ECO 231W: Econometrics

Gun Prevalence and Crimes Rates

Zunran Guo

December 2014

Abstract

This paper examines the relationship between gun prevalence and crime rates based on the panel data collected from National Incident-Based Reporting System. The goal is to understand the relationship between guns and crime, particularly violent crime; and estimate the causal impact of gun prevalence on various types of criminal outcomes.

Gun Prevalence and Crimes Rates

Introduction

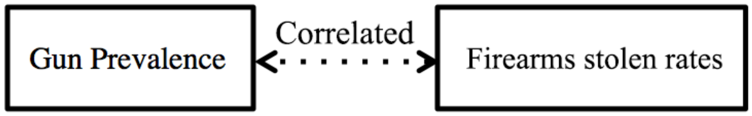
As of 2014, U.S. ranks No.1 in the world for gun ownership[1]. How do changes in gun prevalence influence the crime rates? Does gun prevalence lead to more crimes? If it does, does gun prevalence in general increase all kinds of crimes, from fondling to homicide? Or does prevalence in fact only increase firearm-involved crimes significantly? Is the reason simply that easy accessibility of firearms incites criminals to commit crimes? Or alternatively, does gun prevalence in general leads to few crimes? Is the reason simply that the criminals are deterred from committing crimes because the victim may possess a defense firearm? To answer these questions, we will face and try to overcome some limits mentioned in *More Guns, More Crime* by Duggan and the *Social cost of gun ownership* by Cook and Ludwig, such as the small number of observations in the time series data for an individual city, unknown level of firearm aggregation in different geographic areas.

Instead using gun magazine as a proxy for gun prevalence as Cook did to estimate crime rates, in this paper we assume gun prevalence and the number of firearm stolen are highly correlated with one another. The higher the number of firearm stolen in one region

[1] Chalabi, M. (2013, September 17). Gun crime statistics by US state.

means easier accessibility of guns, such as stealing, compared to purchasing from legal sources.

Figure 1: The relationship between Gun Prevalence



Models of Gun Prevalence and Crime Rates

Because firearms are used in a significant fraction of all violent crimes, but are also frequently used for self-defense purposes, number of stolen firearms often times implies further moves of the criminals, who are highly likely to use the stolen firearms to commit crimes; and may lead to more purchase of firearms for defense purposes. Therefore changes in the number of guns within an area could have an important impact on the level and average seriousness of criminal activity.

There are 11 different kinds of crimes tabulated in the panel data. However, not all of them have direct firearms involvement. Crimes such as fondling, robbery, burglary, theft, motor vehicle theft, and drugs, with no firearm involvement, therefore should not be simply added together and converted to certain rate, since if the reason that criminals steal firearms is for something, they are highly likely to use the stolen firearm to commit crimes, e.g. homicide, aggravated, kidnap etc. Otherwise why do they spend such great effort to obtain firearms?

A) Gun Prevalence and Firearm-Involved Crimes

Therefore when converting the crime rates, we should first consider crime rates with firearm involved, in other words, violent crimes, which we are interested in finding its causality. We obtain the firearms stolen rates and crime rates per 10,000 people by summing the occurrence firearms stolen and firearm-involved crimes respectively first, then dividing the sum respectively by the total population in a given jurisdiction, and multiplying by 10,000. We then convert all rates to natural log for the convenience of calculating the elasticity of different models.

$$\log_{\mathrm{crimerates}_{\mathrm{f,it}}} = \beta_0 + \beta_1 \log_{\mathrm{firearms,it}} + u_{it}$$

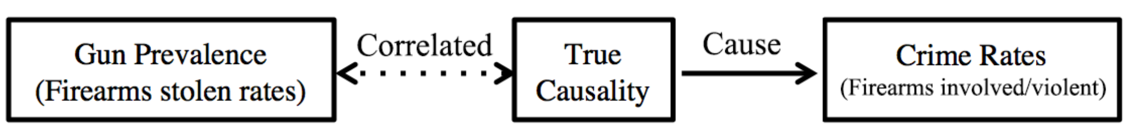
Weighted population caveat

Because cities in our panel data vary quite much on population size, it is necessary to assign different population weights in our regression models because we obtain the crime rates and firearm stolen rates by a population-size-determined conversion. In this paper, we will estimate the causality of gun prevalence on crime rates by taking population weight into account through all regressions.

B) Gun Prevalence, Firearm-Involved Crimes, and Other Crimes

As for concerns of selection and endogeneity, as the saying goes, “Guns don’t kill people. People kill people,” the above regression model, however, may suffer omitted variable bias since the stolen firearms do not kill people; people kill people for many reasons. Cities may have high crime rates due to many unobserved variables, such as the presence of drug-dealing gangs, aggravated mafia, or racial disparity. Gun prevalence (high firearms stolen rates) may look correlated with firearm-involved crime rates, but it may not be the causality

Figure 2: The Omitted Variable Bias (OVB) Problem



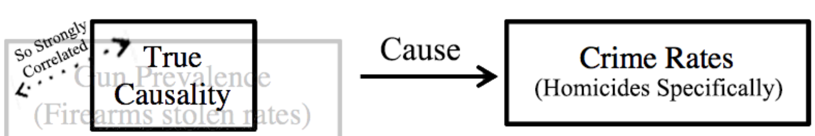
Therefore in this specification additional regressors can be added, such as the number of drug confiscation (converted to rates) or the number of aggravated assault in a given place or a given quarter, to control the possible upward omitted variable bias towards the effect of gun prevalence (firearms stolen rates), since gun prevalence may associated with only a certain degree of causal effect on crime rates. By adding more control variables we can better interpret the true causal effect of gun prevalence on crime rates. Similarly, regressor like mvtheft can also be added to determine the causal effect of gun prevalence on crime rates and whether the bias of the firearms stolen is offset upward, downward, or stay the same (bias and consistency check).

$$\log_{\text{crimerates}_{\text{f,it}}} = \beta_0 + \beta_1 \log_{\text{firearms,it}} + \beta_2 \log_{\text{drug,it}} + \cdots + u_{it}$$

C) Gun Prevalence and Homicide (Specific Crime)

A regression of only the gun prevalence (firearms stolen rates) and a specific type of crime (homicides rates) can actually make the omitted variables and simultaneity problem simpler, because in most cases homicides happened simply because of the high accessibility of firearms. If the criminal wants to kill someone so badly, compared to other aforementioned difficult ways, s/he will extremely likely to just shoot the victim with an available firearm. Therefore we can effectively evaluate the causal impact of gun prevalence on a specific type of crime (homicide) by a simple regression. We can also choose the add control of other reasonable regressors like drug confiscation to much more precisely evaluate the causal impact of gun prevalence on a specific type of crime (homicides), in a same way mentioned above.

Figure 3: Strong Causal Impact of Gun Prevalence on Homicides



$$\log_{\text{crimerates}_{\text{homicide,it}}} = \beta_0 + \beta_1 \log_{\text{firearms,it}} + \beta_2 \log_{\text{drug,it}} + \cdots + u_{it}$$

D) Gun Prevalence and Quarter Lagged Firearm-Involved Crimes

The lag effect of gun prevalence (firearms stolen rates) should also be taken into account for crime rates, by shifting the time frame in the regression by one or two quarters, since it may take time for the stolen firearms to be transferred for the use of criminal activities. As Cook and Lugwig suggested in *The social cost of gun ownership*, in this specification, it's also necessary control the possibility of reversal causation. The regression needs to condition on the log of the area's burglary and robbery rates, which are the kinds of crimes that seem likely to motivate the acquisition of a firearm for self-defense.

1. $\log_{\mathrm{crimerates_f},it} = \beta_0 + \beta_1 \log_{\mathrm{firearms,it-1}} + \beta_2 \log_{\mathrm{burglary,it}} + \beta_3 \log_{\mathrm{robbery,it}} + u_{it}$
2. $\log_{\mathrm{crimerates_f},it} = \beta_0 + \beta_1 \log_{\mathrm{firearms,it-2}} + \beta_2 \log_{\mathrm{burglary,it}} + \beta_3 \log_{\mathrm{robbery,it}} + u_{it}$

E) Gun Prevalence and Firearm-Involved Crimes with Various Fixed Effects

To help solve the aforementioned omitted variable problem, we should also apply fixed effects to our panel data when running regressions, so that some omitted variables that are invariant over time can be cleared way. In fixed effect specifications, α_i is the assumed time-invariant fixed effects for each city or each state. In time fixed effect specifications, α_t is assumed to have the same seasonal effect for all the cities or all the states each quarter, or each year. In fixed time and city or state effect specification, we will fixe both the characteristic traits for each city or state and the seasonal effects for quarter or year. Note that since α_i is unlikely strictly uncorrelated with all regressors in all time periods (strict exogeneity and orthogonality are likely to fail), we will not use any random effect model in this study.

1. $\log_{\mathrm{crimerates_f},it} = \beta_0 + \beta_1 \log_{\mathrm{firearms,it}} + \alpha_i + u_i$
2. $\log_{\mathrm{crimerates_f},it} = \beta_0 + \beta_1 \log_{\mathrm{firearms,it}} + \alpha_t + u_t$
3. $\log_{\mathrm{crimerates_f},it} = \beta_0 + \beta_1 \log_{\mathrm{firearms,it}} + \gamma D_i + \delta B_t + u_{it}$

Data Description and Model Estimation

Given the aforementioned models, we run regressions on Stata and compare the results with our prediction. In our first model, as table 1 shows, we find that gun prevalence (firearms stolen rates) looks highly correlated (over 82% in terms of elasticity) with crime rates, as expected. However when we add potential omitted variables, such as the number of drugs confiscation and the number of aggravated assault, the explanatory power of gun prevalence drops significantly. We also noticed that the number of aggravated assault picks a lot of causal impact on crime rates and R^2 of the regression increases significantly from the previous regression. This result suggests that our original regression suffers serious omitted variable bias and we need to propose a better model to better evaluate the causal impact of gun prevalence on crime rates.

On the other hand, as table 2 shows, when we examine the regression result for homicide, we find that gun prevalence (firearms stolen rates) is correlated with homicides to a reasonable degree, but not as high as expected. When we add other regressors to test the strength of causal impact of gun prevalence, we find that all other variables have almost no causal impact on homicides, a result that confirms with our suggestion.

To further examine the causal impact of gun prevalence (firearms stolen rates), we then apply various fixed effects on our regression, as table 3 and 4 show. One major finding is that the number of aggravated assault has very high correlation with crime rates. Since we deem aggravated assault as a control variable that picks up potential causal from correlated omitted variables, this high correlation implies a serious omitted variable bias problem in our regression model. Across all other regressions, we also find that drug has no strong explanatory power on crime rates. Another major finding from table 4 is that when we control quarter and year fixed effect along with control variables of drug and aggravated assault, we find significant higher explanatory power from gun prevalence on crime rates; while it is not so obvious when we control only the city fixed effect. This implies that a seasonal factor has a great causal impact on crime rates than a unit factor, such as state, does. Given that the causal impact of aggravated assault is almost the same across the tables and that drug confiscation doesn't have much explanatory power under quarter and year fixed effect, we conclude that the causal impact of gun prevalence on crime rates has elasticity around 32% in table 3 and 4.

When we apply the quarter lagged firearm-involved crimes model in table 5, we find that when we control the possibility of reversal causation by conditioning on the log of the area's burglary and robbery rates, which are the kinds of crimes that seem likely to motivate the acquisition of a firearm for self-defense, we don't find any significant change on the causal impact of gun prevalence (firearms stolen rates) on crime rates, the difference in coefficient is less than 0.01, which implies 1% increase in our measure of gun prevalence will cause less than 1% increase in crime rates. This result suggests that the quarter lagged effect model with the regressors we condition on does not change much on the causal impact of the gun prevalence on crime rates.

Conclusion and Other Limitations

Just as the saying goes, "Guns don't kill people. People kill people." Separating the true causal effect is very hard given the limited observations for each city and unknown level of firearm aggregation in different geographic areas. Because the desire to kill someone (with a gun) and incitation to kill someone when seeing a gun can be highly entangled. A criminal may desire to kill someone so badly at the beginning, but choose to give up or retaliate in other ways in the end, due to the difficulty to commit the crime without firearms. However, once the criminal notices that a firearm is potentially available by stealing, s/he may suddenly have a stronger desire to kill someone or commit crimes. In this case, it’s extremely difficult to tell whether the criminal commits a crime because of his/her intense desire, or because of the incitation of the accessible firearms. The accessibility of firearm and the desire to commit crimes act like chemicals that react intensely when mixing together. Both of them could be highly correlated to the causal impact to committing crime.

In this paper, we started with a few models and theories, trying to evaluate the casual impact of gun prevalence (firearms stolen rates) on crime rates. However, not all of our theories and assumptions hold when we obtain the regression results from Stata. The drug variable doesn’t explain much at all. Our systematic approach (20 specifications) is not exhaustive enough yet, as we haven’t tried all combinations of fixed effects and all combination of possible regressors given the page limit allowed.

Based on the regression results, we do find certain degrees of correlation of gun prevalence and crime rates (around 30%). However we can not conclusively prove that gun prevalence is the causality of crime rates, since there is no data of firearms available for us at this point to verify our fundamental assumption on the correlation between the number of firearms stolen and true gun prevalence in one region.

Further studies should be conducted on improving the number of observations on individual cities, the measurement of relevant variables, and the accuracy of the proxy on gun prevalence to better evaluate the causal impact of guns on crimes.

Tables and Graphs

Table 1: Regression of firearm-involved crime rates and firearm stolen rates

	Log (Number of Crimes Per 10,000 Residents)	
Regressor	(1)	(2)
Log (Total Number of Firearms Reported Stolen)	0.8208*** (0.0143)	0.3156 *** (0.0107)
Log (Total Number of Drugs Confiscation)		0.1336* (0.0241)
Log (Total Number of Aggravated Assaults)		0.7631*** (0.0167)
Number of observations	13162	13162
Fixed effect	No	No
T-stats on firearm stolen rates	13162	29.57
F-stats	3316.15	1892.01
R^2	0.3804	0.6921

* Significant at the 10 % level. ** Significant at the 5 % level. *** Significant at the 1 % level. The above result is weighted by population size.

Table 2: Regression of homicides and firearm stolen rates

	Log (Number of Crimes Per 10,000 Residents)	
Regressor	(1)	(2)
Log (Total Number of Firearms Reported Stolen)	0.4753*** (0.0187)	0.3913*** (0.0184)
Log (Total Number of Drugs Confiscation)		-0.00026 (0.0000503)
Log (Total Number of Aggravated Assaults)		0.0006872 (0.000269)
Number of observations	6769	6769
Fixed effect	No	No
T-stats on firearm stolen rates	15.32	21.29
F-stats	645.06	668.92
R^2	0.2576	0.3796

* Significant at the 10 % level. ** Significant at the 5 % level. *** Significant at the 1 % level. The above result is weighted by population size.

Table 3: Regression of firearm-involved crime rates and firearm stolen rates cont'd

	Log (Number of Crimes Per 10,000 Residents)			
Regressor	(1)	(2)	(3)	(4)
Log (Total Number of Firearms Reported Stolen)	0.4953** (.0676)	0.1284*** (0.0106)	0.9101*** (0.0204)	0.2301*** (0.0096)
Log (Number of Drugs Confiscation)		0.0209 (.0107)		0.1765*** (0.0102)
Log (Total Number of Aggravated Assaults)		0.4590*** (0.0141)		0.8010*** (0.0107)
Number of observations	13162	12975	13162	12975
City fixed effect	**Yes**	**Yes**	No	No
State fixed effect	No	No	**Yes**	**Yes**
Quarter fixed effect	No	No	No	No
Year fixed effect	No	No	No	No
City and quarter fixed effect	No	No	No	No
City and year fixed effect	No	No	No	No
State and quarter fixed effect	No	No	No	No
State and year fixed effect	No	No	No	No
T-stats on firearm stolen rates	7.33	12.14	44.57	24.02
F-stats	53.68	424.83	1986.85	5393.98
R^2	0.8521	0.8480	0.4947	0.7095

* Significant at the 10 % level. ** Significant at the 5 % level. *** Significant at the 1 % level. The above result is weighed by population size.

Table 4: Regression of firearm-involved crime rates and firearm stolen rates cont'd

	Log (Number of Crimes Per 10,000 Residents)			
Regressor	1	2	3	4
Log (Total Number of Firearms Reported Stolen)	0.8205*** (0.0142)	0.3105*** (0.0107)	0.8425*** (0.0140)	0.3399*** (0.0109)
Log (Number of Drugs Confiscation)		0.1341* (0.0241)		0.0998* (0.0232)
Log (Total Number of Aggravated Assaults)		0.7691*** (0.0168)		0.7691*** (0.0163)
Number of observations	13162	12975	13162	12975
City fixed effect	No	No	No	No
State fixed effect	No	No	No	No
Quarter fixed effect	**Yes**	**Yes**	No	No
Year fixed effect	No	No	**Yes**	**Yes**
City and quarter fixed effect	No	No	No	No
City and year fixed effect	No	No	No	No
State and quarter fixed effect	No	No	No	No
State and year fixed effect	No	No	No	No
T-stats on firearm stolen rates	57.62	29.06	59.99	31.24
F-stats	3320.27	1878.87	3598.47	1929.60
R^2	0.3808	0.4021	0.6939	0.7033

* Significant at the 10 % level. ** Significant at the 5 % level. *** Significant at the 1 % level. The above result is weighted by population size.

Table 5: Regression of lagged firearm-involved crime rates and firearm stolen rates

	Lagged Log (Number of Crimes Per 10,000 Residents)	
Regressor	1	2
Log (Total Number of Firearms Reported Stolen)	0.1255** (0.01579)	0.1338** (0.0179)
Log (Number of Crimes Per 10,000 Residents)		-0.0104 (0.0218)
Log (Total Number of Burglary)	0.4500*** (0.0353)	0.4534*** (0.0352)
Log (Total Number of Robbery)	-0.3639*** (0.0152)	-0.3495*** (0.0217)
Number of observations	6406	5838
City fixed effect	No	No
Fixed effect	No	No
T-stats on firearm	6.79	7.46
F-stats	46.06	152.87
R^2	0.007	0.2082

* Significant at the 10 % level. ** Significant at the 5 % level. *** Significant at the 1 % level. The above result is weighted by population.

All the above results from table 1 to table 5 are based on our panel data of 14849 observations from 33 states, specifically 416 cities, collected between 1993 and 2010.

Acknowledgement

This paper couldn't be finished without the help from Professor Kinsler and all wonderful TA's, especially Alon Bergman, Umair Khalil, and Cy McGeady. It's been great having you all this semester.

References

Duggan, M. (2011). More Guns, More Crime. Journal of Political Economy, 109(5), 1086-1086.

Cook, P., & Ludwig, J. (2005). The Social Costs Of Gun Ownership. *Journal of Public Economics,* 379-391.

Chalabi, M. (2013, September 17). Gun crime statistics by US state. Retrieved December 12, 2014, from http://www.theguardian.com/news/datablog/2013/sep/17/gun-crime-statistics-by-us-state

Appendix

STATA Code

```
use "/Users/Albert_Guo/Documents/University of Rochester/Academic Curriculum/Fall 2014/ECO 231W/Term Paper/crime_firearms.dta"
describe
gsort city state year quart
set matsize 500

// encode ID's
egen city_id = group(ORI)
encode state, generate (state_id)

// generate regression variabels
generate l_crimerates_f = ln((homicide_f + kidnap_f + rape_f + robbery_f + aggassault_f)/totpop*10000)
generate l_firearms = ln(firearms/totpop*10000)
generate l_drug = ln(drug/totpop*10000)
generate l_mvtheft = ln(mvtheft/totpop*10000)
generate l_theft = ln(theft/totpop*10000)
generate l_burglary = ln(burglary/totpop*10000)
generate l_simassault = ln(simassault/totpop*10000)
generate l_aggassault = ln(aggassault/totpop*1000)
generate l_robbery = ln(robbery/totpop*10000)
generate l_fondling = ln(fondling/totpop*10000)
generate l_rape = ln(rape/totpop*10000)
generate l_kidnap = ln(kidnap/totpop*10000)
generate l_homicide = ln(homicide/totpop*10000)


reg l_crimerates_f l_firearms [aw=totpop], robust
est store rough_reg

reg l_crimerates_f l_firearms l_drug l_aggassault [aw=totpop], robust
est store rough_ovb_reg
n est table rough_reg rough_ovb_reg, b(%7.4f) se(%7.4f) stats(N r2) k(l_firearms l_drug l_aggassault)



areg l_crimerates_f l_firearms [aw=totpop], absorb(city_id) r
est store fixed_city
areg l_crimerates_f l_firearms l_drug l_aggassault, absorb(city_id) r
est store fixed_city_ovb_reg
```

```
areg l_crimerates_f l_firearms [aw=totpop], absorb(state_id) r
est store fixed_state
areg l_crimerates_f l_firearms l_drug l_aggassault, absorb(state_id) r
est store fixed_state_ovb_reg
n est table fixed_city fixed_state fixed_city_ovb_reg fixed_state_ovb_reg, b(%7.4f)
se(%7.4f) stats(N r2) k(l_firearms l_drug l_aggassault)


areg l_crimerates_f l_firearms [aw=totpop], absorb(quart) r
est store quart_fixed_reg
areg l_crimerates_f l_firearms l_drug l_aggassault [aw=totpop], absorb(quart) r
est store quart_fixed_ovb_reg

areg l_crimerates_f l_firearms [aw=totpop], absorb(year) r
est store year_fixed_reg
areg l_crimerates_f l_firearms l_drug l_aggassault [aw=totpop], absorb(year) r
est store year_fixed_ovb_reg
n est table quart_fixed_reg year_fixed_reg quart_fixed_ovb_reg year_fixed_ovb_reg,
b(%7.4f) se(%7.4f) stats(N r2) k(l_firearms l_drug l_aggassault)



xi i.quart
areg l_crimerates_f l_firearms _Iquart_2 - _Iquart_4 [aw=totpop], absorb(city_id) r
est store fixed_quart_city_reg
areg l_crimerates_f l_firearms _Iquart_2 - _Iquart_4 l_drug l_aggassault [aw=totpop],
absorb(city_id) r
est store fixed_quart_city_ovb_reg

xi i.year
areg l_crimerates_f l_firearms _Iyear_1994 - _Iyear_2010 [aw=totpop], absorb(city_id) r
est store fixed_year_city_reg
areg l_crimerates_f l_firearms _Iyear_1994 - _Iyear_2010 l_drug l_aggassault
[aw=totpop], absorb(city_id) r
est store fixed_year_city_ovb_reg

xi i.quart
areg l_crimerates_f l_firearms _Iquart_2 - _Iquart_4 [aw=totpop], absorb(state_id)
est store fixed_quart_state_reg
areg l_crimerates_f l_firearms _Iquart_2 - _Iquart_4 l_drug l_aggassault [aw=totpop],
absorb(state_id)
est store fixed_quart_state_ovb_reg

xi i.year
areg l_crimerates_f l_firearms _Iyear_1994 - _Iyear_2010 [aw=totpop], absorb(state_id)
r
```

```
est store fixed_year_state_reg
areg l_crimerates_f l_firearms _Iyear_1994 - _Iyear_2010 l_drug l_aggassault
[aw=totpop], absorb(state_id)
est store fixed_year_state_ovb_reg
n est table fixed_quart_city_reg fixed_quart_city_ovb_reg fixed_year_city_reg
fixed_year_city_ovb_reg fixed_quart_state_reg fixed_quart_state_ovb_reg
fixed_year_state_reg fixed_year_state_ovb_reg, b(%7.4f) se(%7.4f) stats(N r2)
k(l_firearms l_drug l_aggassault)


// generate lags
tsset city_id quarter
gen lag1_homicide_f = L.1.homicide_f
gen lag1_kidnap_f = L.1.kidnap_f
gen lag1_rape_f = L.1.rape_f
gen lag1_robbery_f = L.1.robbery_f
gen lag1_aggassault_f = L.1.aggassault_f
gen lag1_l_crimerates_f = ln((L.1.homicide_f + L.1.kidnap_f + L.1.rape_f +
L.1.robbery_f + L.1.aggassault_f)/totpop*10000)

reg lag1_l_crimerates_f l_firearms l_burglary l_robbery [aw=totpop], robust

reg lag1_l_crimerates_f l_firearms l_crimerates_f l_burglary l_robbery [aw=totpop],
robust
```