0.044***		0.056***		0.039***	
	(0.010)		(0.011)		(0.014)	
Log VETS	0.013			0.079***	0.013	
	(0.026)			(0.026)	(0.026)	
Log G&A					0.007	0.040***
					(0.013)	(0.011)
F statistic	16.9	30.2	25.6	9.2	12.6	12.4
Number of observations	1,429	1,435	1,430	1,436	1,429	1,430
Number of clusters/fixed effects	49	49	49	49	49	49
Stock and Yogo (2005) critical values for F						
10 % maximal IV size	22.3	16.4	16.4	16.4	24.6	16.4
15 % maximal IV size	12.8	9.0	9.0	9.0	14.0	9.0
20 % maximal IV size	9.5	6.7	6.7	6.7	10.3	6.7

*$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. All regressions are fixed effects OLS estimations with the excluded instruments as specified plus the full set of controls as specified in Table 6; for brevity only the coefficients on the instruments are reported. The sample used corresponds to that for the log gun homicide estimations; results for the sample used for log nongun homicide are very similar. The F statistic is an F test of the joint significance of the instruments. See the main text for the interpretation of the Stock-Yogo test. All test statistics are robust to heteroskedasticity and clustering on state

the roughly 100 counties with populations greater than 500,000 persons. Second, we varied the functional form of the estimating equation by using homicide rates in logs (as in the main results) and in untransformed levels, and using PSG in levels (as in the main results) and in logs. Third, we re-estimated including a lagged measure of the gun homicide rate in the equations to mitigate the possibility that other forms of unobserved heterogeneity (i.e., historical factors besides heterogeneity in the criminal population that may be responsible for current between-county differences in the gun homicide rate) may affect our results. Fourth, we addressed the possibility of a specific form of omitted variable bias, namely failing to include controls for formal deterrence measures by re-estimating our model including as controls county-level measures of police manpower levels and of the rate of solving crimes. Fifth, to see whether the basic results are a stable, long-run feature of the county-level guns-homicide relationship, we assembled comparable data for earlier years and re-estimated the basic model for 1970 and 1980.[23] We checked these various specifications individually and in various combinations. Our basic results stand up to this thorough robustness checking: when treated as exogenous, gun prevalence typically has a positive and significant impact on gun homicide; when treated as endogenous, this result is reversed and the impact of gun prevalence on gun homicide is negative or null. Full details of these robustness checks can be found in Appendix 3.

[23] The latter three robustness checks derive from comments by two anonymous referees.

Table 8 reports the results using all three instruments (column 1) and each of the instruments individually (columns 2–4).[24] Our preferred estimator in column 1 uses all instruments simultaneously since this obtains efficiency gains over using them separately, and indeed the standard error for $\hat{b}_1$ is smallest when all three instruments are used. The small J statistic in column 1 obtained when using all three instruments suggests that we should get similar estimates of $\hat{b}_1$ when using the instruments separately, and this is in fact what we find. All three separate estimates of $\hat{b}_1$ in columns 2–4 are negative, though only one of the three (OMAG) is significant at conventional levels, and, as noted, the precision of the estimates is higher than when the instruments are used together. These findings are consistent with the interpretation of the low J statistic that each of the instruments is identifying the same parameter b_1. Recall also that we had reasons to believe that each of the instruments, if invalid, would be positively correlated with the error term and hence generate estimates of the impact of guns on homicide that would be biased upwards. Since the instruments appear to be identifying the same parameter, any bias would be a shared upwards bias. This implies that the coefficient estimates using all three instruments provides an upper bound for b_1.

We now consider estimations using our fourth instrument, log G&A. Columns 5 and 6 of Table 7 show that, as expected, G&A is correlated with our gun level proxy PSG, though not as strongly as the other three instruments. Table 8, column 5 shows that when G&A is added to the instrument set, the estimated $\hat{b}_1$, while still negative, is smaller in absolute terms compared to column 1, and the J statistic jumps dramatically to 8.3 so that the null of valid instruments should now be rejected. The explanation is that G&A is identifying a b_1 that is significantly different from the one identified by the other three instruments. Column 6 shows that when G&A is used as the single instrument, $\hat{b}_1$=1.43, i.e., the estimated impact of gun prevalence is *positive*, albeit one that is insignificantly different from zero.[25] The GMM distance test of the exogeneity of G&A reported in column 5 confirms this: the null hypothesis that G&A is exogenous (is identifying the same b_1 as the other three instruments) is strongly rejected.

This estimate of $\hat{b}_1$ obtained using G&A as an instrument is not helpful in providing an upper bound for the noncriminal impact of gun prevalence β_1^{NC}; our prior is that β_1^{NC} is either close to zero or negative, and so as an upper bound a positive $\hat{b}_1$ is uninformative, and in any case the coefficient has a large standard error (1.82) attached to it. It is also difficult to interpret it in terms of an estimate of β_1^{C}, i.e., as a LATE estimate that is local to criminals, for two reasons: first, as noted above, the correlation of G&A with our proxy for gun prevalence is probably dominated by the more numerous noncriminal subscribers; and second, as suggested earlier, a more plausible explanation for a positive $\hat{b}_1$ in this case is a large positive endogeneity bias generated by reverse causality. The results are still informative, however. In addition to demonstrating that our exogeneity tests can have the power to detect violations of the null, they suggest that use of *Guns and Ammo* magazine as either an instrument or a proxy for gun prevalence may be vulnerable to reverse causality problems. The results are also relevant for the interpretation of our results using OMAG and PCTREP98; the finding that G&A appears to be affected by reverse causality

[24] The IV estimator is used when there is only one excluded instrument because the 2-step GMM estimator reduces to IV in the just-identified case.

[25] An exogeneity test of PSG using G&A as the sole instrument yields a $\chi^2(1)$ statistic of 0.28 with a *p* value of 0.600. As noted earlier, this is equivalent to a test of the difference between an estimate of b_1 treating PSG as exogenous and an estimate using G&A as an instrument for PSG.

Table 8 Gun homicide equation, various specifications of instruments

Dependent variable: Log CRGMUR (gun homicide per 100,000 population) PSG treated as endogenous						
	(1)	(2)	(3)	(4)	(5)	(6)
Instruments	Log PCTREP88 Log OMAG***Log VETS	Log PCTREP88	Log OMAG	Log VETS	Log PCTREP88 Log OMAG***Log VETS***Log G&A	Log G&A
Estimation method	2-step GMM	IV	IV	IV	2-step GMM	IV
Coefficient on PSG/100	−2.41**	−0.15	−3.08**	−3.21	−1.56*	1.43
	(1.06)	(1.29)	(1.26)	(2.09)	(1.01)	(1.82)
95 % confidence interval	[− 4.48, −0.33]	[− 4.04, 1.00]	[− 5.55, −0.61]	[− 7.31, 0.89]	[− 3.53, 0.41]	[–2.12, 4.99]
J statistic	$\chi^2(2) = 1.04$*** p value = 0.595	Just-identified	Just-identified	Just-identified	$\chi^2(3) = 8.29$*** p value = 0.041	Just-identified
Test of exogeneity of G&A					$\chi^2(1) = 7.23$*** p value = 0.007	
Number of observations	1,429	1,435	1,430	1,436	1,429	1,430
Number of clusters/fixed effects	49	49	49	49	49	49

† $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. All regressions are fixed effects estimations with the excluded instruments as specified plus the full set of controls as specified in Table 6; for brevity only the coefficient on PSG is reported. All test statistics are robust to heteroskedasticity and clustering on states

indirectly supports our suggestion that this is the most plausible source of potential bias in the LATE parameters identified by OMAG and PCTREP88.

To summarize, using our three main instruments for gun prevalence—subscriptions to outdoor sports magazines, voting patterns, and numbers of veterans—we obtain an estimate of a LATE parameter $\hat{b}_1$ that suggests a negative impact of gun prevalence on gun homicide. The low J statistic suggests that all three instruments are identifying the same weighted average of β_1^C and β_1^{NC}. When gun prevalence is treated as exogenous, or instrumented with subscriptions to *Guns and Ammo* magazine, the estimated impact is positive and significantly different from the estimate obtained by instrumenting gun prevalence with our three main instruments, which we interpret as evidence of reverse causality affecting these two variables: high gun homicide rates lead people to acquire guns, and stimulate interest in a magazine that is directly gun-related.

We discussed in the previous section reasons why the LATE parameters identified by OMAG, PCTREP88 or VETS could either put a heavier weight on the impact of criminal gun prevalence (VETS) or possibly be biased upwards via a positive correlation with the error term (OMAG and PCTREP88) generated by reverse causality. In the case of VETS, this prior was supported by empirical evidence about involvement of veterans in violence after reentering civilian life. Our specification tests suggest that these three instruments are estimating approximately the same negative LATE impact of gun prevalence on gun homicide. This implies either that the instruments are all identifying the same underlying LATE parameter, and/or they are generating estimates that are biased upwards and to a similar degree via a correlation with the error term, or some combination of these two channels.

The two most plausible interpretations of our results are therefore as follows: (1) All three instruments are identifying primarily the LATE parameter local to noncriminals; the weight given to criminal effects in the estimated parameter $\hat{b}_1$ is small, and $\hat{b}_1$ can be interpreted as β_1^{NC} scaled by $1/\delta_1^{NC}$. (2) Either because of upward bias, and/or weight put on the impact of criminal gun prevalence β_1^C, the estimated parameter $\hat{b}_1$ provides an upper bound to the noncriminal effect β_1^{NC} scaled by $1/\delta_1^{NC}$. We conclude that our estimates of $\hat{b}_1$ provide evidence of a negative deterrent effect of noncriminal gun prevalence that is statistically significant. In the next section we consider whether this LATE estimate is significant in practical terms, and what we can say about the ATE.

Calibration of the Estimated Impact of Gun Prevalence on Gun Homicide

The calibration exercise earlier in the paper suggests that PSG is already approximately calibrated to HHG, i.e., δ_1 is approximately 1. Our estimate for $\hat{b}_1$ therefore suggests a negative effect of noncriminal gun prevalence on gun homicide that is practically as well as statistically significant. If we use the 1:1 calibration of PSG to HHG, the point estimate of the upper bound for β_1^{NC} is -2.41, implying that an increase of 10 percentage points in noncriminal gun prevalence would reduce gun homicide by at least $\sim$20–25 %. Estimates taken from the conservative end of the 95 % confidence intervals for $\hat{b}_1$ combined with a conservative calibration of PSG to HHG would still generate important negative impacts; e.g., $b_1 = -.33$ and $\delta_1 = 0.7$ implies an upper bound for β_1^{NC} of about -0.23, so that a 10 percentage point increase in noncriminal gun prevalence would reduce gun homicide by $\sim$2 % or more.

The finding that noncriminal gun prevalence has a substantial negative impact on gun homicide is the result of this paper that is most relevant for policy purposes; we will return to this point in the conclusion. However, β_1^{ATE}—the impact of an increase in overall gun prevalence randomly distributed across localities and across criminal and noncriminals—is of some interest, if only because this is the parameter on which previous studies have implicitly focused. We therefore consider here what our estimates suggest as a range of plausible values for the ATE.

The ATE in our model is a weighted average of β_1^{NC} and β_1^{C}, where the weights are given by the relative prevalence of criminal and noncriminal guns, plus the Frishman adjustment for the expectation of a ratio (Eq. 8). For the purposes of this rough calibration we omit this adjustment and treat the overall composition of gun prevalence in the US as an acceptable approximation for the average across counties of the composition of gun prevalence. Evidence is scarce, but plausible figures for the criminal gun stock in the US would be in the range of 10–25 % of the total. Taking total gun prevalence as approximately 50 % of households gives us a range for $\bar{g}^C$ of about 5–13 %. We are not aware of any previous estimates of β_1^{C} or, equivalently, of the elasticity of gun homicide with respect to criminal gun prevalence, which in our model evaluated at the mean would be $\beta_1^{C}\bar{g}^C$.[26]

Simple exercises suggest that an ATE that is positive and important in practical terms is as plausible as one that is negative and important. For example, if we assume $\beta_1^{C} = 7$ and $\bar{g}^C$=13 %, the implied gun homicide-criminal gun prevalence elasticity is a plausible 7*0.13 = 0.9. Combine this with a conservative estimate of β_1^{NC}= −1, and β_1^{ATE}=(0.75* − 1) + (0.25*7) = 1, i.e., a 10 percentage point increase in general gun prevalence would raise gun homicides by about 10 %. Alternatively, if we assume $\beta_1^{C} = 7$ and $\bar{g}^C$=5 %, we get a gun homicide-criminal gun prevalence elasticity of about 0.4, which is also plausible. Combine this with our point estimate of β_1^{NC}= −2.4, and β_1^{ATE}=(0.90* − 2.4) + (0.10*7) = −1.5, which is just the opposite result—an 10 percentage point increase in general gun prevalence reduces gun homicides by about 15 %.

In short, the possibilities that the ATE is positive, negative, or approximately zero, all appear plausible and consistent with our estimates of a significant negative impact of noncriminal gun prevalence.

The Impact of Gun Prevalence on Nongun and Total Homicide

The analysis of nongun homicide is somewhat simpler, because of the three channels through which gun prevalence can affect homicide rates in general—criminality, deterrence and self-defense—the first, criminality, is not relevant because it would generate gun homicides rather than nongun homicides. Nevertheless, heterogeneity in criminality is an issue here as well, and we suggest that our IV/GMM results should be also interpreted as estimates that are primarily "local to noncriminals". We again use PSG as our proxy for general gun prevalence.

The first-stage results are essentially identical to those reported in Table 7 and we do not report them here.[27] Panel B of Table 6 presents the results for the GMM estimations when the dependent variable is the log nongun homicide rate. The patterns in the

[26] Obtained by differentiating Eq. (5) with respect to g^C, noting that h is a log crime rate, and then multiplying both sides by $\bar{g}^C$ to obtain the elasticity at the mean.

[27] The differences arise because of the non-overlap of the small number of missing values for the two dependent variables.

coefficients on the covariates are similar to those for the gun homicide equations. The impact of gun prevalence, however, is rather different compared to the gun homicide case. When gun levels are treated as exogenous, the estimated coefficient on PSG is negative but insignificantly different from zero. When gun levels are treated as endogenous, the estimated coefficient becomes more negative but remains insignificantly different from zero. The test of the endogeneity of PSG suggests it can be treated as exogenous, which implies a null impact on nongun homicide. We report the same robustness checks for nongun homicide as for gun homicide in Appendix 3. The results reinforce the conclusion that a negative deterrent effect on nongun homicide is either weak or nonexistent: in these estimations, PSG is usually insignificantly different from zero, whether treated as exogenous or endogenous. As noted earlier, however, there is a great deal of prior evidence and literature to suggest that the coefficient on gun levels in a crime equation should be biased upwards if the endogeneity problem is not addressed. This would again suggest regarding the estimated coefficient treating PSG as exogenous (column 3) as an upper bound, and modifying our conclusion to the statement that gun levels have no positive impact on nongun homicide.

The separate estimates of gun prevalence on gun and nongun homicide can be combined to obtain an estimate of the impact on total homicide. Because the dependent variables in the two equations are in logs, the impact on total homicide is a weighted average of the two coefficients on gun prevalence, where the weights are the shares of gun and nongun homicide in total homicide (0.634 and 0.366, respectively[28]). Our LATE interpretation applies here as well, i.e., our preferred interpretation of the impact of gun prevalence on total homicide is that it is local to noncriminals. To obtain a standard error for the estimated impact, we estimate the gun and nongun homicide equations as a system by "stacking" the two equations.[29] The use of a cluster-robust covariance estimator that clusters on state enables us to test cross-equation hypotheses—in this case, the significance of a weighted average of the estimated coefficients on PSG in the gun and nongun homicide equations.[30] The results are reported in Table 9; Panel A shows the results when PSG is treated as endogenous in both equations, and Panel B shows the more efficient results when it is treated as exogenous in the nongun homicide equation. Because we use efficient GMM applied to a system of 2 equations, the estimates of the coefficients on PSG differ slightly from those of the equation-by-equation estimates in Table 6. The results in Table 9 are as expected: gun prevalence has a negative impact on total homicide, driven by the underlying negative impact on gun homicide, and with a lower elasticity (coefficient) because of the dilution via the null impact on nongun homicide. Using the 1:1 calibration of PSG to HHG and the point estimate of -1.7 from Panel B, an increase of 10 percentage points in noncriminal gun prevalence would reduce total homicide by $\sim$17 %. A conservative calibration of PSG to HHG ($\delta_1 = 0.7$) and the conservative end of the 95 % confidence interval ($\hat{b}_1 = -0.6$) implies total homicide would instead fall by $\sim$4 %.

[28] These are the county means of the shares of gun and nongun homicide in total homicide for the total sample of 1,456 counties.

[29] That is, we double the size of the dataset so that all counties appear twice, once where the dependent variable is the log gun homicide rate and again where it is the log nongun homicide rate. We then interact all regressors and instruments, including the fixed effects, with a gun/nongun homicide dummy, so that all of the variables in the gun homicide equation take the value of zero when the dependent variable is nongun homicide, and visa-versa for the nongun homicide equation variables.

[30] This works because clustering on states allows for arbitrary within-state correlations. This includes the possible correlation between the two observations on an individual county (one from the gun equation, one from the nongun equation).

Table 9 Gun homicide, nongun homicide and total homicide

Joint estimation of gun and nongun homicide equations Dependent variables are logs of homicides per 100,000 population Estimation method is 2-step GMM

	Panel A			Panel B		
	(1) Log gun homicide	(2) Log nongun homicide	(3) Log total homicide	(4) Log gun homicide	(5) Log nongun homicide	(6) Log total homicide
PSG treated as:	Endogenous	Endogenous	n.a.	Endogenous	Exogenous	n.a.
Weight in total homicide	0.634	0.366	n.a.	0.634	0.366	n.a.
Coefficient on PSG/100	–2.61**	–0.424	–1.81**	–2.60***	–0.21	–1.72***
	(1.03)	(0.68)	(0.80)	(0.90)	(0.18)	(0.57)
95 % confidence interval	[–4.63, –0.59]	[–1.76, 0.92]	[–3.38, –0.25]	[–4.36, –0.84]	[–0.56, 0.15]	[–2.85, –0.60]
J statistic	$\chi^2(4) = 4.23$*** p value = 0.376			$\chi^2(5) = 4.11$*** p value = 0.534		
Number of observations	2,839			2,839		
Number of clusters/fixed effects	49			49		

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Estimation results are obtained by stacking the gun and nongun homicide equations; see text for description of estimation method. Estimation includes separate fixed effects, excluded instruments (Log PCTREP88, Log OMAG and Log VETS) and full sets of controls for gun and nongun homicide. Log total homicide is the weighted average of the estimated gun and nongun homicide coefficients; weights are based on mean county homicide rates for the total sample of 1,456 counties (Table 5). For brevity only the coefficients on PSG are reported. All test statistics are robust to heteroskedasticity and clustering on states

If we take these estimates of gun effects seriously, they suggest that gun levels in the general public may have a net 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, despite the tactical advantages of possessing a deadly weapon, 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.

Conclusions

Prior studies that have attempted to estimate the average treatment effect of gun prevalence on homicide have failed to properly address endogeneity bias or demonstrate that the chosen proxy for gun prevalence is correlated with time-series variation in gun levels. We show that recent studies attempting to address the endogeneity bias problem by using panel data have exacerbated the proxy validity problem by relying on proxies that are essentially uncorrelated in time series with direct measures of gun prevalence. We also demonstrate that even if researchers could eliminate these problems, the estimation frameworks they have used would (1) still not produce a consistent estimate of the average treatment effect of gun prevalence, and (2) be of little value to policymakers as gun law restrictions are usually intended to have differential effects on criminal and noncriminal gun prevalence. That is, an ATE approach cannot address the "heterogeneity in criminality" problem.

This paper used county-level cross-sectional data for 1990 data where the proxy (PSG) has been established by numerous researchers as having strong correlations across space (e.g., states, counties, nations) with direct survey measures of gun ownership. Our strategy is to use instrumental variables in a LATE framework to address the other problems of endogeneity bias and heterogeneity in criminality. The benefit of the LATE approach is that it enables the separation of the effects of criminal and noncriminal gun prevalence. In the context of our LATE framework, the estimated impact of gun prevalence is a weighted average of a possibly negative impact of noncriminal gun prevalence on homicide and a presumed positive impact of criminal gun prevalence. We find evidence of a significant negative impact, and interpret it as primarily "local to noncriminals", i.e., primarily determined by the homicide-reducing effects of noncriminal gun prevalence. We also demonstrate that an ATE for gun prevalence that is positive, negative, or approximately zero are all entirely plausible and consistent with our estimates of a significant negative impact of noncriminal gun prevalence.

The policy implications of our findings are perhaps best understood in the context of two hypothetical gun ban scenarios, the first more optimistic, the second more pessimistic and realistic. First, gun prohibition might reduce gun ownership equiproportionately among criminals and noncriminals, and the traditional ATE interpretation therefore applies. Our results above suggest that plausible estimates of the causal impact of an average reduction in gun prevalence include positive, nil, and negative effects on gun homicide rates, and hence no strong evidence in favor of or against such a measure. But it is highly unlikely that criminals would comply with gun prohibition to the same extent as noncriminals; indeed, it is virtually a tautology that criminals would violate a gun ban at a

higher rate than noncriminals. Thus, under the more likely scenario that gun bans reduced gun levels more among noncriminals than criminals, the LATE interpretation of our results moves the range of possible impacts towards an increase in gun homicide rates because the decline in gun levels would primarily occur among those whose gun possession has predominantly negative effects on homicide.

In sum, the instrumental variables/LATE approach taken in this paper should prove useful to both researchers and policymakers by providing some preliminary estimates on the signs and magnitudes of the separate criminal and noncriminal effects of gun prevalence. Coupled with relevant information on the likely impacts of an existing or proposed gun law restriction, researchers will be able to more accurately assess the effects of such restrictions on violence rates. Future researchers should attempt to develop separate measures of criminal and noncriminal gun prevalence, so as to allow more direct tests of these differing "local" effects.

Acknowledgments Seminar audiences at Aberdeen, Bologna, Dundee, Edinburgh, IZA (Bonn), Moscow, Strathclyde and Zurich provided useful comments and suggestions on earlier versions of this paper. We are grateful to John DiNardo, David Drukker, John Earle, Andrea Ichino, Austin Nichols, Steven Stillman, Jeff Wooldridge, the editor and two anonymous referees for helpful comments, discussions and suggestions. The usual caveat applies.

Appendix 1: Data and Sources

We use cross-sectional data for US counties which had a population of 25,000 or greater in 1990, and for which relevant data were available (N = 1,456). Alaska and Washington, DC were excluded from the analysis: the former, because we did not have compatible data for one of our instruments (voting in 1988); the latter, because it is itself a single county and thus drops out of a fixed-effects specification. Data for most county level variables were obtained from the US Bureau of the Census, *County and City Data Book,* 1994. Other data sources were as follows:

Homicide rates are averages for the 7 years 1987–1993 (bracketing the decennial census year of 1990). 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 (US NCHS 1997). The data include all intentional homicides in the county with the exception of those due to legal intervention (e.g., killings by police and executions).

Similar to homicide, data for the percent of suicides committed with guns are also 1987–93 averages and were obtained using special Part III Mortality Detail File computer tapes made available by the NCHS. Unlike widely available public use versions, the tapes permit the aggregation of death counts for even the smallest counties (US NCHS 1997).

Subscriptions per 100,000 county population to three of the most popular outdoor/sport magazines (*Field and Stream*, *Outdoor Life*, and *Sports Afield*) in 1993 were obtained from Audit Bureau of Circulations (1993). In the earlier version of this paper, we used a principal components index based on the three separate subscription rates; the measure we use here is more convenient and generates almost identical results.

The percent of the county population voting for the Republican candidate in the 1988 Presidential election is from ICPSR (1995). Rurality measures are from US Bureau of the Census (1990).

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.[31] For further discussion of how the estimators and tests are implemented, see Baum et al. (2003, 2007, 2008), Schaffer (2007), and the references therein.

Appendix 2: The OLS and IV Estimators with Population Heterogeneity

Model Setup

The "true model" is one with population heterogeneity (Eq. 7c in the main text):

$$h_i = \beta_0 + \beta_1^{\mathrm{C}} g_{\mathrm{i}}^{\mathrm{C}} + \beta_1^{\mathrm{NC}} g_{\mathrm{i}}^{\mathrm{NC}} + u_i \tag{19}$$

Criminal and noncriminal gun prevalence are not separately observable. A proxy for aggregate gun prevalence is available (Eq. 14 in the text):

$$p_i = \delta_0 + \delta_1^{\mathrm{C}} g_{\mathrm{i}}^{\mathrm{C}} + \delta_1^{\mathrm{NC}} g_{\mathrm{i}}^{\mathrm{NC}} + v_i \tag{20}$$

A single instrument Z_i is available that is correlated with both criminal and noncriminal gun prevalence, but the strength of the correlation may differ (Eqs. 11 and 12 in the text):

$$g_{\mathrm{i}}^{\mathrm{C}} = \pi_0^{\mathrm{C}} + \pi_1^{\mathrm{C}} Z_i + \eta_{\mathrm{i}}^{\mathrm{C}} \tag{21}$$

$$g_{\mathrm{i}}^{\mathrm{NC}} = \pi_0^{\mathrm{NC}} + \pi_1^{\mathrm{NC}} Z_i + \eta_{\mathrm{i}}^{\mathrm{NC}} \tag{22}$$

We assume that $\pi_1^{\mathrm{C}}, \pi_1^{\mathrm{NC}} \geq 0$ and at least one is strictly greater than zero, and similarly for δ_1^{C} and δ_1^{NC}. If gun prevalence is directly observable, the estimating equation is (Eq. 1 in the text):

$$h_i = \beta_0 + \beta_1 g_i + u_i \tag{23}$$

If only the proxy for gun prevalence is observable, the estimating equation is (Eq. 3 in the text):

$$h_i = \mathrm{b}_0 + \mathrm{b}_1 p_i + e_i \tag{24}$$

The derivations below follow the format of those in Stock and Watson (2007), Appendix 13.4.

The Average Treatment Effect (ATE) of Gun Prevalence

Rewrite Eq. (19) as a "random coefficient" model:

$$h_i = \beta_0 + \beta_{1\mathrm{i}} g_{\mathrm{i}} + u_i \tag{25}$$

where

$$\beta_{1\mathrm{i}} \equiv \frac{g_{\mathrm{i}}^{\mathrm{C}}}{g_{\mathrm{i}}^{\mathrm{C}} + g_{\mathrm{i}}^{\mathrm{NC}}} \beta_1^{\mathrm{C}} + \frac{g_{\mathrm{i}}^{\mathrm{NC}}}{g_{\mathrm{i}}^{\mathrm{C}} + g_{\mathrm{i}}^{\mathrm{NC}}} \beta_1^{\mathrm{NC}} \tag{26}$$

and by definition $g_{\mathrm{i}} \equiv g_{\mathrm{i}}^{\mathrm{C}} + g_{\mathrm{i}}^{\mathrm{NC}}$, i.e., our measures of gun prevalence are in levels. The average treatment effect of gun prevalence β^{ATE} is:

[31] 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.

$$\begin{aligned} \mathrm{E}(\beta_{1\mathrm{i}}) &= \mathrm{E}\left(\frac{g_{\mathrm{i}}^{\mathrm{C}}}{g_{\mathrm{i}}^{\mathrm{C}}+g_{\mathrm{i}}^{\mathrm{NC}}}\right)\beta_1^{\mathrm{C}} + \mathrm{E}\left(\frac{g_{\mathrm{i}}^{\mathrm{NC}}}{g_{\mathrm{i}}^{\mathrm{C}}+g_{\mathrm{i}}^{\mathrm{NC}}}\right)\beta_1^{\mathrm{NC}} \\ &= \left(\frac{\mu^{\mathrm{C}} - \mathrm{cov}\left(\mathrm{g_i}, \frac{g_{\mathrm{i}}^{\mathrm{C}}}{g_{\mathrm{i}}}\right)}{\mu^{\mathrm{C}}+\mu^{\mathrm{NC}}}\right)\beta_1^{\mathrm{C}} + \left(\frac{\mu^{\mathrm{NC}} + \mathrm{cov}\left(\mathrm{g_i}, \frac{g_{\mathrm{i}}^{\mathrm{C}}}{g_{\mathrm{i}}}\right)}{\mu^{\mathrm{C}}+\mu^{\mathrm{NC}}}\right)\beta_1^{\mathrm{NC}} \end{aligned} \tag{27}$$

where $\mu^{\mathrm{C}} \equiv \mathrm{E}(g_{\mathrm{i}}^{\mathrm{C}})$ and $\mu^{\mathrm{NC}} \equiv \mathrm{E}(g_{\mathrm{i}}^{\mathrm{NC}})$ and where we make use of the result in Frishman (1971) for the expectation of a ratio; the covariance terms account for the fact that the expectation of the ratio of two random variables does not, in general, equal the ratio of the expectations.[32] In the special case that total gun prevalence is uncorrelated with the criminal/noncriminal share of gun prevalence, the covariance terms in (27) are zero and the ATE takes the following simple form:

$$\mathrm{E}(\beta_{1\mathrm{i}}) = \frac{\mu^{\mathrm{C}}}{\mu^{\mathrm{C}}+\mu^{\mathrm{NC}}}\beta_1^{\mathrm{C}} + \frac{\mu^{\mathrm{NC}}}{\mu^{\mathrm{C}}+\mu^{\mathrm{NC}}}\beta_1^{\mathrm{NC}} \tag{28}$$

Equations (27) and (28) are Eqs. (8) and (9) in the main text.

OLS Estimation

We consider first estimation of Eq. (23), when total gun prevalence is directly observable. The OLS estimator is

$$\hat{\beta}_1^{\mathrm{OLS}} = \frac{\mathrm{s_{gh}}}{\mathrm{s_g^2}} \xrightarrow{\mathrm{P}} \frac{\mathrm{cov}(g_i, h_i)}{\mathrm{var}(g_i)} \tag{29}$$

where s denotes a sample covariance and $\xrightarrow{\mathrm{P}}$ denotes convergence in probability. The numerator is

$$\begin{aligned} \mathrm{cov}(g_i, h_i) &= \mathrm{cov}\left\{\left(g_i^C + g_i^{NC}\right), \left(\beta_0 + \beta_1^C g_i^C + \beta_1^{NC} g_i^{NC} + u_i\right)\right\} \\ &= \beta_1^C \mathrm{var}(g_i^C) + \beta_1^C \mathrm{cov}(g_i^C, g_i^{NC}) + \beta_1^{NC} \mathrm{var}(g_i^{NC}) + \beta_1^{NC} \mathrm{cov}(g_i^C, g_i^{NC}) \\ &\quad + \mathrm{cov}(g_i^C, u_i) + \mathrm{cov}(g_i^{NC}, u_i) \\ &= \beta_1^C \left\{\mathrm{var}(g_i^C) + \mathrm{cov}(g_i^C, g_i^{NC})\right\} + \beta_1^{NC} \left\{\mathrm{var}(g_i^{NC}) + \mathrm{cov}(g_i^C, g_i^{NC})\right\} \\ &\quad + \mathrm{cov}(g_i^C, u_i) + \mathrm{cov}(g_i^{NC}, u_i) \end{aligned} \tag{30}$$

The denominator is simply

$$\begin{aligned} \mathrm{var}(g_i) &= \mathrm{var}\left(g_i^C + g_i^{NC}\right) \\ &= \mathrm{var}(g_i^C) + \mathrm{var}(g_i^{NC}) + 2\mathrm{cov}(g_i^C, g_i^{NC}) \end{aligned} \tag{31}$$

and therefore

[32] If X and Y are two random variables, then $\mathrm{E}\left(\frac{Y}{X}\right) = \frac{\mathrm{E}(Y)}{\mathrm{E}(X)} - \frac{\mathrm{cov}(X,\frac{Y}{X})}{\mathrm{E}(X)}$ (Frishman 1971, p. 333, Eq. 4.2), provided that $\mathrm{E}(X) \neq 0$, i.e., we are assuming that the mean of total gun prevalence is nonzero. The derivation of Eq. (27) also makes use of the fact that $\mathrm{cov}\left(g_i, \frac{g_{\mathrm{i}}^{\mathrm{NC}}}{g_i}\right) = \mathrm{cov}\left(g_i, 1 - \frac{g_{\mathrm{i}}^{\mathrm{C}}}{g_i}\right) = -\mathrm{cov}\left(g_i, \frac{g_{\mathrm{i}}^{\mathrm{C}}}{g_i}\right)$. Note that the weights on β_1^{C} and β_1^{NC} in (27) sum to 1.

$$\hat{\beta}_1^{\mathrm{OLS}} = \frac{s_{gh}}{s_g^2} \xrightarrow{P} \frac{\{\mathrm{var}(g_i^C) + \mathrm{cov}(g_i^C, g_i^{NC})\}}{\mathrm{var}(g_i^C) + \mathrm{var}(g_i^{NC}) + 2\mathrm{cov}(g_i^C, g_i^{NC})} \beta_1^C + \frac{\{\mathrm{var}(g_i^{NC}) + \mathrm{cov}(g_i^C, g_i^{NC})\}}{\mathrm{var}(g_i^C) + \mathrm{var}(g_i^{NC}) + 2\mathrm{cov}(g_i^C, g_i^{NC})} \beta_1^{NC} + \frac{\mathrm{cov}(g_i^C, u_i) + \mathrm{cov}(g_i^{NC}, u_i)}{\mathrm{var}(g_i^C) + \mathrm{var}(g_i^{NC}) + 2\mathrm{cov}(g_i^C, g_i^{NC})} \tag{32}$$

The OLS estimator differs from the ATE Eq. (27) for two reasons. First, if criminal or noncriminal guns are endogenous, then the third term in Eq. (32) is nonzero. Second, even if gun prevalence is exogenous and the third term in Eq. (32) drops out, the resulting OLS estimator is a weighted average of β_1^{C} and β_1^{NC}, but the weights differ from those for the ATE in Eq. (27); whereas the ATE weights are relative gun prevalence (plus the Frishman correction for the expectation of a ratio), the OLS weights are driven by the variances and covariances of gun prevalence, i.e., by gun variability. The intuition is that the identifying variation in the estimation of Eq. (23) comes from the variation in criminal and noncriminal gun prevalence, and these may differ. To take an extreme example, if criminal and noncriminal gun prevalence are uncorrelated so that $\mathrm{cov}(g_i^C, g_i^{NC}) = 0$, and noncriminal gun prevalence varies little or not at all across localities so that $\mathrm{var}(g_i^{NC}) \approx 0$, then the OLS estimator $\hat{\beta}_1^{\mathrm{OLS}}$ will be approximately equal to the impact of criminal guns β_1^{C}, because the identifying variation in the data is driven solely by variation in criminal gun prevalence.

Next we consider OLS estimation of Eq. (24), when only a proxy is available. To simplify the algebra, we assume that the homicide error u_i and the proxy error v_i are uncorrelated with gun prevalence and with each other. The OLS estimator is

$$\hat{b}_1^{\mathrm{OLS}} = \frac{s_{ph}}{s_p^2} \xrightarrow{P} \frac{\mathrm{cov}(p_i, h_i)}{\mathrm{var}(p_i)} \tag{33}$$

The numerator is

$$\begin{aligned} \mathrm{cov}(p_i, h_i) &= \mathrm{cov}\{(\delta_0 + \delta_1^{\mathrm{C}} g_i^{\mathrm{C}} + \delta_1^{\mathrm{NC}} g_i^{\mathrm{NC}} + v_i), (\beta_0 + \beta_1^C g_i^C + \beta_1^{NC} g_i^{NC} + u_i)\} \\ &= \delta_1^{\mathrm{C}} \beta_1^C \mathrm{var}(g_i^{\mathrm{C}}) + \delta_1^{\mathrm{NC}} \beta_1^{NC} \mathrm{var}(g_i^{\mathrm{NC}}) + (\delta_1^{\mathrm{C}} \beta_1^{NC} + \delta_1^{\mathrm{NC}} \beta_1^C) \mathrm{cov}(g_i^{\mathrm{C}}, g_i^{\mathrm{NC}}) \end{aligned} \tag{34}$$

since we've assumed that the error terms are uncorrelated with gun levels. The denominator is

$$\begin{aligned} \mathrm{var}(p_i) &= \mathrm{var}(\delta_0 + \delta_1^{\mathrm{C}} g_i^{\mathrm{C}} + \delta_1^{\mathrm{NC}} g_i^{\mathrm{NC}} + v_i) \\ &= (\delta_1^{\mathrm{C}})^2 \mathrm{var}(g_i^C) + (\delta_1^{\mathrm{NC}})^2 \mathrm{var}(g_i^{NC}) + 2\delta_1^{\mathrm{C}} \delta_1^{\mathrm{NC}} \mathrm{cov}(g_i^C, g_i^{NC}) + \mathrm{var}(v_i) \end{aligned} \tag{35}$$

Thus

$$\hat{b}_1^{\mathrm{OLS}} \xrightarrow{P} \frac{\delta_1^{\mathrm{C}} \mathrm{var}(g_i^{\mathrm{C}}) + \delta_1^{\mathrm{NC}} \mathrm{cov}(g_i^{\mathrm{C}}, g_i^{\mathrm{NC}})}{(\delta_1^{\mathrm{C}})^2 \mathrm{var}(g_i^C) + (\delta_1^{\mathrm{NC}})^2 \mathrm{var}(g_i^{NC}) + 2\delta_1^{\mathrm{C}} \delta_1^{\mathrm{NC}} \mathrm{cov}(g_i^C, g_i^{NC}) + \mathrm{var}(v_i)} \beta_1^C + \frac{\delta_1^{\mathrm{NC}} \mathrm{var}(g_i^{\mathrm{NC}}) + \delta_1^{\mathrm{C}} \mathrm{cov}(g_i^{\mathrm{C}}, g_i^{\mathrm{NC}})}{(\delta_1^{\mathrm{C}})^2 \mathrm{var}(g_i^C) + (\delta_1^{\mathrm{NC}})^2 \mathrm{var}(g_i^{NC}) + 2\delta_1^{\mathrm{C}} \delta_1^{\mathrm{NC}} \mathrm{cov}(g_i^C, g_i^{NC}) + \mathrm{var}(v_i)} \beta_1^{NC} \tag{36}$$

Equation (36) shows the OLS estimator using a proxy for gun levels is a weighted average of the criminal and noncriminal effects. The weights sum to less than one because of the $\mathrm{var}(v_i)$ term; this is the attenuation bias attributable to the measurement error in the proxy.

The weights on β_1^{C} and β_1^{NC} now depend not only on gun variability, but also on the relative strength of the correlations between the proxy and criminal/noncriminal gun levels: if $\delta_1^{NC} \gg \delta_1^{C}$, then the OLS estimator will put a high weight on the noncriminal impact gun prevalence, and vice versa if $\delta_1^{NC} \ll \delta_1^{C}$. Note that even if $\delta_1^{C} = 0$, the weight on β_1^{C} may be positive if criminal and noncriminal gun prevalence are correlated. Note also that $\text{sign}\{\hat{b}_1^{OLS}\}$ is not in general a consistent estimator of $\text{sign}\{\hat{\beta}_1^{OLS}\}$.

IV Estimation

Again we start with the case where gun levels are observable. The IV estimator can be written

$$\hat{\beta}_1^{IV} = \frac{s_{Zh}}{s_{Zg}} \xrightarrow{P} \frac{\text{cov}(Z_i, h_i)}{\text{cov}(Z_i, g_i)} \tag{37}$$

Taking the numerator first,

$$\begin{aligned} \text{cov}(Z_i, h_i) &= \text{cov}\{Z_i, (\beta_0 + \beta_1^{C} g_i^{C} + \beta_1^{NC} g_i^{NC} + u_i)\} \\ &= \text{cov}(Z_i, \beta_1^{C} g_i^{C}) + \text{cov}(Z_i, \beta_1^{NC} g_i^{NC}) + \text{cov}(Z_i, u_i) \\ &= \beta_1^{C} \text{cov}(Z_i, g_i^{C}) + \beta_1^{NC} \text{cov}(Z_i, g_i^{NC}) \end{aligned} \tag{38}$$

since Z is exogenous and orthogonal to the error term. Using Eqs. (21) and (22), we have

$$\begin{aligned} \text{cov}(Z_i, g_i^{C}) &= \text{cov}(Z_i, \pi_0^{C} + \pi_1^{C} Z_i + \eta_i) = \pi_1^{C} \text{var}(Z_i) \\ \text{cov}(Z_i, g_i^{NC}) &= \text{cov}(Z_i, \pi_0^{NC} + \pi_1^{NC} Z_i + \eta_i) = \pi_1^{NC} \text{var}(Z_i) \end{aligned} \tag{39}$$

since Z is also uncorrelated with η. Substituting (39) into (38), we have

$$\text{cov}(Z_i, h_i) = \beta_1^{C} \pi_1^{C} \text{var}(Z_i) + \beta_1^{NC} \pi_1^{NC} \text{var}(Z_i) = \text{var}(Z_i)\{\beta_1^{C} \pi_1^{C} + \beta_1^{NC} \pi_1^{NC}\} \tag{40}$$

Now taking the denominator of (37),

$$\begin{aligned} \text{cov}(Z_i, g_i) &= \text{cov}\{Z_i, (g_i^{C} + g_i^{NC})\} = \text{cov}(Z_i, g_i^{C}) + \text{cov}(Z_i, g_i^{NC}) \\ &= \pi_1^{C} \text{var}(Z_i) + \pi_1^{NC} \text{var}(Z_i) = \text{var}(Z_i)\{\pi_1^{C} + \pi_1^{NC}\} \end{aligned} \tag{41}$$

where we have made use of (39). Substituting (40) and (41) into (37), we obtain

$$\hat{\beta}_1^{IV} = \frac{s_{Zh}}{s_{Zg}} \xrightarrow{P} \frac{\text{var}(Z_i)\{\beta_1^{C} \pi_1^{C} + \beta_1^{NC} \pi_1^{NC}\}}{\text{var}(Z_i)\{\pi_1^{C} + \pi_1^{NC}\}} = \frac{\pi_1^{C}}{\pi_1^{C} + \pi_1^{NC}} \beta_1^{C} + \frac{\pi_1^{NC}}{\pi_1^{C} + \pi_1^{NC}} \beta_1^{NC} \tag{42}$$

which is Eq. (13) in the text, the expression for the LATE estimator when gun levels are observable. The LATE estimator is a weighted average of β_1^{C} and β_1^{NC}, but now the weights are the relative strengths of the correlations between the instrument Z and criminal/noncriminal gun prevalence. Note that, unlike the OLS estimator, the variation in gun prevalence does not affect the weights.

Now consider the case where gun levels are not observable and the IV estimator uses the proxy p:

$$\hat{b}_1^{IV} = \frac{s_{Zh}}{s_{Zp}} \xrightarrow{P} \frac{\text{cov}(Z_i, h_i)}{\text{cov}(Z_i, p_i)} \tag{43}$$

The numerator is the same as in (37) above. The denominator is

$$\begin{aligned}\operatorname{cov}(Z_i, p_i) &= \operatorname{cov}\left\{Z_i, \left(\delta_0 + \delta_1^{\mathrm{C}} g_{\mathrm{i}}^{\mathrm{C}} + \delta_1^{\mathrm{NC}} g_{\mathrm{i}}^{\mathrm{NC}} + v_i\right)\right\} \\ &= \delta_1^{\mathrm{C}} \operatorname{cov}(Z_i, g_i^{C}) + \delta_1^{\mathrm{NC}} \operatorname{cov}(Z_i, g_i^{NC}) \\ &= \delta_1^{\mathrm{C}} \pi_1^{\mathrm{C}} \operatorname{var}(Z_i) + \delta_1^{\mathrm{NC}} \pi_1^{\mathrm{NC}} \operatorname{var}(Z_i) = \operatorname{var}(Z_i)\left\{\delta_1^{\mathrm{C}} \pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}} \pi_1^{\mathrm{NC}}\right\}\end{aligned} \tag{44}$$

Substituting (40) and (44) into (43) yields

$$\begin{aligned}\hat{b}_1^{\mathrm{IV}} &= \frac{\mathrm{s_{Zh}}}{\mathrm{s_{Zp}}} \xrightarrow{\mathrm{P}} \frac{\operatorname{cov}(Z_i, h_i)}{\operatorname{cov}(Z_i, p_i)} = \frac{\operatorname{var}(\mathrm{Z_i})\left\{\beta_1^{\mathrm{C}} \pi_1^{\mathrm{C}} + \beta_1^{\mathrm{NC}} \pi_1^{\mathrm{NC}}\right\}}{\operatorname{var}(Z_i)\left\{\delta_1^{\mathrm{C}} \pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}} \pi_1^{\mathrm{NC}}\right\}} \\ &= \frac{\pi_1^{\mathrm{C}}}{\delta_1^{\mathrm{C}} \pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}} \pi_1^{\mathrm{NC}}} \beta_1^{\mathrm{C}} + \frac{\pi_1^{\mathrm{NC}}}{\delta^{\mathrm{C}} \pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}} \pi_1^{\mathrm{NC}}} \beta_1^{\mathrm{NC}} \\ &= \frac{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}}{\delta_1^{\mathrm{C}} \pi_1^{\mathrm{C}} + \delta_1^{\mathrm{NC}} \pi_1^{\mathrm{NC}}} \left\{\frac{\pi_1^{\mathrm{C}}}{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}} \beta_1^{\mathrm{C}} + \frac{\pi_1^{\mathrm{NC}}}{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}} \beta_1^{\mathrm{NC}}\right\}\end{aligned} \tag{45}$$

which is Eq. (15) in the text, the expression for the LATE estimator when gun prevalence is proxied by p. Note that this is a scaling parameter (assumed positive) times the probability limit of $\hat{\beta}_1^{\mathrm{IV}}$ given in Eq. (42). Thus sign $\{\widehat{\mathrm{b}}_1^{\mathrm{IV}}\}$ is a consistent estimator of sign $\{\hat{\beta}_1^{\mathrm{IV}}\}$, irrespective of the strength of the correlation between the proxy and criminal/noncriminal gun prevalence.

Appendix 3: Robustness Checks

Checks Using 1990 Data

We tested the robustness of our 1990 results by varying the main specification in a number of ways:

1. Weights. We estimated the main specification but weighted by county population in 1990. Because the results using population weights were sometimes sensitive to the inclusion/exclusion of a small number of large counties, we estimated using both the sample of counties with populations in excess of 25,000 as in the results discussed in the main text, and using a subset of this sample that excludes the roughly 100 counties with populations greater than 500,000.
2. Functional form. We varied the functional form of the estimating equation by using homicide rates in logs and in levels, and by using PSG in logs and in levels.
3. Lagged dependent variable (LDV). One of the anonymous referees suggested we include a lagged measure of the gun homicide rate in the equations to mitigate possible problems of unobserved heterogeneity, i.e., historical factors besides heterogeneity in the criminal population. We note, however, that although this is a useful robustness check, the results of such an estimation are not easily interpreted. In our preferred LATE model, the IV estimator has a very clear interpretation in the presence of heterogeneity in criminality: it is a weighted average of the effects of criminal and non-criminal gun prevalence. Including lagged homicide as a regressor eliminates this clear interpretation offered by our LATE model.
4. Criminal justice (CJ) controls. Another referee suggested our specifications may suffer from a specific form of omitted variable bias, namely failing to include controls for formal deterrence measures such as police manpower, incarceration rates, or arrest rates. We examine this possibility by re-estimating our model including as controls two of the most widely used measures in the macro-level deterrence literature and for

which data for all US counties are readily available: ICPSR county-level data on sworn police officers per capita in 1992[33] (as a measure of police manpower levels) and a measure of the rate of solving crimes constructed as the ratio of arrests for violent crimes 1989–91 to reports of violent crimes 1989–91.[34] We use both controls in log form. Unfortunately, incarceration data are not available for all US counties We should note, however, that to the extent that incarceration levels and other omitted criminal justice measures operate at the state-level to reduce county-level rates of homicide, these effects would be captured by our inclusion of state fixed effects.

We re-estimated the main equation using various combinations of the above specifications. The results for gun homicide are reported below in Table 10; the results for nongun homicide are in Table 11. The first row in each table corresponds to the specification discussed in the main text.

GMUR is the gun murder rate; NGMUR is the non-gun murder rate; PSG is defined as in the paper. "LDV" and "CJ" indicate whether a lagged dependent variable or criminal justice measures are included as regressors. In the LDV specifications, the table reports the long-run coefficient on PSG, equal to the coefficient on PSG * $1/(1 - \alpha)$ where α is the coefficient on lagged homicide. The magnitude of the long-run coefficient can therefore be compared directly to the coefficient on PSG when the LDV is omitted. The "Wt" column indicates whether or not the results weight by 1990 population. The total sample includes counties with a population of at least 25,000 persons in 1990; the "U lim" column indicates whether a subset of counties with a population upper limit of 500,000 persons is used. The HOLS column reports the coefficient on the gun prevalence proxy when it is treated as exogenous; the GMM2S column is the coefficient when treated as endogenous; 2-step efficient GMM is used in both cases. The "F 1st St" column reports the first-stage F statistic; J is the J overidentification statistic; p(J) is the corresponding p value; and N is the sample size. Stars are as in the paper (1, 5, 10 %). Tests are robust to heteroskedasticity and clustering.

The various specifications show that the main results reported in the paper are indeed robust. In the gun homicide estimations, when gun prevalence is treated as exogenous, the estimated impact on gun homicide is generally positive and statistically significant; when it is treated as endogenous, the impact on gun homicide is significantly negative or null. For nongun homicide, the impact of gun prevalence is generally null, whether or not gun prevalence is treated as exogenous or endogenous. The instrument relevance tests are generally satisfactory (a first-stage F statistic in excess of 10), as are the tests of instrument orthogonality (insignificant J statistics).

The different functional forms generate broadly similar qualitative results, but the specifications in which the homicide rate is in logs, as in the results discussed in the main text, tend to generate smaller quantitative impacts than those in which the homicide rate is used in levels. For the reasons discussed in the text, we regard the log specification, and

[33] This is the ICPSR dataset 2266, "Directory of Law Enforcement Agencies: 1992". We use total sworn police, aggregating agencies by county. We exclude agencies with state-wide jurisdiction. We do not calculate the measure for the 5 counties of New York City, because sworn police for NYC are not available in the dataset disaggregated by county. The per capita measure uses county population according to the 1990 census.

[34] The measure is constructed using the ICPSR datasets 9573, 9785 and 6036, "Uniform Crime Reporting Data [United States}: County-level Detailed Arrest and Offense Data". The numerator uses data on arrests (all ages) for violent crimes; the denominator uses the sum of reported murders, rapes, robberies and aggravated assaults. We aggregate over 1989–91 because of small numbers for some counties.

Table 10 1990 gun homicide estimates

Spec	Years	Dep var	PSG	LDV	CJ	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1.	1990	Log GMUR	PSG/100					0.70***	−2.41**	16.9	1.0	0.595	1,429
2.	1990	Log GMUR	PSG/100				500 k 1990	0.73***	−1.26	14.0	0.1	0.944	1,334
3.	1990	Log GMUR	PSG/100			Yes		0.94***	−1.25*	29.7	4.6	0.101	1,429
4.	1990	Log GMUR	PSG/100			Yes	500 k 1990	0.82***	−0.56	11.4	0.4	0.829	1,334
5.	1990	Log GMUR	Log PSG					0.38***	−1.16***	12.3	0.5	0.786	1,429
6.	1990	Log GMUR	Log PSG				500 k 1990	0.42***	−0.66	12.5	0.3	0.873	1,334
7.	1990	Log GMUR	Log PSG			Yes		0.25	−0.65**	25.4	4.1	0.126	1,429
8.	1990	Log GMUR	Log PSG			Yes	500 k 1990	0.44***	−0.39	8.9	0.3	0.874	1,334
9.	1990	GMUR	PSG/100					3.68***	−28.09***	14.0	0.6	0.731	1,449
10.	1990	GMUR	PSG/100				500 k 1990	3.14***	−20.19**	11.0	0.6	0.755	1,354
11.	1990	GMUR	PSG/100			Yes		8.25***	−34.99***	29.3	0.5	0.770	1,449
12.	1990	GMUR	PSG/100			Yes	500 k 1990	3.92***	−19.98	11.0	1.7	0.422	1,354
13.	1990	GMUR	Log PSG					2.08***	−12.35***	10.6	0.4	0.819	1,449
14.	1990	GMUR	Log PSG				500 k 1990	1.58***	−11.26**	10.2	0.8	0.684	1,354
15.	1990	GMUR	Log PSG			Yes		3.25**	−14.80***	24.8	0.3	0.856	1,449
16.	1990	GMUR	Log PSG			Yes	500 k 1990	2.13***	−6.65	8.7	2.5	0.280	1,354
17.	1990	Log GMUR	PSG/100	Yes				0.90***	−2.00	13.2	0.6	0.734	1,342
18.	1990	Log GMUR	PSG/100	Yes			500 k 1990	0.89***	−0.97	10.9	0.5	0.792	1,247
19.	1990	Log GMUR	PSG/100	Yes		Yes		1.31***	−0.11	28.4	5.7	0.058	1,342
20.	1990	Log GMUR	PSG/100	Yes		Yes	500 k 1990	0.93***	0.79	9.9	0.1	0.952	1,247
21.	1990	Log GMUR	Log PSG	Yes				0.49***	−0.95	10.1	0.5	0.793	1,342
22.	1990	Log GMUR	Log PSG	Yes			500 k 1990	0.52***	−0.45	9.8	0.6	0.728	1,247
23.	1990	Log GMUR	Log PSG	Yes		Yes		0.43*	−0.16	26.2	5.6	0.061	1,342
24.	1990	Log GMUR	Log PSG	Yes		Yes	500 k 1990	0.48***	0.29	7.9	0.2	0.921	1,247
25.	1990	GMUR	PSG/100	Yes				4.25***	−27.83***	12.6	0.3	0.872	1,448
26.	1990	GMUR	PSG/100	Yes			500 k 1990	3.55***	−21.28**	10.0	1.3	0.534	1,353

Table 10 continued

Spec	Years	Dep var	PSG	LDV	CJ	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
27.	1990	GMUR	PSG/100	Yes		Yes		6.81**	−34.71**	24.0	1.8	0.410	1,448
28.	1990	GMUR	PSG/100	Yes		Yes	500 k 1990	5.52**	−6.82	10.7	4.6	0.099	1,353
29.	1990	GMUR	Log PSG	Yes				2.17**	−11.20***	9.8	1.7	0.433	1,448
30.	1990	GMUR	Log PSG	Yes			500 k 1990	1.83**	−10.51*	9.1	2.7	0.255	1,353
31.	1990	GMUR	Log PSG	Yes		Yes		1.41	−15.80**	21.7	1.8	0.405	1,448
32.	1990	GMUR	Log PSG	Yes		Yes	500 k 1990	2.95**	−0.35	8.3	5.1	0.079	1,353
33.	1990	Log GMUR	PSG/100		Yes			0.72***	−1.57	17.7	1.4	0.489	1,398
34.	1990	Log GMUR	PSG/100		Yes		500 k 1990	0.74***	−0.81	15.0	0.3	0.878	1,307
35.	1990	Log GMUR	PSG/100		Yes	Yes		0.78***	−0.53	23.8	5.8	0.054	1,398
36.	1990	Log GMUR	PSG/100		Yes	Yes	500 k 1990	0.81***	−0.25	12.1	0.2	0.891	1,307
37.	1990	Log GMUR	Log PSG		Yes			0.39***	−0.86*	16.2	1.1	0.572	1,398
38.	1990	Log GMUR	Log PSG		Yes		500 k 1990	0.43***	−0.42	13.5	0.4	0.838	1,307
39.	1990	Log GMUR	Log PSG		Yes	Yes		0.36***	−0.35	33.3	5.4	0.066	1,398
40.	1990	Log GMUR	Log PSG		Yes	Yes	500 k 1990	0.44***	−0.21	9.2	0.2	0.909	1,307
41.	1990	GMUR	PSG/100		Yes			3.42***	−19.37***	14.4	0.5	0.790	1,418
42.	1990	GMUR	PSG/100		Yes		500 k 1990	2.93***	−15.48*	11.6	0.2	0.904	1,327
43.	1990	GMUR	PSG/100		Yes	Yes		8.89***	−26.10**	23.7	1.1	0.591	1,418
44.	1990	GMUR	PSG/100		Yes	Yes	500 k 1990	3.91***	−23.97	11.4	0.9	0.623	1,327
45.	1990	GMUR	Log PSG		Yes			1.72***	−9.49***	13.6	0.0	0.985	1,418
46.	1990	GMUR	Log PSG		Yes		500 k 1990	1.51***	−8.86**	10.7	0.1	0.975	1,327
47.	1990	GMUR	Log PSG		Yes	Yes		5.02***	−13.14**	32.2	0.7	0.722	1,418
48.	1990	GMUR	Log PSG		Yes	Yes	500 k 1990	2.14***	−9.32	8.9	1.6	0.448	1,327

* 10 %, ** 5 %, *** 1 %

Table 11 1990 nongun homicide estimates

Spec	Years	Dep var	PSG	LDV	CJ	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1.	1990	Log NGMUR	PSG/100					−0.16	−0.31	13.6	2.3	0.322	1,410
2.	1990	Log NGMUR	PSG/100				500 k 1990	−0.14	−0.58	10.8	1.8	0.413	1,315
3.	1990	Log NGMUR	PSG/100			Yes		−0.22	0.35	29.4	3.6	0.169	1,410
4.	1990	Log NGMUR	PSG/100			Yes	500 k 1990	−0.20	0.30	10.5	0.9	0.636	1,315
5.	1990	Log NGMUR	Log PSG					−0.09	−0.06	10.1	2.4	0.296	1,410
6.	1990	Log NGMUR	Log PSG				500 k 1990	−0.08	−0.20	9.8	2.0	0.376	1,315
7.	1990	Log NGMUR	Log PSG			Yes		−0.08	0.21	24.9	3.2	0.204	1,410
8.	1990	Log NGMUR	Log PSG			Yes	500 k 1990	−0.10	0.26	8.4	0.8	0.682	1,315
9.	1990	NGMUR	PSG/100					−0.48	−6.60**	14.0	0.4	0.808	1,449
10.	1990	NGMUR	PSG/100				500 k 1990	−0.39	−4.49	11.0	0.3	0.847	1,354
11.	1990	NGMUR	PSG/100			Yes		−0.67	−3.77	29.3	2.9	0.232	1,449
12.	1990	NGMUR	PSG/100			Yes	500 k 1990	−0.17	−2.42	11.0	2.2	0.335	1,354
13.	1990	NGMUR	Log PSG					−0.44	−2.62**	10.6	1.3	0.510	1,449
14.	1990	NGMUR	Log PSG				500 k 1990	−0.28	−2.34	10.2	0.8	0.686	1,354
15.	1990	NGMUR	Log PSG			Yes		−0.58*	−1.47	24.8	3.4	0.184	1,449
16.	1990	NGMUR	Log PSG			Yes	500 k 1990	−0.09	−0.42	8.7	2.5	0.292	1,354
17.	1990	Log NGMUR	PSG/100	Yes				−0.04	0.37	15.3	1.9	0.394	1,227
18.	1990	Log NGMUR	PSG/100	Yes			500 k 1990	−0.06	−0.04	14.0	1.3	0.514	1,132
19.	1990	Log NGMUR	PSG/100	Yes		Yes		−0.07	0.85	29.8	3.9	0.141	1,227
20.	1990	Log NGMUR	PSG/100	Yes		Yes	500 k 1990	−0.03	1.17	10.3	0.2	0.889	1,132
21.	1990	Log NGMUR	Log PSG	Yes				0.02	0.23	11.4	1.7	0.438	1,227
22.	1990	Log NGMUR	Log PSG	Yes			500 k 1990	0.00	0.06	12.6	1.3	0.515	1,132
23.	1990	Log NGMUR	Log PSG	Yes		Yes		0.00	0.44	26.7	3.2	0.199	1,227
24.	1990	Log NGMUR	Log PSG	Yes		Yes	500 k 1990	0.00	0.66	8.4	0.1	0.943	1,132
25.	1990	NGMUR	PSG/100	Yes				−0.51	−6.64**	13.6	1.2	0.553	1,448
26.	1990	NGMUR	PSG/100	Yes			500 k 1990	−0.41	−5.43	10.3	0.6	0.736	1,353

Table 11 continued

Spec	Years	Dep var	PSG	LDV	CJ	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
27.	1990	NGMUR	PSG/100	Yes		Yes		−1.09	−3.66	28.7	6.3	0.043	1,448
28.	1990	NGMUR	PSG/100	Yes		Yes	500 k 1990	−0.24	−4.41	10.5	3.0	0.218	1,353
29.	1990	NGMUR	Log PSG	Yes				−0.42	−2.58**	10.6	2.8	0.246	1,448
30.	1990	NGMUR	Log PSG	Yes			500 k 1990	−0.29	−2.91	9.6	1.2	0.560	1,353
31.	1990	NGMUR	Log PSG	Yes		Yes		−0.56	−1.28	28.3	7.1	0.029	1,448
32.	1990	NGMUR	Log PSG	Yes		Yes	500 k 1990	−0.15	−1.56	8.5	3.6	0.165	1,353
33.	1990	Log NGMUR	PSG/100		Yes			−0.15	−0.30	13.8	2.2	0.328	1,379
34.	1990	Log NGMUR	PSG/100		Yes		500 k 1990	−0.15	−0.66	11.2	1.6	0.455	1,288
35.	1990	Log NGMUR	PSG/100		Yes	Yes		−0.29	0.36	23.7	3.0	0.223	1,379
36.	1990	Log NGMUR	PSG/100		Yes	Yes	500 k 1990	−0.15	0.42	11.0	1.0	0.601	1,288
37.	1990	Log NGMUR	Log PSG		Yes			−0.08	−0.02	12.9	2.4	0.307	1,379
38.	1990	Log NGMUR	Log PSG		Yes		500 k 1990	−0.07	−0.22	10.2	1.8	0.406	1,288
39.	1990	Log NGMUR	Log PSG		Yes	Yes		−0.13	0.23	32.1	2.7	0.265	1,379
40.	1990	Log NGMUR	Log PSG		Yes	Yes	500 k 1990	−0.07	0.34	8.5	0.8	0.665	1,288
41.	1990	NGMUR	PSG/100		Yes			−0.61	−4.80	14.4	0.4	0.807	1,418
42.	1990	NGMUR	PSG/100		Yes		500 k 1990	−0.56	−3.82	11.6	0.3	0.881	1,327
43.	1990	NGMUR	PSG/100		Yes	Yes		−0.70	−2.89	23.7	2.9	0.235	1,418
44.	1990	NGMUR	PSG/100		Yes	Yes	500 k 1990	−0.21	−2.67	11.4	1.7	0.425	1,327
45.	1990	NGMUR	Log PSG		Yes			−0.43	−2.16	13.6	0.9	0.652	1,418
46.	1990	NGMUR	Log PSG		Yes		500 k 1990	−0.37	−2.16	10.7	0.4	0.818	1,327
47.	1990	NGMUR	Log PSG		Yes	Yes		−0.52	−0.94	32.2	3.5	0.177	1,418
48.	1990	NGMUR	Log PSG		Yes	Yes	500 k 1990	−0.11	−0.82	8.9	1.9	0.380	1,327

* 10 %, ** 5 %, *** 1 %

Table 12 1970 gun homicide estimates

Spec	Years	Dep var	PSG	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1.	1970	Log GMUR	PSG/100			0.66***	−4.76	2.1	1.2	0.275	1,288
2.	1970	Log GMUR	PSG/100		500 k 1990	0.70***	−2.67	0.6	0.7	0.408	1,193
3.	1970	Log GMUR	PSG/100			0.70***	−4.26	0.9	0.5	0.484	1,127
4.	1970	Log GMUR	PSG/100		500 k 1970	0.73***	7.15	0.0	0.2	0.687	1,052
5.	1970	Log GMUR	PSG/100	Yes		0.69***	−6.38*	3.8	0.1	0.707	1,288
6.	1970	Log GMUR	PSG/100	Yes	500 k 1990	0.83***	−5.67	0.6	0.5	0.487	1,193
7.	1970	Log GMUR	PSG/100	Yes		0.69***	−7.20*	2.9	0.2	0.696	1,127
8.	1970	Log GMUR	PSG/100	Yes	500 k 1970	0.83***	−14.83	0.0	0.5	0.479	1,052
9.	1970	Log GMUR	Log PSG			0.33***	−1.86	4.0	0.6	0.429	1,287
10.	1970	Log GMUR	Log PSG		500 k 1990	0.38***	−2.05	0.9	0.4	0.517	1,192
11.	1970	Log GMUR	Log PSG			0.33***	−1.16	3.4	0.4	0.523	1,127
12.	1970	Log GMUR	Log PSG		500 k 1970	0.39***	−0.92	0.3	0.3	0.612	1,052
13.	1970	Log GMUR	Log PSG	Yes		0.25***	−1.95*	7.8	0.1	0.709	1,287
14.	1970	Log GMUR	Log PSG	Yes	500 k 1990	0.40***	−1.86	1.4	0.7	0.410	1,192
15.	1970	Log GMUR	Log PSG	Yes		0.24***	−1.96*	7.1	0.2	0.664	1,127
16.	1970	Log GMUR	Log PSG	Yes	500 k 1970	0.39***	−2.11	0.6	0.8	0.371	1,052
17.	1970	GMUR	PSG/100			3.05***	−38.01	1.9	0.7	0.396	1,379
18.	1970	GMUR	PSG/100		500 k 1990	2.86***	−43.11	0.5	0.7	0.395	1,284
19.	1970	GMUR	PSG/100			2.20***	−32.85	1.1	0.0	0.906	1,198
20.	1970	GMUR	PSG/100		500 k 1970	2.05**	−59.60	0.1	0.0	0.963	1,123
21.	1970	GMUR	PSG/100	Yes		3.67***	−34.37*	3.9	0.3	0.567	1,379
22.	1970	GMUR	PSG/100	Yes	500 k 1990	2.71***	−84.48	0.7	0.4	0.514	1,284
23.	1970	GMUR	PSG/100	Yes		3.81***	−38.76*	3.2	0.3	0.604	1,198
24.	1970	GMUR	PSG/100	Yes	500 k 1970	2.29**	−153.20	0.1	0.2	0.689	1,123
25.	1970	GMUR	Log PSG			1.33***	−13.15*	4.1	0.2	0.628	1,378

Table 12 continued

Spec	Years	Dep var	PSG	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
26.	1970	GMUR	Log PSG		500 k 1990	1.27***	−26.23	0.8	0.3	0.578	1,283
27.	1970	GMUR	Log PSG			1.10***	−8.99	3.6	0.1	0.707	1,198
28.	1970	GMUR	Log PSG		500 k 1970	0.99**	−18.14	0.5	0.2	0.686	1,123
29.	1970	GMUR	Log PSG	Yes		1.76***	−10.61**	7.9	0.3	0.575	1,378
30.	1970	GMUR	Log PSG	Yes	500 k 1990	1.23***	−26.92	1.5	0.9	0.349	1,283
31.	1970	GMUR	Log PSG	Yes		1.81***	−10.98**	7.5	0.3	0.598	1,198
32.	1970	GMUR	Log PSG	Yes	500 k 1970	1.01**	−23.79	0.9	1.3	0.250	1,123

* 10 %, ** 5 %, *** 1 %

Table 13 1970 nongun homicide estimates

Spec	Years	Dep var	PSG	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1.	1970	Log NGMUR	PSG/100			0.04	1.85	1.5	0.2	0.656	1,195
2.	1970	Log NGMUR	PSG/100		500 k 1990	0.06	8.45	0.3	1.4	0.238	1,100
3.	1970	Log NGMUR	PSG/100			−0.06	−0.54	0.5	0.0	0.938	1,067
4.	1970	Log NGMUR	PSG/100		500 k 1970	0.00	−30.17	0.0	0.0	0.927	992
5.	1970	Log NGMUR	PSG/100	Yes		−0.13	−3.74	3.6	0.0	0.937	1,195
6.	1970	Log NGMUR	PSG/100	Yes	500 k 1990	0.10	0.88	0.5	0.8	0.361	1,100
7.	1970	Log NGMUR	PSG/100	Yes		−0.22	−5.58	2.8	0.0	0.945	1,067
8.	1970	Log NGMUR	PSG/100	Yes	500 k 1970	0.01	13.12	0.0	0.3	0.599	992
9.	1970	Log NGMUR	Log PSG			−0.03	0.65	4.2	0.1	0.710	1,194
10.	1970	Log NGMUR	Log PSG		500 k 1990	0.02	3.90	0.7	0.9	0.339	1,099
11.	1970	Log NGMUR	Log PSG			−0.08	−0.12	2.8	0.0	0.944	1,067
12.	1970	Log NGMUR	Log PSG		500 k 1970	−0.01	10.76	0.1	0.0	0.902	992
13.	1970	Log NGMUR	Log PSG	Yes		−0.13	−1.17	7.5	0.0	0.916	1,194
14.	1970	Log NGMUR	Log PSG	Yes	500 k 1990	0.01	0.14	1.3	0.8	0.367	1,099
15.	1970	Log NGMUR	Log PSG	Yes		−0.17	−1.59*	6.9	0.0	0.973	1,067
16.	1970	Log NGMUR	Log PSG	Yes	500 k 1970	−0.03	0.12	0.6	1.3	0.260	992
17.	1970	NGMUR	PSG/100			−0.22	−15.53	1.9	0.0	0.870	1,379
18.	1970	NGMUR	PSG/100		500 k 1990	−0.11	−18.68	0.5	0.8	0.381	1,284
19.	1970	NGMUR	PSG/100			−0.55	−29.39	1.1	0.0	0.885	1,198
20.	1970	NGMUR	PSG/100		500 k 1970	−0.34	−33.97	0.1	0.1	0.728	1,123
21.	1970	NGMUR	PSG/100	Yes		−1.51**	−26.25**	3.9	0.6	0.446	1,379
22.	1970	NGMUR	PSG/100	Yes	500 k 1990	−0.04	−13.03	0.7	1.0	0.316	1,284
23.	1970	NGMUR	PSG/100	Yes		−1.93**	−34.28**	3.2	0.3	0.556	1,198
24.	1970	NGMUR	PSG/100	Yes	500 k 1970	−0.27	−54.83	0.1	0.3	0.589	1,123
25.	1970	NGMUR	Log PSG			−0.21	−4.69	4.1	0.4	0.542	1,378

Table 13 continued

Spec	Years	Dep var	PSG	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
26.	1970	NGMUR	Log PSG		500 k 1990	−0.01	−6.59	0.8	1.2	0.281	1,283
27.	1970	NGMUR	Log PSG			−0.36*	−7.15	3.6	0.3	0.566	1,198
28.	1970	NGMUR	Log PSG		500 k 1970	−0.12	−4.12	0.5	0.8	0.386	1,123
29.	1970	NGMUR	Log PSG	Yes		−1.24***	−7.80**	7.9	0.6	0.446	1,378
30.	1970	NGMUR	Log PSG	Yes	500 k 1990	−0.02	−3.87	1.5	1.5	0.217	1,283
31.	1970	NGMUR	Log PSG	Yes		−1.40***	−9.41***	7.5	0.4	0.531	1,198
32.	1970	NGMUR	Log PSG	Yes	500 k 1970	−0.13	−5.86	0.9	1.7	0.187	1,123

* 10 %, ** 5 %, *** 1 %

hence the corresponding quantitative results, as preferable. We discuss here, for illustration, the calibrations for the main specifications for gun homicide (full sample, unweighted, no criminal justice controls or LDV) corresponding to a one percentage point increase in PSG. (1) The specification in line 1 of Table 10 is the specification discussed in the main text; the coefficient of −2.41 on PSG/100 implies that a one percentage point increase in PSG would reduce the gun homicide rate by 0.01 times 2.41 or about 2.4 %. (2) When both the gun homicide rate and PSG are in logs, the estimated coefficient on log(PSG) in the main specification is −1.16 (line 5). A one percentage point increase in PSG evaluated at the sample mean of PSG (67 %; see Table 5) is equivalent to a 1/67th or 1.5 % increase; 0.015 times 1.16 means a fall in the gun homicide rate of 1.7 %. (3) When both the gun homicide rate and PSG are in levels, the coefficient on PSG/100 in the main specification is −28.09 (line 9). A one percentage point increase in PSG would therefore reduce gun homicide by 0.28 persons per 100,000 population. At the sample average of 4.11 gun homicides per 100,000 persons (Table 5), this is equivalent to 6.8 % fall in the gun homicide rate. (4) When the gun homicide rate is in levels and PSG is in logs, the estimated coefficient on log(PSG) in the main specification is −12.35 (line 13). A one percentage point increase in PSG, equivalent to a 1.5 % increase, implies that gun homicide would fall by 0.015 times 12.35 or 0.19 persons per 100,000, equivalent to a 4.5 % fall in the gun homicide rate evaluated at the sample average of 4.11.

Checks Using 1970 and 1980 Data

We constructed datasets for 1970 and 1980 using Census variables, CDC homicide and suicide data, and ICPSR election data. The only variable used in the main specification we did not have available for these earlier years is outdoor sports magazine subscriptions (OMAG), and hence only the voting and veterans variables were available as instruments.

The gun and nongun homicide results for 1970 are reported in Tables 12 and 13, respectively. Because of data availability constraints, some of the control variables are defined slightly differently from the 1990 sample. Age structure categories are 0–14, 15–19, 20–24, 25–44, 45–64, and 65 + . The cutoff for low-income households is $6,000. Inequality is based on household income cutoffs of $6,000 and $25,000. "Hispanic" is based on use of Spanish language. Homicide and PSG are based on 5-year averages, 1968–72. The voting instrument is the percentage voting Republican in the 1968 presidential election. Because 1960 CDC homicide data were not available to us, we do not report the robustness check using a lagged dependent variable. For comparability with the 1990 county coverage, we also report the results of limiting the sample of counties based on 1990 as well as on 1970 population.

The 1970 results for gun homicide are similar to those for 1990: when PSG is treated as exogenous, it has a positive and significant coefficient, but when it is treated as endogenous, the significance disappears or (in a few cases) the coefficient becomes negative and significant. The nongun homicide results are also similar to our 1990 results: PSG has a null or (in a few cases) a negative impact on nongun homicide whether treated as exogenous or endogenous. The overidentification statistics are satisfactory. However, the instruments are weak to very weak, and hence the results should be treated with some caution.

Robustness checks using 1980 data are reported below in Tables 14 and 15. Data definitions are the same as for the 1990 sample, except that the cutoff for low-income households is $8,000, and inequality is based on household income cutoffs of $8,000 and $50,000. Homicide and PSG are based on 5-year averages, 1978–82. The voting instrument

Table 14 1980 gun homicide estimates

Spec	Years	Dep var	PSG	LDV	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1.	1980	Log GMUR	PSG/100				0.35**	−4.18	1.8	4.2	0.041	1,411
2.	1980	Log GMUR	PSG/100			500 k 1990	0.35**	3.72	0.8	4.3	0.039	1,315
3.	1980	Log GMUR	PSG/100				0.29*	−2.66	3.2	4.7	0.031	1,338
4.	1980	Log GMUR	PSG/100			500 k 1980	0.29*	2.57	1.8	5.7	0.017	1,256
5.	1980	Log GMUR	PSG/100		Yes		0.70***	−7.63***	8.8	0.6	0.437	1,411
6.	1980	Log GMUR	PSG/100		Yes	500 k 1990	0.65***	−1.02	1.9	2.1	0.144	1,315
7.	1980	Log GMUR	PSG/100		Yes		0.66**	−7.58***	8.5	0.5	0.466	1,338
8.	1980	Log GMUR	PSG/100		Yes	500 k 1980	0.58***	−2.41	1.6	2.0	0.157	1,256
9.	1980	Log GMUR	Log PSG				0.11	−1.48	3.6	4.5	0.034	1,411
10.	1980	Log GMUR	Log PSG			500 k 1990	0.17*	0.73	2.1	7.0	0.008	1,315
11.	1980	Log GMUR	Log PSG				0.04	−1.24	4.7	4.8	0.029	1,338
12.	1980	Log GMUR	Log PSG			500 k 1980	0.11	0.56	3.3	6.6	0.010	1,256
13.	1980	Log GMUR	Log PSG		Yes		0.16	−2.64***	13.2	0.6	0.451	1,411
14.	1980	Log GMUR	Log PSG		Yes	500 k 1990	0.32***	−0.30	4.1	2.3	0.126	1,315
15.	1980	Log GMUR	Log PSG		Yes		0.14	−2.66***	12.6	0.5	0.468	1,338
16.	1980	Log GMUR	Log PSG		Yes	500 k 1980	0.26***	−0.86	3.2	2.4	0.119	1,256
17.	1980	GMUR	PSG/100				2.29***	−39.39	1.0	4.8	0.028	1,456
18.	1980	GMUR	PSG/100			500 k 1990	1.71**	28.17	0.6	3.1	0.078	1,360
19.	1980	GMUR	PSG/100				1.52*	−18.85	2.0	4.6	0.031	1,382
20.	1980	GMUR	PSG/100			500 k 1980	1.13	16.58	1.2	3.9	0.049	1,300
21.	1980	GMUR	PSG/100		Yes		8.24***	−89.02***	7.7	0.3	0.587	1,456
22.	1980	GMUR	PSG/100		Yes	500 k 1990	3.79***	−5.27	1.3	2.4	0.125	1,360
23.	1980	GMUR	PSG/100		Yes		8.07***	−87.82***	7.5	0.2	0.640	1,382
24.	1980	GMUR	PSG/100		Yes	500 k 1980	3.25***	−21.03	1.1	1.5	0.213	1,300
25.	1980	GMUR	Log PSG				1.31**	−14.32*	2.9	2.6	0.109	1,455
26.	1980	GMUR	Log PSG			500 k 1990	1.08**	−2.91	1.4	5.1	0.024	1,359

Table 14 continued

Spec	Years	Dep var	PSG	LDV	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
27.	1980	GMUR	Log PSG				0.78	−9.48	3.9	3.3	0.071	1,381
28.	1980	GMUR	Log PSG			500 k 1980	0.64	2.45	2.4	4.8	0.028	1,299
29.	1980	GMUR	Log PSG		Yes		3.27**	−28.84**	12.9	0.3	0.596	1,455
30.	1980	GMUR	Log PSG		Yes	500 k 1990	2.23***	−2.05	3.6	2.5	0.113	1,359
31.	1980	GMUR	Log PSG		Yes		3.15**	−28.93**	12.4	0.2	0.634	1,381
32.	1980	GMUR	Log PSG		Yes	500 k 1980	1.80***	−7.31	2.9	1.8	0.186	1,299
33.	1980	Log GMUR	PSG/100	Yes			0.41**	−9.31	0.9	1.5	0.228	1,315
34.	1980	Log GMUR	PSG/100	Yes		500 k 1990	0.39**	3.29	0.2	1.0	0.326	1,219
35.	1980	Log GMUR	PSG/100	Yes			0.34*	−4.97	2.7	2.9	0.088	1,256
36.	1980	Log GMUR	PSG/100	Yes		500 k 1980	0.32	1.11	1.1	5.6	0.018	1,174
37.	1980	Log GMUR	PSG/100	Yes	Yes		0.79**	−9.35***	8.6	0.9	0.339	1,315
38.	1980	Log GMUR	PSG/100	Yes	Yes	500 k 1990	0.67***	−5.24	1.3	2.3	0.127	1,219
39.	1980	Log GMUR	PSG/100	Yes	Yes		0.71**	−9.25***	8.8	0.9	0.351	1,256
40.	1980	Log GMUR	PSG/100	Yes	Yes	500 k 1980	0.54***	−5.74	1.4	1.9	0.163	1,174
41.	1980	Log GMUR	Log PSG	Yes			0.19	−2.33*	2.8	3.0	0.083	1,315
42.	1980	Log GMUR	Log PSG	Yes		500 k 1990	0.23*	−1.42	1.0	6.4	0.012	1,219
43.	1980	Log GMUR	Log PSG	Yes			0.12	−1.90	4.3	3.5	0.060	1,256
44.	1980	Log GMUR	Log PSG	Yes		500 k 1980	0.15	−0.84	2.6	6.3	0.012	1,174
45.	1980	Log GMUR	Log PSG	Yes	Yes		0.23	−3.03***	13.3	0.8	0.366	1,315
46.	1980	Log GMUR	Log PSG	Yes	Yes	500 k 1990	0.35***	−1.62	3.2	3.1	0.079	1,219
47.	1980	Log GMUR	Log PSG	Yes	Yes		0.19	−3.03***	12.9	0.8	0.373	1,256
48.	1980	Log GMUR	Log PSG	Yes	Yes	500 k 1980	0.26**	−2.07	3.0	2.8	0.093	1,174
49.	1980	GMUR	PSG/100	Yes			3.25**	−69.97	0.7	2.1	0.146	1,447
50.	1980	GMUR	PSG/100	Yes		500 k 1990	2.47**	33.13	0.4	2.6	0.109	1,351
51.	1980	GMUR	PSG/100	Yes			1.50	−33.23	2.2	2.0	0.155	1,375

Table 14 continued

Spec	Years	Dep var	PSG	LDV	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
52.	1980	GMUR	PSG/100	Yes		500 k 1980	1.08	12.06	1.3	2.8	0.097	1,293
53.	1980	GMUR	PSG/100	Yes	Yes		19.90*	−220.72	9.7	1.5	0.218	1,447
54.	1980	GMUR	PSG/100	Yes	Yes	500 k 1990	4.45**	−18.38	1.3	2.6	0.109	1,351
55.	1980	GMUR	PSG/100	Yes	Yes		18.74*	−225.25	9.8	1.5	0.224	1,375
56.	1980	GMUR	PSG/100	Yes	Yes	500 k 1980	2.51	−16.89	1.5	1.3	0.248	1,293
57.	1980	GMUR	Log PSG	Yes			2.05**	−20.14*	2.5	1.1	0.297	1,446
58.	1980	GMUR	Log PSG	Yes		500 k 1990	1.72**	−13.13	1.1	3.1	0.080	1,350
59.	1980	GMUR	Log PSG	Yes			0.75	−13.60	4.1	1.3	0.261	1,374
60.	1980	GMUR	Log PSG	Yes		500 k 1980	0.61	−1.66	2.7	2.6	0.104	1,292
61.	1980	GMUR	Log PSG	Yes	Yes		8.23	−58.55	16.7	1.4	0.240	1,446
62.	1980	GMUR	Log PSG	Yes	Yes	500 k 1990	2.67**	−5.56	3.4	2.8	0.092	1,350
63.	1980	GMUR	Log PSG	Yes	Yes		7.58	−59.61	16.1	1.4	0.239	1,374
64.	1980	GMUR	Log PSG	Yes	Yes	500 k 1980	1.32	−5.20	3.4	1.5	0.217	1,292

* 10 %, ** 5 %, *** 1 %

Table 15 1980 nongun homicide estimates

Years	Dep var	PSG	LDV	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1980	Log NGMUR	PSG/100				−0.16	−0.76	0.7	0.4	0.526	1,348
1980	Log NGMUR	PSG/100			500 k 1990	−0.15	4.77	0.2	0.0	0.827	1,252
1980	Log NGMUR	PSG/100				−0.23	−0.86	1.8	0.1	0.797	1,293
1980	Log NGMUR	PSG/100			500 k 1980	−0.21	0.96	1.0	0.0	0.878	1,211
1980	Log NGMUR	PSG/100		Yes		−0.12	−2.54	7.7	0.0	0.939	1,348
1980	Log NGMUR	PSG/100		Yes	500 k 1990	0.02	1.17	1.2	0.0	0.981	1,252
1980	Log NGMUR	PSG/100		Yes		−0.15	−2.45	7.7	0.0	0.972	1,293
1980	Log NGMUR	PSG/100		Yes	500 k 1980	0.06	−0.22	1.1	0.0	0.987	1,211
1980	Log NGMUR	Log PSG				−0.08	−0.23	2.8	0.4	0.513	1,348
1980	Log NGMUR	Log PSG			500 k 1990	−0.05	1.09	1.3	0.3	0.606	1,252
1980	Log NGMUR	Log PSG				−0.13	−0.30	3.9	0.0	0.844	1,293
1980	Log NGMUR	Log PSG			500 k 1980	−0.08	0.47	2.6	0.0	0.950	1,211
1980	Log NGMUR	Log PSG		Yes		−0.06	−0.85	12.7	0.0	0.878	1,348
1980	Log NGMUR	Log PSG		Yes	500 k 1990	0.04	0.40	3.4	0.0	0.983	1,252
1980	Log NGMUR	Log PSG		Yes		−0.08	−0.84	12.1	0.0	0.942	1,293
1980	Log NGMUR	Log PSG		Yes	500 k 1980	0.08	−0.08	2.9	0.0	0.988	1,211
1980	NGMUR	PSG/100				−0.66	−11.88	1.0	0.9	0.343	1,456
1980	NGMUR	PSG/100			500 k 1990	−0.54	1.34	0.6	0.0	0.827	1,360
1980	NGMUR	PSG/100				−0.59	−11.36*	2.0	0.4	0.529	1,382
1980	NGMUR	PSG/100			500 k 1980	−0.39	−3.00	1.2	0.0	0.889	1,300
1980	NGMUR	PSG/100		Yes		−0.57	−28.33***	7.7	0.0	0.952	1,456
1980	NGMUR	PSG/100		Yes	500 k 1990	−0.33	1.46	1.3	0.1	0.738	1,360
1980	NGMUR	PSG/100		Yes		−0.72	−27.52***	7.5	0.0	0.906	1,382
1980	NGMUR	PSG/100		Yes	500 k 1980	−0.35	−4.22	1.1	0.0	0.939	1,300
1980	NGMUR	Log PSG				−0.38	−3.82*	2.9	0.4	0.514	1,455

Table 15 continued

Years	Dep var	PSG	LDV	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1980	NGMUR	Log PSG			500 k 1990	−0.19	0.22	1.4	0.1	0.784	1,359
1980	NGMUR	Log PSG				−0.40*	−3.99**	3.9	0.1	0.764	1,381
1980	NGMUR	Log PSG			500 k 1980	−0.14	−1.27	2.4	0.1	0.791	1,299
1980	NGMUR	Log PSG		Yes		−0.72*	−9.59***	12.9	0.0	0.897	1,455
1980	NGMUR	Log PSG		Yes	500 k 1990	−0.07	0.46	3.6	0.1	0.725	1,359
1980	NGMUR	Log PSG		Yes		−0.81*	−9.49***	12.4	0.0	0.872	1,381
1980	NGMUR	Log PSG		Yes	500 k 1980	−0.08	−1.61	2.9	0.0	0.952	1,299
1980	Log NGMUR	PSG/100	Yes			0.01	−2.11	2.2	1.0	0.315	1,176
1980	Log NGMUR	PSG/100	Yes		500 k 1990	0.03	1.02	0.8	1.4	0.231	1,080
1980	Log NGMUR	PSG/100	Yes			−0.08	−1.03	4.8	1.0	0.309	1,139
1980	Log NGMUR	PSG/100	Yes		500 k 1980	−0.03	0.65	2.9	1.0	0.306	1,057
1980	Log NGMUR	PSG/100	Yes	Yes		−0.24	−2.70	11.0	0.2	0.659	1,176
1980	Log NGMUR	PSG/100	Yes	Yes	500 k 1990	−0.02	1.21	3.1	0.3	0.562	1,080
1980	Log NGMUR	PSG/100	Yes	Yes		−0.28	−2.56	10.9	0.2	0.664	1,139
1980	Log NGMUR	PSG/100	Yes	Yes	500 k 1980	0.04	−0.51	2.8	0.3	0.560	1,057
1980	Log NGMUR	Log PSG	Yes			0.04	−0.63	4.8	1.4	0.235	1,176
1980	Log NGMUR	Log PSG	Yes		500 k 1990	0.09	0.03	2.9	1.4	0.237	1,080
1980	Log NGMUR	Log PSG	Yes			−0.04	−0.45	6.3	1.1	0.297	1,139
1980	Log NGMUR	Log PSG	Yes		500 k 1980	0.04	0.09	5.6	1.0	0.314	1,057
1980	Log NGMUR	Log PSG	Yes	Yes		−0.11	−0.95	14.0	0.3	0.600	1,176
1980	Log NGMUR	Log PSG	Yes	Yes	500 k 1990	0.04	0.53	6.3	0.3	0.580	1,080
1980	Log NGMUR	Log PSG	Yes	Yes		−0.13	−0.92	13.2	0.2	0.630	1,139
1980	Log NGMUR	Log PSG	Yes	Yes	500 k 1980	0.08	−0.21	4.9	0.4	0.547	1,057
1980	NGMUR	PSG/100	Yes			−0.71	−7.75	0.6	1.3	0.261	1,447
1980	NGMUR	PSG/100	Yes		500 k 1990	−0.58	6.04	0.4	0.0	0.860	1,351

Table 15 continued

Years	Dep var	PSG	LDV	Wt	U lim	HOLS	GMM2S	F 1st St	J	p(J)	N
1980	NGMUR	PSG/100	Yes			−0.82	−9.29	2.1	0.6	0.446	1,375
1980	NGMUR	PSG/100	Yes		500 k 1980	−0.55	−2.09	1.3	0.0	0.992	1,293
1980	NGMUR	PSG/100	Yes	Yes		−0.04	−34.56***	6.7	0.0	0.943	1,447
1980	NGMUR	PSG/100	Yes	Yes	500 k 1990	−0.71	7.28	1.2	0.3	0.561	1,351
1980	NGMUR	PSG/100	Yes	Yes		−0.31	−33.48***	6.6	0.0	0.967	1,375
1980	NGMUR	PSG/100	Yes	Yes	500 k 1980	−0.69	−0.82	1.2	0.1	0.740	1,293
1980	NGMUR	Log PSG	Yes			−0.39	−3.12	2.4	0.7	0.395	1,446
1980	NGMUR	Log PSG	Yes		500 k 1990	−0.20	1.78	1.0	0.2	0.674	1,350
1980	NGMUR	Log PSG	Yes			−0.54*	−3.63	4.1	0.2	0.626	1,374
1980	NGMUR	Log PSG	Yes		500 k 1980	−0.23	−1.00	2.6	0.0	0.931	1,292
1980	NGMUR	Log PSG	Yes	Yes		−0.11	−11.13***	12.5	0.0	0.993	1,446
1980	NGMUR	Log PSG	Yes	Yes	500 k 1990	−0.25	2.41	3.3	0.4	0.532	1,350
1980	NGMUR	Log PSG	Yes	Yes		−0.23	−11.04***	11.8	0.0	0.995	1,374
1980	NGMUR	Log PSG	Yes	Yes	500 k 1980	−0.23	−0.54	2.9	0.1	0.740	1,292

* 10 %, ** 5 %, *** 1 %

is the percentage voting Republican in the 1980 presidential election. Lagged homicide is based on the 1968–72 5-year average. For comparability with the 1990 county coverage, we report the results of limiting the sample of counties based on 1990 as well as on 1980 population. The LDV specifications report the long-run impact of gun prevalence, calculated as noted above.

The results for both gun and nongun homicide in 1980 are similar to those for 1990 and 1970. As with the 1990 data, including a lagged dependent variable does not noticeably change the results. The main difference with the results in term of the specification tests are that J statistic is sometimes high enough to reject at the 5 % level, and that the instruments are weak less often. We note that when the first-stage F statistic is satisfactorily high (near or above 10), the J statistic is also satisfactorily low, and these particular estimations are consistent with our 1990 results (i.e., PSG typically has a negative and significant coefficient in the gun homicide estimations). Again, however, because of the weakness of the instruments, the results should be treated with some caution.

References

Angrist J, Imbens G (1995) Two-stage least squares estimates of average causal effects in models with variable treatment intensity. J Am Stat Assoc 90(430):431–442

Annest JL, Mercy JA, Gibson DR, Ryan GW (1995) National estimates of nonfatal firearm-related injuries. J Am Med Assoc 273:1749–1754

Audit Bureau of Circulations (1993) Supplementary data report, covering county paid circulation for gun and related sports magazines. Audit Bureau of Circulations, Schaumburg, IL

Ayres I, Donohue JJ III (2003) Shooting down the 'More Guns, Less Crime' hypothesis. Stanf Law Rev 55:1193–1312

Azrael D, Cook PJ, Miller M (2004) State and local prevalence firearms ownership: measurement structure, and trends. J Quant Criminol 20:43–62

Baum CF, Schaffer ME, Stillman S (2003) Instrumental variables and GMM: estimation and testing. Stata J 3:1–31

Baum CF, Schaffer ME, Stillman S (2007) Enhanced routines for instrumental variables/generalized method of moments estimation and testing. Stata J 7:465–506

Baum C, Schaffer ME, Stillman S (2008) IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation. http://ideas.repec.org/c/boc/bocode/s425401.html

Bordua DJ, Lizotte AJ (1979) Patterns of legal firearms ownership: a cultural and situational analysis of Illinois counties. Law Policy Quart 1:147–175

Clarke RV, Mayhew P (1988) The British gas suicide story and its criminological implications. In: Tonry M, Morris N (eds) Crime and justice, vol 10. University of Chicago Press, Chicago, pp 79–116

Cook PJ, Ludwig J (1997) Guns in America. Police Foundation, Washington, DC

Cook PJ, Ludwig J (2003a) Guns and burglary. In: Ludwig J, Cook PJ (eds) Evaluating gun policy. Brookings Institution Press, Washington, DC

Cook PJ, Ludwig J (2003b) Pragmatic gun policy. In: Cook and Ludwig (2003a), *op. cit*

Cook PJ, Ludwig J (2004) The social costs of gun ownership. NBER working paper 10736. http://www.nber.org/papers/w10736

Cook PJ, Ludwig J (2006) The social costs of gun ownership. J Public Econ 90:379–391

Deaton A (1997) The analysis of household surveys: a microeconometric approach to development policy. World Bank/Johns Hopkins University Press, Washington, DC

Decker SH, Curry GD, Catalano S, Watkins A (2005) Strategic approaches to Community Safety Initiative (SACSI) in St. Louis. Final Report. Research report submitted to the US Department of Justice, Document Number 210361, Award Number 2000-IJ-CX-K008, June. http://www.ncjrs.gov/pdffiles1/nij/grants/210361.pdf

Duggan M (2001) More guns, more crime. J Polit Econ 109:1086–1114

Duggan M (2003) Guns and suicide. In: Cook and Ludwig (2003a), *op. cit*

Frishman F (1971) On the arithmetic means and variances of products and ratios of random variables. Army Research Office, Durham, North Carolina, publication AD-785 623. http://handle.dtic.mil/100.2/AD0785623
Greene, WH (2008) Econometric analysis, 6th edn. Prentice Hall, New Jersey
Hayashi F (2000) Econometrics. Princeton University Press, Princeton
Heckman J (1997) Instrumental variables: a study of implicit behavioral assumptions used in making program evaluations. J Human Resourc 32(3):441–462
ICPSR (1995) General election data for the United States, 1950–1990 [Computer file]. ICPSR ed. Inter-university Consortium for Political and Social Research [producer and distributor], Ann Arbor, MI
Imbens G, Angrist J (1994) Identification and estimation of local average treatment effects. Econometrica 62(2):467–475
Kates DB, Mauser G (2007) Would banning firearms reduce murder and suicide? A review of international and some domestic evidence. Harv J Law nd Public Policy 30(2):649–694
Killias M (1993) Gun ownership, suicide, and homicide: an international perspective. In: del Frate A, Zvekic U, van Dijk JJM (eds) Understanding crime: experiences of crime and crime control. UNICRI, Rome, pp 289–303
Killias M, van Kesteren J, Rindlisbacher M (2001) Guns, violent crime, and suicide in 21 countries. Can J Criminol 43:429–448
Kleck G (1979) Capital punishment, gun ownership, and homicide. Am J Sociol 84:882–910
Kleck G (1984) The relationship between gun ownership levels and rates of violence in the United States. In: Kates DB Jr. (ed) Firearms and violence: issues of public policy. Ballinger, Cambridge, MA
Kleck G (1988) Crime control through the private use of armed force. Soc Probl 35:1–21
Kleck G (1997) Targeting guns: firearms and their control. Aldine, NY
Kleck G (2004) Measures of gun ownership levels for macro-level crime and violence research. J Res Crime Delinquency 41(1):3–36
Kleck G, DeLone M (1993) Victim resistance and offender weapon effects in robbery. J Quant Criminol 9:55–82
Kleck G, Hogan M (1999) A national case-control study of homicide offending and gun ownership. Soc Probl 46(2):275–293
Kleck G, Kates DB (2001) Armed: new perspectives on gun control. Prometheus, Amherst, NY
Kleck G, McElrath K (1991) The effects of weaponry on human violence. Soc Forces 69:669–692
Kleck G, Patterson EB (1993) The impact of gun control and gun ownership levels on violence rates. J Quant Criminol 9:249–288
Kovandzic T, Vieraitis LM, Yeisley MR (1998) The structural covariates of urban homicide. Criminology 36:569–600
Kovandzic T, Schaffer ME, Kleck G (2012) Gun prevalence, homicide rates and causality: a GMM approach to endogeneity bias. In: Gadd D, Karstedt S, Messner SF (eds) The Sage handbook of criminological research methods. Sage, London
Land K, McCall PL, Cohen LE (1990) Structural covariates of homicide rates. Am J Sociol 95:922–963
Lott JR Jr (2000) More guns, less crime, 2nd edn. University of Chicago Press, Chicago
Marvell TB, Moody CE Jr (1991) Age structure and crime rates: the conflicting evidence. J Quant Criminol 7(3):237–273
McDowall D, Loftin C (1983) Collective security and the demand for handguns. Am J Sociol 88:1146–1161
Moody CE, Marvell TB (2005) Guns and crime. South Econ J 71:720–736
Newey WK (1985) Generalized method of moments specification testing. J Econom 29:229–256
Okoro CA, Nelson DE, Mercy JA, Balluz LS, Crosby AE, Mokdad AH (2005) Prevalence of household firearms and firearm-storage practices in the 50 states and the District of Columbia. Pediatrics 116:e370–e376
Rice DC, Hemley DD (2002) The market for new handguns. J Law Econ 45:251–265
Sampson RJ (1986) Crime in cities. In: Reiss AJ Jr, Tonry M (eds) Communities and crime. University of Chicago Press, Chicago, pp 271–311
Schaffer ME (2007) XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. http://ideas.repec.org/c/boc/bocode/s456501.html
Sloan JH, Kellermann AL, Reay DT, Ferris JA, Koepsell T, Rivara FP, Rice C, Gray L, LoGerfo J (1990) Handgun regulations, crime, assaults and homicide. N Engl J Med 319:1256–1262
Southwick L Jr (2000) Self-defense with guns: the consequences. J Criminal Just 28:351–370
Staiger D, Stock JH (1997) Instrumental variables regression with weak instruments. Econometrica 65:557–586
Stock JH, Watson MW (2007) Introduction to econometrics, 2nd edn. Addison-Wesley, Pearson

Stock JH, Yogo M (2005) Testing for weak instruments in linear IV regression. In: Andrews DWK, Stock JH (eds) Identification and inference for econometric models: essays in honor of Thomas Rothenberg. Cambridge University Press, Cambridge, pp 80–108. Working paper version: NBER Technical Working Paper 284, 2002. http://www.nber.org/papers/T0284

Tark J, Kleck G (2004) Resisting crime: the effects of victim action on the outcomes of crimes. Criminology 42(4):861–909

US Bureau of the Census (1990) Census 1990 summary file 3 (SF3)—sample data, table P006 urban and rural. Retrieved 7 February 2005 from US Census http://factfinder.census.gov

US Bureau of the Census (1994) County and City Data Book, 1994. US Government Printing Office, Washington, DC

US Federal Bureau of Investigation (FBI) 1990–2000, 2006. Crime in the United States 1989 [–1999] Uniform crime reports. US Government Printing Office, Washington, DC

US National Center for Health Statistics (1997) Limited access versions of mortality detail files, 1987–1993, with location detail, supplied to third author. US Department of Health and Human Services, Hyattsville, MD

Vieraitis LM (2000) Income inequality, poverty, and violent crime: a review of the empirical evidence. Soc Pathol 6(1):24–45

White H (1994) Estimation, inference and specification analysis. Cambridge University Press, Cambridge

Wright JD, Rossi PH (1986) Armed and considered dangerous. Aldine, New York

Zimring FE, Hawkins G (1997) Crime is not the problem. NY, Oxford