ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Kovandzic, Tomislav; Schaffer, Mark E.

Working Paper

Estimating the causal effect of gun prevalence on homicide rates: a local average treatment effect approach

IZA Discussion Papers, No. 3589

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Kovandzic, Tomislav; Schaffer, Mark E. (2008) : Estimating the causal effect of gun prevalence on homicide rates: a local average treatment effect approach, IZA Discussion Papers, No. 3589, Institute for the Study of Labor (IZA), Bonn, https://nbn-resolving.de/urn:nbn:de:101:1-20080711142

This Version is available at:
https://hdl.handle.net/10419/34758

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 3589

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

Tomislav Kovandzic
Mark E. Schaffer
Gary Kleck

July 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

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

Tomislav Kovandzic
University of Texas at Dallas

Mark E. Schaffer
CERT, Heriot-Watt University, CEPR and IZA

Gary Kleck
Florida State University

Discussion Paper No. 3589
July 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 3589
July 2008

ABSTRACT

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

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, 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 U.S. 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. 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 beneficiaries of the reduced level of violence may include substantial numbers of (urban) criminals, the murders of whom decrease via spillovers from noncriminal gun prevalence.

JEL Classification: K42, C51, C52

Keywords: crime, homicide, gun levels, endogeneity

Corresponding author:

Mark Schaffer
Centre for Economic Reform and Transformation
Department of Economics
Heriot-Watt University
Edinburgh EH14 4AS
United Kingdom
E-mail: m.e.schaffer@hw.ac.uk

* Seminar audiences at IZA, EERC-Moscow, and the Universities of Aberdeen, Bologna, Dundee, Edinburgh and Strathclyde 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, and Jeff Wooldridge for helpful comments and discussions on modeling and econometric issues. The usual caveat applies.

1. Introduction

Guns are heavily involved in violence in America, especially homicide. In 2005, 68 percent of homicides were committed by criminals armed with guns (U.S. Federal Bureau of Investigation 2006). Probably an additional 100,000 to 150,000 individuals were medically treated for nonfatal gunshot wounds (Kleck, 1997, p. 5; Annest et al., 1995). Further, relative to other industrialized nations, the United States has higher rates of violent crime, both fatal and nonfatal, a larger private civilian gun stock, 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 reduce their chances of injury or death. Individual-level research (e.g., Kleck and McElrath, 1990; Kleck and DeLone, 1993; Tark and Kleck 2004) can assess such effects of gun use in crime incidents, but it is less useful for detecting deterrent effects of gun ownership among prospective victims. 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

[1] See, e.g., Sloan et al. (1990), Killias (1993), and Zimring and Hawkins (1999). 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).

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 in fact also include the impact of crime rates on gun levels, and estimates of the former will 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 largely 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 vs. 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 U.S. 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. Our results also suggest that the beneficiaries of the reduced level of violence may include substantial numbers of (urban) criminals, the murders of whom decrease via spillovers from noncriminal gun prevalence.

The paper is organized as follows. Section 2 critically reviews previous research on the gun-homicide relationship. In Section 3 we set out how we model heterogeneity of criminality and discuss estimation and specification testing in an IV/GMM/LATE framework. Section 4 describes the data and model specification that we use. Estimation results are presented in Section 5, and Section 6 concludes.

2. Prior research

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 disrupt an attack. The impact of the self-defense effect is ambiguous: gun use by potential victims may simply disrupt attacks, or may lead to escalation and fatalities either because victims shoot their attackers or provoke their attackers into shooting them. $\partial h_i / \partial g_i >$ or < 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

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

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.

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 (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"

[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.

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.

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

$$h_i = b_0 + b_1 p_i + e_i \quad (2)$$

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

$$p_i = \delta_0 + \delta_1 g_i + v_i \quad (3)$$

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, Cook and Miller 2004; Moody and Marvell 2005; Cook and Ludwig 2004, 2006) have also concluded that PSG is currently the best available proxy. Proxies for time-series variation in gun levels, on the other hand, are particularly poor, a point we return to shortly.

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. These studies have relied primarily or exclusively on proxies for gun prevalence, and have not explicitly addressed the heterogeneity in criminality problem.

Critics of the IV approach 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 and the notion of Granger causality to try to establish whether past gun levels helps predict current crime rates. Panel data models allow the researcher to control for all time-invariant unobserved heterogeneity through the use of fixed effects or first-differences. For example, in

$$\Delta h_{it} = a_1 \Delta h_{it-1} + b_1 \Delta p_{it-1} + u_{it} \quad (4)$$

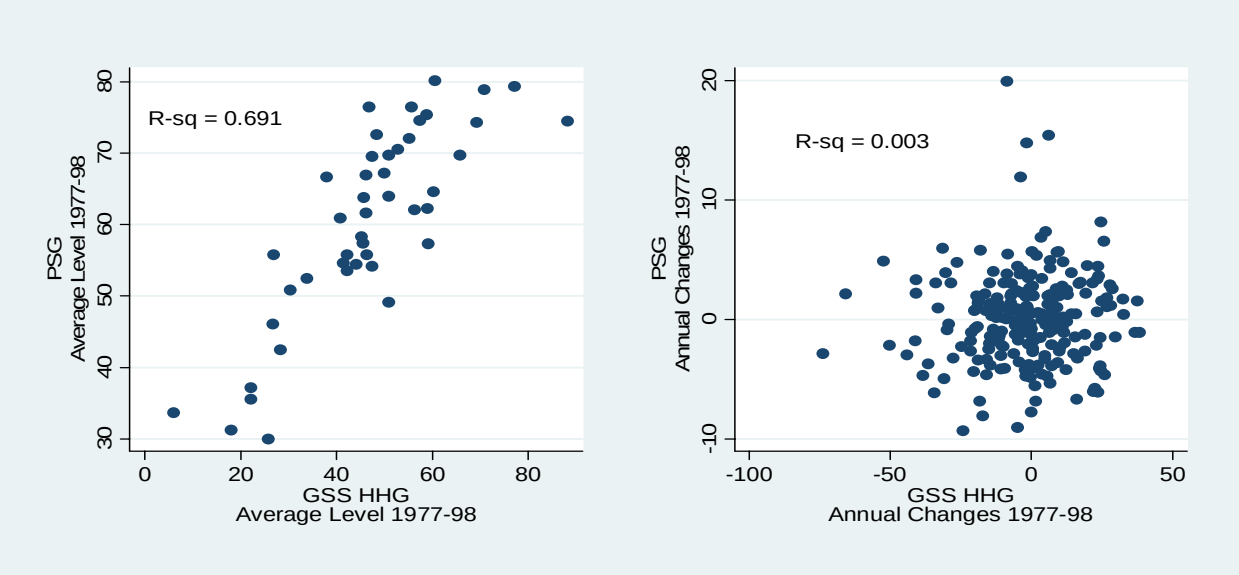
where Δ denotes the first-differences operator, the coefficient b_1 on Δp_{it-1} indicates whether lagged changes in the gun prevalence proxy predict current changes in homicide rates. A significant b_1 is interpreted as evidence of a causal effect of gun prevalence on homicide rates.

Moody and Marvell (2005) use first-differences on state level panel data and a combination of a direct measure of gun levels derived from the General Social Survey (GSS) with interpolations using the percentage of suicides with a gun (PSG) as their proxy; they find no effect of gun levels on homicide rates. Duggan (2001) uses first-differences on county-level panel data and *Guns & Ammo* magazine subscription rates (GAR) as a proxy for gun levels; Cook and Ludwig (2004, 2006) use fixed effects on county-level panel data and PSG as their gun proxy. Both Duggan and Cook-Ludwig find a positive impact of gun levels on homicide rates.

Although using time-series variation helps address the endogeneity problem, it aggravates the proxy problem, for two reasons. First, panel data methods amplify the measurement error problem inherent in using a proxy for gun prevalence instead of a direct measure (see, e.g., Griliches and Hausman (1986) for a general discussion). Second and more directly, the three panel studies between them used two gun level proxies, PSG and GAR, that show little or no time series correlation with the best available direct survey measure of gun prevalence at the national level, the GSS.

This is readily demonstrated using the Moody and Marvell (2005) dataset, who have kindly made their data publicly available.[3] Moody and Marvell report a high overall correlation of PSG with the percent of households with a gun (HHG) from the GSS (Moody and Marvell 2005, Table 1, p. 723), but this is driven entirely by the cross-sectional correlation between PSG and the GSS gun measure; the intertemporal correlations are tiny. The left-hand panel of Figure 1 below shows that state-average PSG and state-average HHG are highly correlated – the R^2 is 69%. The right-hand panel, on the other hand, shows that the correlation of first-differences in state PSG and state HHG is essentially nil, with an R^2 of approximately zero. The estimations reported by Cook and Ludwig (2004, pp. 13-14 and Table 2, p. 43) are similar: when they regress various GSS gun level measures on PSG, fixed effects for 9 Census Divisions, and in some specifications, time dummies, the t-statistics on PSG correspond to low R^2s of about 3-7%.[4]

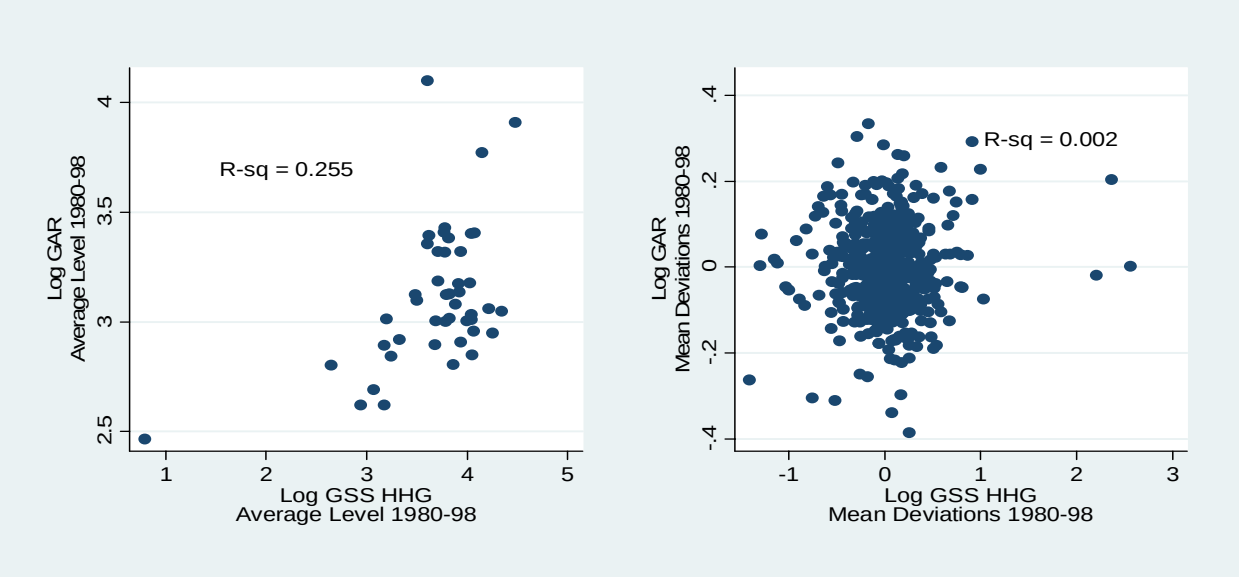
Figure 1: PSG and General Social Survey HHG, State-level Data 1977-98



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

[4] More precisely, these are "partial R^2s": the R^2 that would be obtained in a residual regression after partialing-out the fixed effects and time dummies. The t-statistics on PSG reported by Cook and Ludwig vary from 1.94 to 3.02 and refer to heteroskedastic-robust tests for a regression weighted by population. These can be converted into a partial-R^2-like statistics using the formula $R^2 = F/(df+F)$ where F is the F statistic (obtained as the square of the t-statistic) and df is the degrees of freedom of the test.

Figure 2: GAR and General Social Survey HHG, State-level Data 1980-98



The Moody-Marvell dataset also allows us to compare the cross-section and time-series validity of the proxy used by Duggan (2001), subscription rates to *Guns & Ammo* magazine. For consistency with Duggan's specification, we use logged values of HHG and GAR in mean-deviation form, i.e., with state-fixed effects partialed-out. The left-hand panel of Figure 2 shows that state-average log GAR and state-average log HHG are correlated, though not as strongly as were PSG and HHG (consistent with the consensus view that PSG is the superior proxy) – the R^2 is 26%. The right-hand panel, however, shows that the intertemporal variations in GAR and HHG are essentially uncorrelated, with an R^2 of approximately zero. The validation results reported by Duggan (2001, Table 3, column 4) are qualitatively similar: in a regression of log HHG on log GAR using state-level data for 1980-98, including both state and year fixed effects, and weighting observations by the number of GSS survey respondents, the t-statistic on log GAR is 3.10, which corresponds to an R^2 of about 2%.[5]

This problem of invalid time series proxies for gun levels can be attributed at least in part to the nature of the GSS. The GSS is available only annually or less since 1973, and has been shown to be affected by intertemporal variations in reporting bias, such as the artificial drop in gun prevalence in the

[5] This is also a "partial-R^2"; see note 4 above.

early 1990s following the enactment of the Brady laws. 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. This is apparent in Figures 1 and 2 above: in Figure 1, for example, the standard deviation for annual changes in GSS state-level HHG is 18 percentage points. The GSS survey data are, nevertheless, by consensus the best available direct data on gun prevalence in the US. The problem is therefore quite general. 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 the GSS measures. The unfortunate consequence of this is that researchers using panel data and relying on intertemporal variation in gun level proxies have not demonstrated that the proxies provide valid identifying variation.

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 the next section, we use a simple model to show that estimation of (3) or (4) 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 (3) or (4) by IV, as in Kleck and Patterson (1993), 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 an instrumental variables/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.

3. 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 (and also with gun prevalence), 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.[6] Most homicides of noncriminals are by criminals.

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

[6] 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.

numbers of criminals and noncriminals holding guns, respectively, per member of county i. We specify two DGPs, one for homicides of criminals (5a) and one for homicides of noncriminals (5b):

$$h_i^C = \beta_0^C + \beta_1^{C\text{-}C} g_i^C + \beta_1^{NC\text{-}C} g_i^{NC} + u_i^C \quad (5a)$$

$$h_i^{NC} = \beta_0^{NC} + \beta_1^{C\text{-}NC} g_i^C + \beta_1^{NC\text{-}NC} g_i^{NC} + u_i^{NC} \quad (5b)$$

In both equations, changes in criminal and noncriminal gun prevalence have separate impacts on homicide rates. Thus in equation (5a), the causal impact of criminal gun prevalence g_i^C on homicides of other criminals h_i^C is given by $\beta_1^{C\text{-}C}$, and $\beta_1^{NC\text{-}C}$ is the causal impact of noncriminal gun prevalence g_i^{NC} on homicides of criminals. In equation (5b), $\beta_1^{C\text{-}NC}$ and $\beta_1^{NC\text{-}NC}$ 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 $\beta_1^{C\text{-}C}$ 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 also expect $\beta_1^{C\text{-}NC}$ 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 stylized facts tell us less about what to expect for $\beta_1^{NC\text{-}NC}$ and $\beta_1^{C\text{-}NC}$. The parameter $\beta_1^{NC\text{-}NC}$ 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 $\beta_1^{NC\text{-}NC}$ may be positive (gun homicide "crimes of passion" dominate), negative (deterrence of criminal attacks dominates), or somewhere in between.

The sign of the parameter $\beta_1^{NC\text{-}C}$ is also ambiguous. The use of guns by noncriminals for self-defense against attacks by criminals could lead to fewer or more homicides, depending on whether the more likely result is disruption or escalation. However, there is also an interesting possibility that has not received much attention in the literature: noncriminal gun prevalence may have a negative deterrent effect

on homicides of criminals by other criminals, either by making non-local criminals less willing to travel to a locality to operate, or by making indigenous local criminals more cautious and less violent.

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 (5a) and (5b), we obtain

$$h_i = \beta_0 + \beta_1^C g_i^C + \beta_1^{NC} g_i^{NC} + u_i \qquad (5c)$$

where

$$h_i \equiv (h_i^C + h_i^{NC})$$
$$u_i \equiv (u_i^C + u_i^{NC})$$
$$\beta_0 \equiv (\beta_0^C + \beta_0^{NC})$$
$$\beta_1^C \equiv (\beta_1^{C\text{-}C} + \beta_1^{C\text{-}NC})$$
$$\beta_1^{NC} \equiv (\beta_1^{NC\text{-}C} + \beta_1^{NC\text{-}NC})$$

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.[7] The impact of noncriminal gun prevalence β_1^C is ambiguous, and depends on the relative magnitudes of negative deterrent effects of g_i^{NC} on homicides of noncriminals and criminals vs. the possible positive effects on homicides of criminals (self defense) and the positive effects on homicides 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 a very simple form (see Appendix 2): in a large sample it will be approximately the weighted average of the criminal and noncriminal impacts,

[7] The use of guns by criminals for self-defense against other criminals would reinforce this if it is a positive effect; if it is a negative effect, it would be insufficient to offset the positive criminality effects.

$$\beta_1^{ATE} \approx \text{plim}(\beta_{1i}) = \frac{\mu^C}{\mu^C + \mu^{NC}}\beta_1^C + \frac{\mu^{NC}}{\mu^C + \mu^{NC}}\beta_1^{NC} \quad (6)$$

where μ^C and μ^{NC} are the county means of criminal and noncriminal gun prevalence, E(g_i^C) and E(g_i^{NC}), respectively. The weights are the shares of guns in the population that are criminal and noncriminal.[8]

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 $g_i \equiv g_i^C + g_i^{NC}$ 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 equation (1), repeated here for convenience:

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

Previous researchers have interpreted OLS and IV estimates of β_1 in (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

$$\hat{\beta}_1^{OLS} \xrightarrow{P} \frac{\left\{\text{var}(g_i^C) + \text{cov}(g_i^C, g_i^{NC})\right\}}{\text{var}(g_i^C) + \text{var}(g_i^{NC}) + 2\,\text{cov}(g_i^C, g_i^{NC})}\beta_1^C + \frac{\left\{\text{var}(g_i^{NC}) + \text{cov}(g_i^C, g_i^{NC})\right\}}{\text{var}(g_i^C) + \text{var}(g_i^{NC}) + 2\,\text{cov}(g_i^C, g_i^{NC})}\beta_1^{NC} \quad (7)$$

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 vs. noncriminal gun prevalence. There is some

[8] The equality is approximate because E(g_i^C/($g_i^C + g_i^{NC}$)) ≠ E(g_i^C)/(E(g_i^C)+E(g_i^{NC})) and similarly for E(g_i^{NC}). However, plim(g_i^C/($g_i^C + g_i^{NC}$)) = plim(g_i^C)/(plim(g_i^C)+plim(g_i^{NC})), which gives us a large-sample justification for the approximation.

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 equation (7) are therefore likely to be positive, but of unknown magnitude. The magnitude of the variances in (7) 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 2003c).

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 (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:

$$g_i^C = \pi_0^C + \pi_1^C Z_i + \eta_i^C \quad (8)$$

$$g_i^{NC} = \pi_0^{NC} + \pi_1^{NC} Z_i + \eta_i^{NC} \quad (9)$$

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 equation (1) could be estimated by IV. We show in Appendix 2 that IV estimation of (1) generates an estimator of β_1 that converges in probability to a weighted average of β_1^C and β_1^{NC}:

$$\hat{\beta}_1^{IV} \xrightarrow{P} \frac{\pi_1^C}{\pi_1^C + \pi_1^{NC}}\beta_1^C + \frac{\pi_1^{NC}}{\pi_1^C + \pi_1^{NC}}\beta_1^{NC} \tag{10}$$

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 equation (10) 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 define a different LATE estimator and would converge to a different weighted average.

In the polar case of $\pi_1^C=0$, 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 $\pi_1^{NC}=0$, $\hat{\beta}_1^{IV}$ 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

$$p_i = \delta_0 + \delta_1^C g_i^C + \delta_1^{NC} g_i^{NC} + v_i \tag{11}$$

where 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

$$h_i = b_0 + b_1 p_i + e_i \tag{2}$$

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}:

$$\begin{aligned} \hat{b}_1^{IV} &\xrightarrow{P} \frac{\pi_1^C}{\delta_1^C\pi_1^C + \delta_1^{NC}\pi_1^{NC}}\beta_1^C + \frac{\pi_1^{NC}}{\delta_1^C\pi_1^C + \delta_1^{NC}\pi_1^{NC}}\beta_1^{NC} \\ &= \frac{\pi_1^C + \pi_1^{NC}}{\delta_1^C\pi_1^C + \delta_1^{NC}\pi_1^{NC}}\left\{\frac{\pi_1^C}{\pi_1^C + \pi_1^{NC}}\beta_1^C + \frac{\pi_1^{NC}}{\pi_1^C + \pi_1^{NC}}\beta_1^{NC}\right\} \end{aligned} \tag{12}$$

Equation (12) shows that $\hat{\beta}_1^{IV}$ and $\hat{b}_1^{IV}$ 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 $\hat{b}_1^{IV}$ as results for the sign of the infeasible LATE estimator $\hat{\beta}_1^{IV}$.

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, 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 b_1^{IV}, or the instruments are correlated with the error *u* in the data-generating process (5c).

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 (equation 10). Different instruments can therefore define different LATE parameters b_1^{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 in the DGP (5c), i.e., E(*Z'u*)=0. The second reason the null may be rejected is precisely because the orthogonality assumption E(*Z'u*)=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.[9] Intuitively, a failure of an instrument to be exogenous generates a bias in the estimate $\hat{b}_1^{IV}$ of the parameter b_1^{IV}. If other instruments are uncorrelated with the error *e* and generate consistent estimates of b_1^{IV}, the invalid and valid instruments will effectively identify different

[9] 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."

parameters, and 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 *e,* the test can detect this failure if the induced biases in the estimates of β_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 when it is false. The J statistic will be small when the null hypothesis of a single LATE parameter and valid instruments is correct; but it will also be small if the biases induced in $\hat{\beta}_1^{IV}$ 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.[10] 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. 233-34; Newey 1985), for the case when there are only 2 instruments and 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.[11]

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 asymptotic bias in $\hat{\beta}_1$.

[10] 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.

[11] 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.

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 can sign the direction of the induced bias in $\hat{b}_1^{IV}$, 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 $\hat{b}_1^{IV}$ 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).[12] 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.

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

As noted above, different instruments may be uncorrelated with the error *u* in the DGP (5c), 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 $\hat{b}_1^{IV}$, 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.[13] Again, we are not aware of any evidence that voting Republican should either be directly related to homicide rates or correlated more or less strongly with noncriminal or criminal gun prevalence. We note that Ayres and Donohue (2003, p. 1256) question the exogeneity of

[13] 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.

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.

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.[14] 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

[14] 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.

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 important and influential 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,

$$p_i = \theta_0 + \theta_1 Z_i + err_i \qquad (13)$$

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 $\hat{b}_1^{IV}$ 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, equation (12) reduces to

$$\hat{b}_1^{IV} \xrightarrow{P} \frac{\pi_1^{NC}}{\delta^{NC}\pi_1^{NC}}\left\{\frac{\pi_1^{NC}}{\pi_1^{NC}}\beta_1^{NC}\right\} = \frac{1}{\delta^{NC}}\beta_1^{NC} \tag{14}$$

We can apply an upper bound argument here, should we find that $\hat{b}_1^{IV} < 0$. Note that the weight on β_1^{NC} in equation (12) can be rewritten as:

$$\frac{1}{\delta^{NC}} * \frac{\pi_1^{NC}}{\frac{\delta^C}{\delta^{NC}}\pi_1^C + \pi_1^{NC}} \tag{15}$$

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 $\hat{b}_1^{IV}$ is an estimate of $(\beta_1^{NC} * 1/\delta_1^{NC})$ as in (14). As π_1^C increases, (15) shows that the weight on β_1^{NC} falls, and from (12) 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 $\hat{b}_1^{IV}$ will now exceed $(\beta_1^{NC} * 1/\delta_1^{NC})$, i.e., $\hat{b}_1^{IV}$ scaled by δ_1^{NC} provides an upper bound for β_1^{NC}.

4. Data and model specification

Data

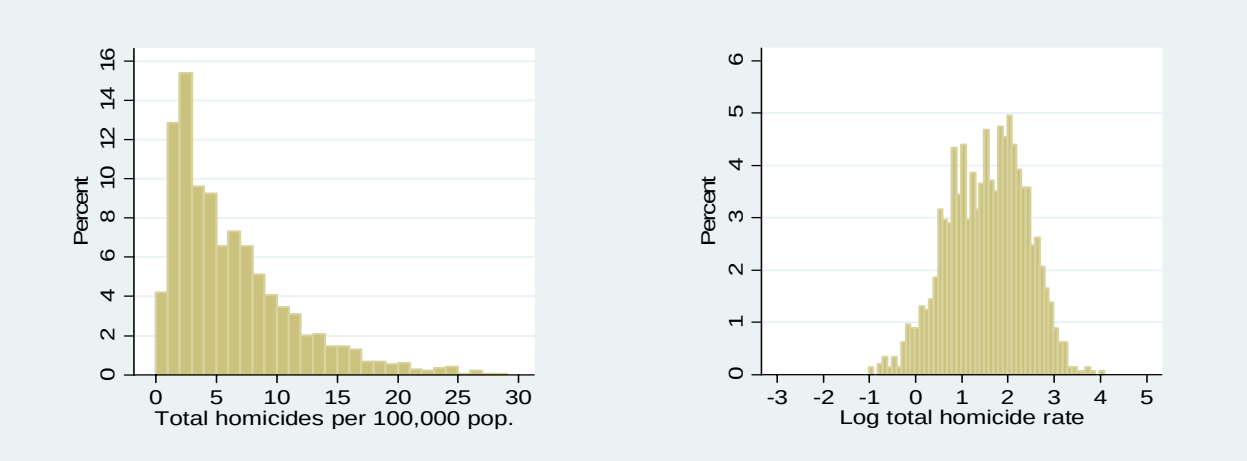
The dependent variables in our model are the gun and nongun homicide rates per 100,000 county population. Homicide rates are averages for the seven years 1987 to 1993, thus bracketing the census year of 1990 for which data on many of the control variables were available. Seven years were covered to reduce the influence of random year-to-year aberrations, e.g., misclassification of homicides as other kinds of deaths such as suicides or unintentional deaths. 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 U.S. 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 U.S. counties but over 90% of the U.S. 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.

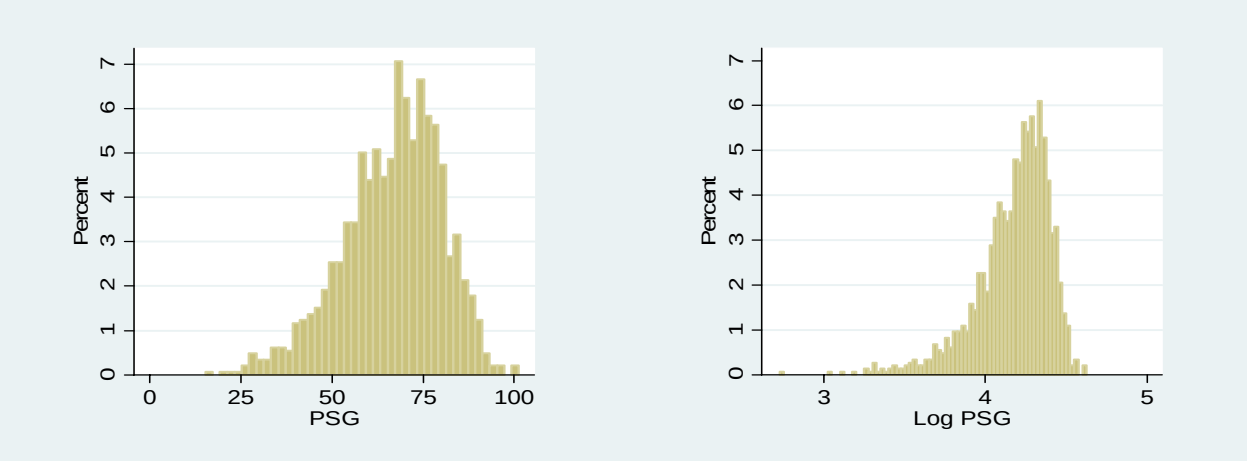
Although the simple model outlined in the previous section specifies untransformed homicide rates as the dependent variable, in our main estimations we specify homicide rates in logs. Figure 3 illustrates why: the distribution of the total homicide rate is skewed to the right with a number of outlier 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 equation (5c) as a simple additive aggregation of (5a) and (5b) is not available. There is also a "log of zero" problem, because even though we are using 7-year averages and excluding the smallest counties, a small number of counties have zero murders: of the 1,456 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, and using logged homicide rates where the zero murder counties are set to the lowest non-zero homicide rate instead of dropped.

Figure 3: Distributions of Total Homicide Rate and Log Total Homicide Rate



Note: 8 counties with homicide rates greater than 30 per 100,000 are omitted from the left-hand panel.

Figure 4: Distributions of PSG and Log PSG



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 4 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 Figure 1 suggests that PSG ≈ 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 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 δ_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 to 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 to 24 and 25 to 34, and the African-American share of the population because people in these groups are less likely to own guns, but more likely to commit homicide, than other people (Kleck, 1997; Cook and Ludwig, 1997; U.S. FBI, 2000). The other controls used were percent Hispanic, population density, average education level, unemployment rate, transient population (born out-of-state), vacant housing units, female-headed households with children, median household income, households earning less than $15,000, and inequality (ratio of households earning more than $75,000 to households earning less than $15,000). The sets of controls for rurality and age structure are 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 1 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 equation (2) 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.

It is common in this literature to weight observations by the population of the locality. One rationale for doing so is to correct for possible heteroskedasticity, although this will be the efficient feasible GLS procedure only if the heteroskedasticity takes a particular form (variance proportional to the square of the population). We report both weighted and unweighted results.

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 efficient in the presence of both arbitrary heteroskedasticity and arbitrary within-state correlation of the error. Particularly important for our purposes is that inference and testing using this estimator will be similarly robust, including GMM tests of exogeneity.

5. 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 Table 2. Column 1 reports the results of population-weighted 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. Some readers might be surprised that the size of the 18-24 age group is negatively associated with gun homicide rates, but this is actually a common finding (Marvell and Moody, 1991), perhaps because persons in this physically vigorous age group are harder to victimize. The overall fit of the regressions is quite good, with the population-weighted PSG-exogenous specification explaining 79% of the within-state variation in county-level log gun homicide rates; in the unweighted specification, the R^2 is still quite high, at 44%.[15]

The main result of interest is the coefficient of PSG/100. Column 1 of Table 2 shows that when PSG/100 is treated as exogenous, the estimated coefficient is 0.968 and is statistically significant at the 1% level. The calibration exercise above implies this would be a significant effect in practical as well statistical terms: an increase of 10 percentage points (0.10) in gun prevalence would imply an increase in gun homicide of about 10%. 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 that are most plausibly exogenous, 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 (on the grounds that the IV estimator is a special case of two-step GMM for homoskedastic and independent errors).[16] The Stock-Yogo test we report is the one for maximal size distortion, 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 test is based on the rejection rate r (10%, 15%, 20%) that the researcher is willing to accept

[15] *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.

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 3. 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. 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 2. 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 -2.08 that is significant at the 1% level.

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 2, column 1, when PSG is treated as exogenous, is 10.9 with a p-level of 1%. 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 2.31, with a corresponding p-value of 0.32. 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 J statistics for the

[16] Stock and Yogo do not tabulate critical values for the non-homoskedastic case. 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.

PSG-exogenous and PSG-endogenous estimations. The χ^2 test statistic reported in column 2 is 7.72, significant at the 1% level.[17] 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.

Table 4 reports the results using all three instruments (column 1) and each of the instruments individually (columns 2-4). 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 two of the three (using OMAG and VETS) are 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 3 show that, as expected, G&A is correlated with our gun level proxy PSG, though not as strongly as PCTREP88 and OMAG. Table 4, 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 11.4 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, $\hat{b}_1$=1.90, i.e., the estimated impact of gun prevalence is *positive* and comparable in magnitude to the $\hat{b}_1$ obtained when

[17] 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.

PSG is treated as exogenous.[18] 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.

The large positive estimate $\hat{b}_1$ 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. 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 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

[18] An exogeneity test of PSG using G&A as the sole instrument yields a $\chi^2(1)$ statistic of 1.40 with a p-value of 0.236. 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.

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 -2.1, implying that an increase of 10 percentage points in noncriminal gun prevalence would reduce gun homicide by at least ~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., $b_1 = 1.2$ and $\delta_1 = 0.7$ implies an upper bound for β_1^{NC} of about -0.8, so that a 10 percentage point increase in noncriminal gun prevalence would reduce gun homicide by ~8% 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 (equation 6). 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 $\beta_1^{C}\bar{g}^C$.[19]

One way to proceed is to use this range of weights and our estimates of β_1^{NC} to calculate what the elasticity would have to be for β_1^{ATE} to be approximately zero. A value of -2 for β_1^{NC} and an assumed criminal gun prevalence of 5% (criminal share of total=5/50=10%) would require β_1^{C}=18 for β_1^{ATE} to be zero. This implies an elasticity of gun homicide with respect to criminal gun prevalence of about $\beta_1^{C}\bar{g}^C$=18*0.05=0.9. Raising the assumed criminal gun prevalence to 13% (criminal share of total=13/50=25%) would require a value of 6 for β_1^{C} if β_1^{ATE} is to be zero; the implied elasticity $\beta_1^{C}\bar{g}^C$ is 6*0.13 or about 0.8. If we use a more conservative estimate of -1 for β_1^{NC} and $\bar{g}^C$=5%, β_1^{ATE}=0 would require β_1^{C}=9 which implies an elasticity of about 0.5. Raising the weight on criminal guns to 25% means that β_1^{C} would have to be 3 for β_1^{ATE} to be zero, and the implied elasticity is about 0.4. All of these implied elasticities appear plausible *a priori*.

[19] Obtained by differentiating equation (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.

Similar exercises suggest that an ATE that is positive and important in practical terms is about as plausible as one that is negative and important. For example, using a conservative estimate of β_1^{NC} = -1, β_1^{C} = 7, and $\bar{g}^{C}$ =13%, the gun homicide-criminal gun prevalence elasticity would be 7*0.13 or about 0.9, and β_1^{ATE} =(0.75*-1)+(0.25*7)=1, i.e., an X percentage point increase in general gun prevalence would raise gun homicides by about X%. Alternatively, using our point estimate of β_1^{NC} = -2, β_1^{C} = 7, and $\bar{g}^{C}$ =5%, we get a gun homicide-criminal gun prevalence elasticity of about 0.4, and β_1^{ATE} =(0.90*-2)+(0.10*7)=-1.1, which is just the opposite result – an X percentage point increase in general gun prevalence reduces gun homicides by about X%.

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.

Gun homicide estimations: robustness checks

We tested the robustness of our results in a number of ways. First, we estimated as above but without weighting by population. The key results were all unaffected by the change: the LATE estimate of b_1 was negative and statistically highly significant, and the specification testing generated the same results (OMAG, PCTREP88, VETS valid instruments; PSG endogenous; G&A an invalid instrument). The main differences vs. the weighted regressions were a more negative point estimate of b_1 when instrumented (-3.62, vs. -2.08 weighted), and first-stage regressions suggesting the instruments were weaker than in the weighted case.

We tried reestimating the gun homicide equation using log PSG instead of PSG as the proxy for gun levels, both with and without population weights. The qualitative results were essentially identical to those reported above: treated as exogenous, the coefficient on log PSG was positive and statistically significant; treated as endogenous, the coefficient became negative and significant at the 1% level; endogeneity tests of log PSG strongly rejected the null that log PSG could be treated as exogenous; and the J test of overidentifying restrictions safely failed to reject the null that the instruments were valid, with

p-values of 99% in the unweighted case and 41% in the weighted case. Log PCTREP, log OMAG and log VETS were about as relevant for log PSG as for the level of PSG.

We also estimated both the PSG and log PSG specifications using as the dependent variable the "one-sided winsorized" log gun homicide rate, so that the 20 zero gun homicide rate counties, instead of being dropped, had log crime rates set to the same value as the county with the lowest observed nonzero rate (-2.02, corresponding to 0.13 gun homicides per 100,000 population). The results were little changed from those that simply dropped the 20 counties.

The unweighted estimates using all three excluded instruments and PSG or log PSG bordered on being weakly identified, with first-stage F statistics of 12.4 (PSG) and 9.7 (log PSG). Recent research (see, e.g., Stock et al. 2002) suggests that there are various estimators that are at least partially robust to weak instruments. One such estimator is the GMM continuous-updating estimator (CUE) of Hansen et al. (1996). As a last robustness check, we estimated both our PSG and log PSG specifications with all three excluded instruments using CUE.[20] The estimates were essentially identical to those using 2-step efficient GMM. Moreover, the first-stage F statistics using all three excluded instruments are well in excess of the Stock-Yogo (2005) critical value of 6.5 for 10% maximal size for the LIML estimator (on the grounds that the LIML estimator is a special case of CUE for homoskedastic and independent errors). We conclude that weak instruments are not a major problem for the unweighted as well as the weighted results.

Who benefits from noncriminal deterrence? Spillovers and urban crime

The results above suggest that noncriminal gun prevalence has a negative impact on total gun homicide. One of the stylized facts cited in Section 3, however, is that most homicides victims are themselves criminals. Our results can be reconciled with this stylized fact if we also recall that noncriminal gun prevalence can deter homicides of criminals as well as noncriminals. It is possible that,

[20] We employ CUE after partialing out the exogenous regressors, on the grounds that the coefficient on the endogenous regressor is the only coefficient of interest and that the numerical optimization required by CUE is most reliable when only one parameter is being estimated.

in quantitative terms, noncriminal gun prevalence may prevent more homicides of criminals than of noncriminals. Note that this does not necessarily mean noncriminal gun prevalence would reduce the aggregate gun homicide rate. Total gun homicide would decrease if noncriminal gun prevalence made criminals more cautious or less willing to commit crimes, but not if it displaced criminal activity from areas of high gun prevalence to areas of lower prevalence.

We do not have separate county-level data on the numbers of homicides of criminals and noncriminals. However, we can indirectly test whether noncriminal gun prevalence is having a significant impact on homicides of criminals by looking separately at gun homicide in rural and urban areas. Gun homicide is found disproportionately in urban areas and is associated with drug crime, a primarily urban phenomenon. If noncriminal gun prevalence reduces significant numbers of homicides of criminals, it should be particularly apparent in primarily urban counties. This is, in fact, what we find.

We split our sample into two subsamples based on whether 50% or more of the county population reside in rural areas (using the U.S. Census Bureau definition). This gives us 771 largely-urban counties, with a total population of 195 million and an aggregate gun homicide rate of 6.9 per 100,000 persons, and 685 largely-rural counties, with a total population of 33 million and an aggregate gun homicide rate of 3.7 per 100,000 persons. We estimate the same equations as reported in Tables 2 and 4, but for the two subsamples separately.

The results for the largely-urban counties are – predictably, given that we are weighting by population and this group accounts for about 85% of the population in the total sample – very similar to the results we report for the total sample. The estimated coefficient on PSG/100 is -1.91 and significant at the 5% level; the first-stage F statistic is 28.0, suggesting the equation is not weakly identified; the J statistic is 2.59 with a p-value of 0.27, evidence that again the instruments are valid and are identifying a single LATE parameter; and estimation with the three instruments separately generate three negative estimates of $\hat{b}_1$. The results for the largely-rural counties, however, are essentially null. They show no significant impact of gun prevalence on homicide, whether estimating with all three instruments or with the instruments separately. The J statistic again provides evidence that the instruments are valid (J=1.69, p-value=0.43). However, these results should be treated with caution because the equation shows

evidence of weak identification, particularly when estimating with all three instruments or with OMAG or LVETS.

The impact of gun prevalence on nongun 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". For consistency with the gun homicide results, we continue to use PSG as our proxy for general gun prevalence.

The first-stage results are essentially identical to those reported in Table 3 and we do not report them here.[21] Panel B of Table 2 presents the results for the GMM estimations when the dependent variable is the log nongun homicide rate. The patterns in the coefficients on the covariates are similar to those for the gun homicide equations; the main exception is that rurality is now significantly associated with lower nongun homicide rates. 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 coefficients become negative and significantly different from zero at the 10% level, suggesting a negative deterrent impact of gun levels on nongun homicide. The results are weaker, however, than they are in the gun homicide case. The standard errors are larger than the gun homicide estimations (about 3.5 vs. 0.8) and the point estimates, though large, bring with them large confidence intervals. The test of the endogeneity of PSG suggests it can be treated as exogenous, which implies a null impact on nongun homicide. Reestimating using log PSG as the gun level proxy provides a robustness check and reinforces the conclusion that a negative deterrent effect is either weak or nonexistent: in these estimations (not reported here), PSG is insignificantly different from zero, whether treated as exogenous or endogenous.

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

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.

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.

6. 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 applying Granger-type causality tests to 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. More importantly, however, we 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. It is possible that the beneficiaries of the reduced level of violence may include substantial numbers of (urban) criminals who benefit via spillovers from noncriminal gun prevalence, but we have no direct evidence that would enable us to estimate the magnitude of these spillovers.

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 favour 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.

Table 1: Descriptive Statistics (N=1456 except where noted)		**Mean**	**Std. Dev.**	**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=1436)	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=1417)	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=1455)	56.55	9.90	14.83	81.40
Log PCTREP88	PCTREP88 logged (N=1455)	4.02	0.19	2.70	4.40
OMAG	Subscriptions to 3 top outdoor/sport magazines per 100,000 pop. (N=1450)	2,259	928	220	6,296
Log OMAG	OMAG logged (N=1450)	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=1450)	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	2064	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	58.48	7.02	33.40	80.90
Log PCTFEM18	PCTFEM18 logged	4.06	0.12	3.51	4.39
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
PCTINCLT15K	% households with income < $15,000	27.91	8.86	5.00	65.20
Log PCTINCLT15K	PCTINCLT15K logged	3.27	0.36	1.61	4.18
INEQUALITY	% HHs w/income <$15k / % income >$75k	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

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

	A. Gun homicide		B. Nongun homicide	
	(1)	(2)	(3)	(4)
	PSG-exogenous	PSG-endogenous	PSG-exogenous	PSG-endogenous
PSG/100	0.968***	-2.078***	-0.441	-6.405*
	(0.217)	(0.791)	(0.774)	(3.504)
Log DENSITY	0.266***	0.200***	0.686***	0.549***
	(0.027)	(0.030)	(0.112)	(0.122)
PCTSUBURBAN	-0.000	-0.003*	0.013**	0.009
	(0.002)	(0.002)	(0.005)	(0.006)
PCTURBAN	0.000	-0.001	0.003	0.000
	(0.001)	(0.001)	(0.004)	(0.004)
PCT0T17	0.034***	0.050***	0.056*	0.106**
	(0.011)	(0.015)	(0.034)	(0.052)
PCT18T24	-0.008	0.012	-0.071**	-0.045
	(0.007)	(0.010)	(0.029)	(0.035)
PCT25T44	0.022***	0.028***	0.036*	0.045***
	(0.007)	(0.008)	(0.021)	(0.017)
PCT45T64	0.053***	0.086***	-0.050	-0.001
	(0.016)	(0.018)	(0.052)	(0.055)
Log PCTBLK	0.234***	0.262***	0.493***	0.457***
	(0.020)	(0.020)	(0.065)	(0.064)
Log PCTHISP	0.032	-0.021	-0.090	-0.155
	(0.032)	(0.036)	(0.125)	(0.146)
Log PCTFEM18	-0.296	-0.004	0.718	1.173*
	(0.206)	(0.213)	(0.698)	(0.697)
Log PCTEDUC	-0.625***	-0.755***	-0.872***	-1.086***
	(0.107)	(0.133)	(0.296)	(0.378)
Log PCTTRANS	0.045	0.050	-0.036	0.056
	(0.052)	(0.052)	(0.148)	(0.174)
Log MEDHHINC	0.111	0.146	1.192	0.727
	(0.341)	(0.404)	(1.148)	(1.184)
Log PCTINCLT15K	0.750*	0.657	2.701***	3.388***
	(0.414)	(0.405)	(0.910)	(0.955)
Log INEQUALITY	0.582***	0.488***	0.940**	1.187**
	(0.118)	(0.138)	(0.429)	(0.482)
Log PCTPOOR	0.903***	0.820***	1.356***	1.125***
	(0.223)	(0.177)	(0.474)	(0.406)
Log PCTUNEMP	0.076	0.085	0.780*	0.876**
	(0.107)	(0.117)	(0.440)	(0.443)
Log PCTVACANT	0.204***	0.193***	0.641***	0.627***
	(0.059)	(0.051)	(0.135)	(0.119)
J statistic	$\chi^2(3)=10.90$	$\chi^2(2)=2.31$	$\chi^2(3)=6.82$	$\chi^2(2)=3.72$
p-value	0.012	0.315	0.078	0.156
Test of exogeneity of PSG		$\chi^2(1)=7.72$		$\chi^2(1)=1.47$
p-value		0.005		0.226
within-R^2	0.785	n.a.	0.768	n.a.
N	1429	1429	1449	1449
Number of clusters/fixed effects	49	49	49	49

Notes: * 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.

Table 3: Tests of Instrument Relevance (first-stage regression results)
Dependent variable: PSG/100
Estimation method: OLS

	(1)	(2)	(3)	(4)	(5)	(6)
Log PCTREP88	0.109***	0.129***			0.105***	
	(0.026)	(0.024)			(0.024)	
Log OMAG	0.050***		0.069***		0.037***	
	(0.010)		(0.011)		(0.012)	
Log VETS	0.036			0.104***	0.036	
	(0.027)			(0.026)	(0.028)	
Log G&A					0.027	0.073***
					(0.018)	(0.020)
F statistic	22.9	28.0	37.3	16.1	16.3	13.8
Number of observations	1429	1435	1430	1436	1429	1430
Number of clusters/fixed effects	49	49	49	49	49	49
Stock-Yogo (2005) critical values for F						
10% maximal IV size	22.3		16.4		24.6	16.4
15% maximal IV size	12.8		9.0		14.0	9.0
20% maximal IV size	9.5		6.7		10.3	6.7

Notes: *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 2; 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.

Table 4: 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	–2.08***	–0.727	–3.25***	–2.59†	–1.44*	1.90*
Standard error	(0.791)	(1.19)	(1.16)	(1.59)	(0.783)	(0.977)
95% confidence interval	[–6.25, –1.14]	[–3.06, 1.61]	[–5.52, –0.98]	[–5.70, 0.52]	[–2.89, 0.01]	[–0.01, 3.82]
J statistic	$\chi^2(2)$=2.31 p-value=0.315	Just-identified	Just-identified	Just-identified	$\chi^2(2)$=11.4 p-value=0.010	Just-identified
Test of exogeneity of G&A					$\chi^2(1)$=8.82 p-value=0.003	
Number of observations	1429	1435	1430	1436	1429	1430
Number of clusters/fixed effects	49	49	49	49	49	49

Notes: † 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 2; for brevity only the coefficient on PSG is reported. All test statistics are robust to heteroskedasticity and clustering on state.

Appendix 1: Data and sources

We use cross-sectional data for U.S. 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 U.S. Bureau of the Census, *County and City Data Book, 1994*. Other data sources were as follows:

Homicide rates are averages for the seven years 1987 to 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 (U.S. 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 (U.S. 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 U.S. Census Bureau (2000).

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, and can be freely downloaded via the software database of RePEc.[22] For further discussion of how the estimators and tests are implemented, see Baum, Schaffer and Stillman (2003; 2007; 2008), Schaffer (2007), and the references therein.

[22] 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.

Appendix 2: The OLS and IV estimators with population heterogeneity

Model setup

The "true model" is one with population heterogeneity (equation 5c in the main text):

$$h_i = \beta_0 + \beta_1^{C} g_i^{C} + \beta_1^{NC} g_i^{NC} + u_i \quad \text{(A2.1)}$$

Criminal and noncriminal gun prevalence are not separately observable. A proxy for aggregate gun prevalence is available (equation 11 in the text):

$$p_i = \delta_0 + \delta_1^{C} g_i^{C} + \delta_1^{NC} g_i^{NC} + v_i \quad \text{(A2.2)}$$

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 (equations 8 and 9 in the text):

$$g_i^{C} = \pi_0^{C} + \pi_1^{C} Z_i + \eta_i^{C} \quad \text{(A2.3)}$$

$$g_i^{NC} = \pi_0^{NC} + \pi_1^{NC} Z_i + \eta_i^{NC} \quad \text{(A2.4)}$$

We assume that $\pi_1^{C}, \pi_1^{NC} \geq 0$ 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 (equation 1 in the text):

$$h_i = \beta_0 + \beta_1 g_i + u_i \quad \text{(A2.5)}$$

If only the proxy for gun prevalence is observable, the estimating equation is (equation 2 in the text):

$$h_i = b_0 + b_1 p_i + e_i \quad \text{(A2.6)}$$

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 equation (A2.1) as a "random coefficient" model:

$$h_i = \beta_0 + \beta_{1i} g_i + u_i \quad \text{(A2.7)}$$

where

$$\beta_{1i} \equiv \frac{g_i^{C}}{g_i^{C} + g_i^{NC}} \beta_1^{C} + \frac{g_i^{NC}}{g_i^{C} + g_i^{NC}} \beta_1^{NC} \quad \text{(A2.8)}$$

and by definition $g_i \equiv g_i^{C} + g_i^{NC}$, i.e., our measures of gun prevalence are in levels (rather than, say, logs). The average treatment effect of gun prevalence β^{ATE} is $E(\beta_{1i})$. This does not have a simple expression because the quantities in (A2.8) are ratios, but the probability limit does take a simple form:

$$\text{plim}(\beta_{1i}) = \frac{\mu^{C}}{\mu^{C}+\mu^{NC}}\beta_1^{C} + \frac{\mu^{NC}}{\mu^{C}+\mu^{NC}}\beta_1^{NC} \tag{A2.9}$$

where $\mu^{C} \equiv E(g_i^{C})$ and $\mu^{NC} \equiv E(g_i^{NC})$. Equation (A2.9) is equation (6) in the main text, and shows that, for a large sample of localities, the average treatment effect for a randomly distributed increase in gun prevalence is a weighted average of the criminal and noncriminal effects, with weights given by average criminal and noncriminal gun prevalence.

OLS estimation

We consider first estimation of equation (A2.5), when total gun prevalence is directly observable. The OLS estimator is

$$\hat{\beta}_1^{OLS} = \frac{s_{gh}}{s_g^2} \xrightarrow{P} \frac{\text{cov}(g_i, h_i)}{\text{var}(g_i)} \tag{A2.10}$$

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

$$\begin{aligned}
\text{cov}(g_i, h_i) &= \text{cov}\left\{\left(g_i^{C} + g_i^{NC}\right), \left(\beta_0 + \beta_1^{C} g_i^{C} + \beta_1^{NC} g_i^{NC} + u_i\right)\right\} \\
&= \beta_1^{C}\text{var}(g_i^{C}) + \beta_1^{C}\text{cov}(g_i^{C}, g_i^{NC}) + \beta_1^{C}\text{var}(g_i^{C}) + \beta_1^{C}\text{cov}(g_i^{C}, g_i^{NC}) \\
&\quad + \text{cov}(g_i^{C}, u_i) + \text{cov}(g_i^{NC}, u_i) \\
&= \beta_1^{C}\left\{\text{var}(g_i^{C}) + \text{cov}(g_i^{C}, g_i^{NC})\right\} + \beta_1^{NC}\left\{\text{var}(g_i^{NC}) + \text{cov}(g_i^{C}, g_i^{NC})\right\} \\
&\quad + \text{cov}(g_i^{C}, u_i) + \text{cov}(g_i^{NC}, u_i)
\end{aligned} \tag{A2.11}$$

The denominator is simply

$$\begin{aligned}
\text{var}(g_i) &= \text{var}\left(g_i^{C} + g_i^{NC}\right) \\
&= \text{var}(g_i^{C}) + \text{var}(g_i^{NC}) + 2\,\text{cov}(g_i^{C}, g_i^{NC})
\end{aligned} \tag{A2.12}$$

and therefore

$$\begin{aligned}
\hat{\beta}_1^{OLS} = \frac{s_{gh}}{s_g^2} \xrightarrow{P} & \frac{\left\{\text{var}(g_i^{C}) + \text{cov}(g_i^{C}, g_i^{NC})\right\}}{\text{var}(g_i^{C}) + \text{var}(g_i^{NC}) + 2\,\text{cov}(g_i^{C}, g_i^{NC})}\beta_1^{C} \\
& + \frac{\left\{\text{var}(g_i^{NC}) + \text{cov}(g_i^{C}, g_i^{NC})\right\}}{\text{var}(g_i^{C}) + \text{var}(g_i^{NC}) + 2\,\text{cov}(g_i^{C}, g_i^{NC})}\beta_1^{NC} \\
& + \frac{\text{cov}(g_i^{C}, u_i) + \text{cov}(g_i^{NC}, u_i)}{\text{var}(g_i^{C}) + \text{var}(g_i^{NC}) + 2\,\text{cov}(g_i^{C}, g_i^{NC})}
\end{aligned} \tag{A2.13}$$

The OLS estimator differs from the ATE (equation A2.9) for two reasons. First, if criminal or noncriminal guns are endogenous, then the third term in (A2.13) is nonzero. Second, even if gun prevalence is exogenous and the third term in (A2.13) 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 (A2.9); whereas the ATE weights are relative gun prevalence, 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 (A2.5) 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 by variation in criminal gun prevalence.

Next we consider OLS estimation of (A2.6), 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

$$\hat{\mathrm{b}}_1^{\mathrm{OLS}} = \frac{\mathrm{s}_{\mathrm{ph}}}{\mathrm{s}_{\mathrm{p}}^2} \xrightarrow{\mathrm{P}} \frac{\mathrm{cov}(p_i, h_i)}{\mathrm{var}(p_i)} \tag{A2.14}$$

The numerator is

$$\begin{aligned} \mathrm{cov}(p_i, h_i) &= \mathrm{cov}\left\{\left(\delta_0 + \delta_1^{\mathrm{C}} g_{\mathrm{i}}^{\mathrm{C}} + \delta_1^{\mathrm{NC}} g_{\mathrm{i}}^{\mathrm{NC}} + \nu_i\right), \left(\beta_0 + \beta_1^C g_i^C + \beta_1^{NC} g_i^{NC} + u_i\right)\right\} \\ &= \delta_1^{\mathrm{C}}\beta_1^C\ \mathrm{var}(g_{\mathrm{i}}^{\mathrm{C}}) + \delta_1^{\mathrm{NC}}\beta_1^{NC}\ \mathrm{var}(g_{\mathrm{i}}^{\mathrm{NC}}) + (\delta_1^{\mathrm{C}}\beta_1^{NC} + \delta_1^{\mathrm{NC}}\beta_1^C)\,\mathrm{cov}(g_{\mathrm{i}}^{\mathrm{C}}, g_{\mathrm{i}}^{\mathrm{NC}}) \end{aligned} \tag{A2.15}$$

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

$$\begin{aligned} \mathrm{var}(p_i) &= \mathrm{var}\left(\delta_0 + \delta_1^{\mathrm{C}} g_{\mathrm{i}}^{\mathrm{C}} + \delta_1^{\mathrm{NC}} g_{\mathrm{i}}^{\mathrm{NC}} + \nu_i\right) \\ &= (\delta_1^{\mathrm{C}})^2\,\mathrm{var}(g_i^C) + (\delta_1^{\mathrm{NC}})^2\ \mathrm{var}(g_i^{NC}) + 2\delta_1^{\mathrm{C}}\delta_1^{\mathrm{NC}}\,\mathrm{cov}(g_i^C, g_i^{NC}) + \mathrm{var}(\nu_i) \end{aligned} \tag{A2.16}$$

Thus

$$\begin{aligned} \hat{\mathrm{b}}_1^{\mathrm{OLS}} \xrightarrow{\mathrm{P}} & \frac{\delta_1^{\mathrm{C}}\,\mathrm{var}(g_{\mathrm{i}}^{\mathrm{C}}) + \delta_1^{\mathrm{NC}}\,\mathrm{cov}(g_{\mathrm{i}}^{\mathrm{C}}, g_{\mathrm{i}}^{\mathrm{NC}})}{(\delta_1^{\mathrm{C}})^2\,\mathrm{var}(g_i^C) + (\delta_1^{\mathrm{NC}})^2\ \mathrm{var}(g_i^{NC}) + 2\delta_1^{\mathrm{C}}\delta_1^{\mathrm{NC}}\,\mathrm{cov}(g_i^C, g_i^{NC}) + \mathrm{var}(\nu_i)}\beta_1^C \\ & + \frac{\delta_1^{\mathrm{NC}}\,\mathrm{var}(g_{\mathrm{i}}^{\mathrm{NC}}) + \delta_1^{\mathrm{C}}\,\mathrm{cov}(g_{\mathrm{i}}^{\mathrm{C}}, g_{\mathrm{i}}^{\mathrm{NC}})}{(\delta_1^{\mathrm{C}})^2\,\mathrm{var}(g_i^C) + (\delta_1^{\mathrm{NC}})^2\ \mathrm{var}(g_i^{NC}) + 2\delta_1^{\mathrm{C}}\delta_1^{\mathrm{NC}}\,\mathrm{cov}(g_i^C, g_i^{NC}) + \mathrm{var}(\nu_i)}\beta_1^{NC} \end{aligned} \tag{A2.17}$$

Equation (A2.17) 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 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 $\delta_1^{\mathrm{NC}} >> \delta_1^{\mathrm{C}}$, then the OLS estimator will put a high weight on the noncriminal impact gun prevalence, and vice-versa if $\delta_1^{\mathrm{NC}} << \delta_1^{\mathrm{C}}$. Note that even if δ_1^{C}=0, the weight on β_1^{C} may be positive if criminal and noncriminal gun prevalence are correlated. Note also that sign{$\hat{b}_1^{\mathrm{OLS}}$} is not in general a consistent estimator of sign{$\hat{\beta}_1^{\mathrm{OLS}}$}.

IV estimation

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

$$\hat{\beta}_1^{\mathrm{IV}} = \frac{s_{Zh}}{s_{Zg}} \xrightarrow{p} \frac{\mathrm{cov}(Z_i, h_i)}{\mathrm{cov}(Z_i, g_i)} \tag{A2.18}$$

Taking the numerator first,

$$\begin{aligned} \mathrm{cov}(Z_i, h_i) &= \mathrm{cov}\left\{Z_i, \left(\beta_0 + \beta_1^C g_i^C + \beta_1^{NC} g_i^{NC} + u_i\right)\right\} \\ &= \mathrm{cov}(Z_i, \beta_1^C g_i^C) + \mathrm{cov}(Z_i, \beta_1^{NC} g_i^{NC}) + \mathrm{cov}(Z_i, u_i) \\ &= \beta_1^C \mathrm{cov}(Z_i, g_i^C) + \beta_1^{NC} \mathrm{cov}(Z_i, g_i^{NC}) \end{aligned} \tag{A2.19}$$

since *Z* is exogenous and orthogonal to the error term. Using equations (A2.3) and (A2.4), we have

$$\begin{aligned} \mathrm{cov}(Z_i, g_i^C) &= \mathrm{cov}(Z_i, \pi_0^{\mathrm{C}} + \pi_1^{\mathrm{C}} Z_i + \eta_i) = \pi_1^{\mathrm{C}} \mathrm{var}(Z_i) \\ \mathrm{cov}(Z_i, g_i^{NC}) &= \mathrm{cov}(Z_i, \pi_0^{\mathrm{NC}} + \pi_1^{\mathrm{NC}} Z_i + \eta_i) = \pi_1^{\mathrm{NC}} \mathrm{var}(Z_i) \end{aligned} \tag{A2.20}$$

since *Z* is also uncorrelated with η. Substituting (A2.20) into (A2.19), we have

$$\mathrm{cov}(Z_i, h_i) = \beta_1^C \pi_1^{\mathrm{C}} \mathrm{var}(Z_i) + \beta_1^{NC} \pi_1^{\mathrm{NC}} \mathrm{var}(Z_i) = \mathrm{var}(Z_i)\left\{\beta_1^C \pi_1^{\mathrm{C}} + \beta_1^{NC} \pi_1^{\mathrm{NC}}\right\} \tag{A2.21}$$

Now taking the denominator of (A2.18),

$$\begin{aligned} \mathrm{cov}(Z_i, g_i) &= \mathrm{cov}\left\{Z_i, \left(g_i^C + g_i^{NC}\right)\right\} = \mathrm{cov}(Z_i, g_i^C) + \mathrm{cov}(Z_i, g_i^{NC}) \\ &= \pi_1^{\mathrm{C}} \mathrm{var}(Z_i) + \pi_1^{\mathrm{NC}} \mathrm{var}(Z_i) = \mathrm{var}(Z_i)\left\{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}\right\} \end{aligned} \tag{A2.22}$$

where we have made use of (A2.20). Substituting (A2.21) and (A2.22) into (A2.18), we obtain

$$\hat{\beta}_1^{\mathrm{IV}} = \frac{s_{Zh}}{s_{Zg}} \xrightarrow{p} \frac{\mathrm{var}(Z_i)\left\{\beta_1^C \pi_1^{\mathrm{C}} + \beta_1^{NC} \pi_1^{\mathrm{NC}}\right\}}{\mathrm{var}(Z_i)\left\{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}\right\}} = \frac{\pi_1^{\mathrm{C}}}{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}} \beta_1^{\mathrm{C}} + \frac{\pi_1^{\mathrm{NC}}}{\pi_1^{\mathrm{C}} + \pi_1^{\mathrm{NC}}} \beta_1^{\mathrm{NC}} \tag{A2.23}$$

which is equation (10) 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}, where 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. LATE (A2.23) will be a consistent estimator of the ATE (A2.9) only in the special cases (a) the correlations of *Z* with criminal/noncriminal gun prevalence are proportional to the levels of criminal/noncriminal gun prevalence, and (b) the impacts of criminal and noncriminal gun prevalence are equal, $\beta_1^C = \beta_1^{NC}$.

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

$$\hat{b}_1^{IV} = \frac{s_{Zh}}{s_{Zp}} \xrightarrow{P} \frac{\text{cov}(Z_i, h_i)}{\text{cov}(Z_i, p_i)} \tag{A2.24}$$

The numerator is the same as in (A.5) above. The denominator is

$$\begin{aligned} \text{cov}(Z_i, p_i) &= \text{cov}\left\{Z_i, \left(\delta_0 + \delta_1^C g_i^C + \delta_1^{NC} g_i^{NC} + \nu_i\right)\right\} \\ &= \delta_1^C \text{cov}(Z_i, g_i^C) + \delta_1^{NC} \text{cov}(Z_i, g_i^{NC}) \\ &= \delta_1^C \pi_1^C \text{ var}(Z_i) + \delta_1^{NC} \pi_1^{NC} \text{ var}(Z_i) = \text{var}(Z_i)\left\{\delta_1^C \pi_1^C + \delta_1^{NC} \pi_1^{NC}\right\} \end{aligned} \tag{A2.25}$$

Substituting (A2.21) and (A2.25) into (A2.24) yields

$$\begin{aligned} \hat{b}_1^{IV} = \frac{s_{Zh}}{s_{Zp}} \xrightarrow{P} \frac{\text{cov}(Z_i, h_i)}{\text{cov}(Z_i, p_i)} &= \frac{\text{var}(Z_i)\left\{\beta_1^C \pi_1^C + \beta_1^{NC} \pi_1^{NC}\right\}}{\text{var}(Z_i)\left\{\delta_1^C \pi_1^C + \delta_1^{NC} \pi_1^{NC}\right\}} \\ &= \frac{\pi_1^C}{\delta_1^C \pi_1^C + \delta_1^{NC} \pi_1^{NC}} \beta_1^C + \frac{\pi_1^{NC}}{\delta^C \pi_1^C + \delta_1^{NC} \pi_1^{NC}} \beta_1^{NC} \\ &= \frac{\pi_1^C + \pi_1^{NC}}{\delta_1^C \pi_1^C + \delta_1^{NC} \pi_1^{NC}} \left\{ \frac{\pi_1^C}{\pi_1^C + \pi_1^{NC}} \beta_1^C + \frac{\pi_1^{NC}}{\pi_1^C + \pi_1^{NC}} \beta_1^{NC} \right\} \end{aligned} \tag{A2.26}$$

which is equation (11) in the text, the expression for the LATE estimator when gun prevalence is proxied by *p*. Note that this a scaling parameter (assumed positive) times the probability limit of $\hat{\beta}_1^{IV}$ given in equation (A2.23). Thus sign{ $\hat{b}_1^{IV}$ } is a consistent estimator of sign{ $\hat{\beta}_1^{IV}$ }, irrespective of the strength of the correlation between the proxy and criminal/noncriminal gun prevalence.

REFERENCES

Angrist, Joshua and Guido Imbens. 1995. "Two-stage least squares estimates of average causal effects in models with variable treatment intensity." *Journal of the American Statistical Association* 90(430):431-442.

Annest, Joseph L., James A. Mercy, Delinda R. Gibson, and George W. Ryan. 1995. "National estimates of nonfatal firearm-related injuries." *Journal of the American Medical Association* 273:1749-54.

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

Ayres, Ian and John J. Donohue III. 2003. "Shooting Down the 'More Guns, Less Crime' Hypothesis." *Stanford Law Review*, 55:1193-1312.

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

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

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2007. "Enhanced routines for instrumental variables/generalized method of moments estimation and testing." *The Stata Journal* 7:465-506.

Baum, Christopher, Mark E. Schaffer and Steven Stillman. 2008. "IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation". http://ideas.repec.org/c/boc/bocode/s425401.html.

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

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

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

Cook, Philip J., and Jens Ludwig (eds.). 2003a. *Evaluating Gun Policy*, edited by Jens Ludwig and Philip J. Cook. Washington, D.C.: Brookings Institution Press.

Cook, Philip J., and Jens Ludwig. 2003b. "Pragmatic gun policy." In Cook and Ludwig (2003a), *op. cit*.

Cook, Philip J., and Jens Ludwig. 2003c. "Guns and burglary." In Cook and Ludwig (2003a), *op. cit*.

Cook, Philip J., and Jens Ludwig. 2004. "The social costs of gun ownership." NBER Working Paper 10736. http://www.nber.org/papers/w10736.

Cook, Philip J., and Jens Ludwig. 2006. "The social costs of gun ownership." *Journal of Public Economics* 90:379-391.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington, DC: World Bank/Johns Hopkins University Press.

Decker, Scott H., G. David Curry, Shannan Catalano, Adam Watkins. 2005. "Strategic Approaches to Community Safety Initiative (SACSI) in St. Louis. Final Report." Research report submitted to the U.S. Department of Justice, Document Number 210361, Award Number 2000-IJ-CX-K008, June. http://www.ncjrs.gov/pdffiles1/nij/grants/210361.pdf.

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

Duggan, Mark. 2003. "Guns and Suicide." In Cook and Ludwig (2003a), *op. cit.*

Griliches, Zvi and Jerry Hausman. 1986. "Errors in variables in panel data." *Journal of Econometrics* 31(1): 93-118.

Hansen, L.P. J. Heaton and A. Yaron. 1996. "Finite Sample Properties of Some Alternative GMM Estimators." *Journal of Business and Economic Statistics*, 14(3): 262-280.

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

Heckman, James. 1997. "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *Journal of Human Resources*, 32(3): 441-462.

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

Imbens, Guido and Joshua Angrist. 1994. "Identification and estimation of local average treatment effects." *Econometrica* 62(2): 467-475.

Kates, Don B. and Gary Mauser. 2007. "Would Banning Firearms Reduce Murder and Suicide? A Review of International and Some Domestic Evidence." *Harvard Journal of Law and Public Policy* 30(2): 649-694.

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

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

Kleck, Gary. 1988. "Crime control through the private use of armed force." *Social Problems* 35:1-21.

Kleck, Gary. 1997. *Targeting guns: Firearms and their control*. N.Y.: Aldine.

Kleck, Gary and Michael Hogan. 1999. "A National Case-control Study of Homicide Offending and Gun Ownership." Social Problems 46(2):275-293.

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

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

Kleck, Gary, and Don B. Kates. 2001. *Armed: New perspectives on gun control*. Amherst, NY: Prometheus.

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

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

Kovandzic, Tomislav, Lynne M. Vieraitis, and Mark R. Yeisley. 1998. “The structural covariates of urban homicide.” *Criminology* 36:569-600.

Kovandzic, Tomislav, Mark E. Schaffer and Gary Kleck. 2005. “Gun Prevalence, Homicide Rates and Causality: A GMM Approach to Endogeneity Bias.” CEPR Discussion Paper 5357. http://ideas.repec.org/p/cpr/ceprdp/5357.html.

Land, Kenneth, Patricia L. McCall, and Lawrence E. Cohen. 1990. “Structural covariates of homicide rates.” *American Journal of Sociology* 95:922-963.

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

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

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

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

Newey, Whitney K. 1985. “Generalized method of moments specification testing.” *Journal of Econometrics* 29:229-256.

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

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

Sampson, Robert J. 1986. “Crime in cities.” Pp. 271-311in Albert J. Reiss Jr. and Michael Tonry (eds.), *Communities and Crime*. Chicago: University of Chicago Press.

Schaffer, Mark E. 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.

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

Sloan, John Henry, Arthur L. Kellermann, Donald T. Reay, James A. Ferris, Thomas Koepsell, Frederick P. Rivara, Charles Rice, Laurel Gray, and James LoGerfo. 1990. "Handgun regulations, crime, assaults and homicide." *New England Journal of Medicine* 319:1256-62.

Stock, James H. and Mark W. Watson. 2007. *Introduction to Econometrics*. 2nd edition. Pearson/Addison-Wesley.

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

Stock, J.H., and M. Yogo. 2005. "Testing for weak instruments in linear IV regression." In D.W.K. Andrews and J.H. Stock, eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge: Cambridge University Press, pp. 80-108. Working paper version: NBER Technical Working Paper 284, 2002. http://www.nber.org/papers/T0284.

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

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

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

U.S. Federal Bureau of Investigation (FBI). 1990-2000a, 2006. Crime in The United States 1989 [-1999] Uniform Crime Reports. Washington, D.C.: U.S. Government Printing Office.

U.S. National Center for Health Statistics. 1997. Limited access versions of Mortality Detail Files, 1987-1993, with location detail, supplied to third author. Hyattsville, MD: U.S. Department of Health and Human Services.

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

White, Halbert. 1994. *Estimation, inference and specification analysis*. Cambridge University Press.

Wooldridge, Jeffrey M. 2006. *Introductory econometrics: A modern approach*. Thomson/South-Western.

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

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