Advances in Economic Analysis & Policy

Using Placebo Laws to Test "More Guns, Less Crime"

Eric Helland*   Alexander Tabarrok†

*Claremont-McKenna College, eric.helland@claremontmckenna.edu

†George Mason University, Tabarrok@GMU.edu

Using Placebo Laws to Test "More Guns, Less Crime"*

Eric Helland and Alexander Tabarrok

Abstract

We reexamine Mustard and Lott's controversial study on the affect of "shall-issue" gun laws on crime using an empirical standard error function randomly generated from "placebo" laws. We find that the effect of shall-issue laws on crime is much less well-estimated than the Mustard and Lott (1997) and Lott (2000) results suggest. We also find, however, that the cross equation restrictions implied by the Lott-Mustard theory are supported. A boomlet has occurred in recent years in the use of quasi-natural experiments to answer important questions of public policy. The intuitive power of this approach, however, has sometimes diverted attention from the statistical assumptions that must be made, particularly regarding standard errors. Failing to take into account serial correlation and grouped data can dramatically reduce standard errors suggesting greater certainty in effects than is actually the case. We find that the placebo law technique (Bertrand, Duflo and Mullainathan 2002) is a useful addition to the econometrician's toolkit.

KEYWORDS: guns, crime, placebo laws

I. Introduction

Lott and Mustard's (1997) paper on "shall-issue" gun laws (expanded in Lott's book *More Guns, Less Crime* Second Edition 2000) is one of the most controversial policy analyses of recent times. Shall-issue laws essentially allow any citizen to obtain a permit for a concealed weapon, much like obtaining a driver's license. Lott and Mustard estimate that shall-issue laws dramatically decrease murder, rape, and other violent crimes. Unlike most policy analyses, Lott's work has been enormously influential. For example, he was cited by 18 state attorneys general in their letter (dated July 8, 2002) to Attorney General John Ashcroft in support of Ashcroft's decision to interpret the second amendment as protecting an individual right. Lott's work also has been cited in Congressional testimony and in testimony before a number of state legislatures. Therefore, testing Lott's findings for robustness is of great importance.

Lott and Mustard estimate the impact of shall-issue laws with the following specification (among others):

$$y_{it} = \alpha\, year_t + \beta x_{it} + \delta\, shall_{it} + \phi_i + \varepsilon_{it,}$$

(1)

where y_{it} is the natural log of the relevant crime category per 100,000 people in county i in year t, x_{it} is a vector of control variables, $shall_{it}$ is a dummy variable equal to one if county i was covered by a shall-issue law in period t, and $\varepsilon_{it,}$ is a robust (White) standard error weighted by county population.

There is considerable evidence that county-level crime data are subject to "shocks." These shocks are common to a group at a particular point in time and extend through time.[1] Such shocks would confound conventional standard error corrections, and, if they exist, could cause serious downward bias in the standard errors.

The problem of serial correlation is common to many differences in difference studies (DD) and is not unique to Lott and Mustard (Bertrand et al. 2002). Positive serial correlation in the error term will cause standard errors to be understated, and will do so to a greater extent when the independent variable is also serially correlated. Since the variable of interest in a DD study is typically a policy variable (represented by a dummy set to zero up to some point in time and set to one thereafter), the relevant independent variable in DD studies is very serially correlated. State laws, for example, are not typically reevaluated every year. Therefore, knowing that a law is in effect today leads us to believe that it will be in effect tomorrow. It is this dependence that needs to be taken into account when calculating standard errors if, as is common, the error term also exhibits serial correlation.

2. The Placebo Law Technique

It is possible to fix the serial correlation problem by making assumptions about the true error distribution and "correcting" the ordinary least squares (OLS) standard errors in light of these assumptions. An alternative and generally superior approach that also can correct a number of other problems is to create standard errors and critical values using an estimate of the true error distribution bootstrapped from the data.[2] Consider a typical DD study that estimates the effect of a state law using panel data from across the United States. For example,

---

[1] For example, Grogger and Willis (2000) look at the impact of crack cocaine on local crime rates. Mustard (2002) examines the importance of omitted variable bias in estimating models using county-level crime data and Maltz and Targonski (2002) examine how the FBI's data imputation method may create local shocks even when none actually exist.

[2] This technique is discussed extensively in Bertrand et al. An earlier less extensive use is discussed in Bound et al.

$$y_{it} = \alpha \, year_t + \beta x_{it} + \delta T_{it} + \phi_i + \varepsilon_{it} \qquad\qquad (2)$$

where y_{it} is the outcome variable in question for county (or state) i in year t, $year_t$ represents year-specific intercepts, x_{it} represents control variables, T_{it} is an indicator variable equal to 1 if the law in question applies to observation i in period t, ϕ_i represents county-specific intercepts, and ε_{it} is the standard error.

      To estimate the true error distribution, randomly choose a time and a state, and then create a variable which indicates that at that time and in that state a policy was begun (see the example below). That is, replace the actual law indicator T_{it} with P_{it} and estimate

$$y_{it} = \alpha \, year_t + \beta x_{it} + \delta P_{it} + \phi_i + \varepsilon_{it} \qquad\qquad (3)$$

Store the coefficients on the relevant variables. Repeat the procedure enough times to generate a probability distribution function for the relevant coefficients. The empirical PDF, called the placebo PDF, can then be used to identify the correct standard errors and critical values for the model estimated on the true data. Since the placebo PDF is generated from the actual data, it correctly reflects the presence of grouped data, serial correlation, non-normality, or other irregularities that would otherwise confound appropriate statistical inference.

      The placebo PDF is so-called because of its analogy to clinical trials for new pharmaceuticals. In a randomized clinical trial, the effects of a new drug are not tested against a theoretical null of zero, but against the effects of a placebo. The idea is similar here except our primary interest is not in comparing the size of a coefficient, but in properly estimating the uncertainty with which the coefficient is measured.

      We apply the "placebo law" technique to better estimate the uncertainty surrounding the more guns, less crime hypothesis. Much of the criticism of Lott and Mustard's work has focused on the point estimates of shall-issue gun laws on crime rates. Different sub-samples or aggregations of the data (e.g., Black and Nagin, 1998, Ayres and Donohue, 2002, Olson and Maltz, 2001), different time periods (Ayres and Donohue, 2002), or different functional forms (Dezhbakhsh and Rubin, 1998, Ayres and Donohue, 1999, Moody, 2001, Plassmann and Tideman, 2001, Duggan, 2001) have been tested, but relatively little attention has focused on the standard errors. We did not attempt to survey the earlier literature or reevaluate every aspect of Lott's data and statistical model. In particular, Ayres and Donohue (2003) have reestimated the model using many different specifications, functional forms, and so on, and have numerous references to the literature. In this paper, we focus attention on the uncertainty surrounding the estimated effects, and we primarily follow Lott's formulation in the spirit of examining his results under the *best-case* scenario.[3] Our goal is to explain, illustrate, and apply the placebo law technique to an issue of importance.

3. Results

The original Lott and Mustard paper used data from 1977–1992, which is where we also begin. Ten shall-issue gun laws were passed during this time frame, all post–1984. To generate our placebo laws, we created two urns: the first was filled with state names, and the second with the years 1985–1992. We drew a state name from urn 1 with each name equally likely to be drawn, and similarly we drew a year from urn 2. We then created a dummy variable set equal to zero in that state prior to the year drawn and

---

[3] Ayres and Donohue (2003), for example, argue that the model should include a state-specific time trend while Lott et al. (2002) argue that it should not. We follow Lott's model not because we necessarily agree that it is the most appropriate, but because we wish to focus attention on the use of placebo laws to better estimate standard errors.

to one thereafter—this is the simulated placebo law. We did not replace the state, but we did replace the year. That is, we did not allow a state to retract its placebo law during our sample or to pass shall-issue gun laws more than once, but we allowed different states to pass placebo laws in the same year. We drew 10 state/year combinations and ran a regression equivalent to that run by Lott and Mustard. From the regression we computed beta coefficients. We then repeated the entire procedure 1,000 times and computed the placebo PDF. The resulting distribution of coefficients is the distribution under the null hypothesis of "no effect." We then computed the standard deviation and the 97.5% and 2.5% percentiles of the placebo PDF to find the standard error and critical values under the null hypothesis of no effect. Since the placebo laws are drawn randomly, we should assume that their true effect is zero.[4] Thus, if the difference in the difference method generates appropriate standard errors, we would expect to reject the null hypothesis no more than 5% of the time. Below, we show that the null hypothesis of no effect is rejected much more often.

The first section of Table 1 presents the results of the Lott and Mustard model where shall-issue gun laws are modeled as a dummy variable taking on the value of zero prior to the law and one thereafter.[5] A number of other variables, including county- and state-fixed effects, arrest rates, population controls, measures of income, and so on, are also included in the model, but are not shown in the table. As the table indicates, the standard errors computed from the placebo distribution are typically three to four times larger than the original standard errors. The corrected 95% failure-to-reject region (i.e., the region bounded by the critical values) is written below the standard errors. In every case, the estimated coefficient is well within the failure-to-reject region. Note that the failure-to-reject region is computed directly from the placebo PDF, rather than based upon the corrected standard error. Thus, the critical values do not rely on a normality assumption (although in this case the differences are not substantial).

As noted above, many papers have reestimated the Lott and Mustard paper using somewhat different specifications or functional forms. The failure-to-reject regions in Table 1 highlight the problem with this approach—essentially all the "revised" figures are within the critical values making these results uninformative. In other words, conventional statistical inference finds these estimates to be indistinguishable.

For comparison purposes, we have also included standard errors clustered on states, the usual prescription for Moulton-type problems. Although the clustered standard errors are closer to the placebo distribution than the robust standard errors, they are still a significant underestimate (by some 20% to 50%), which is not surprising since clustering will not solve autocorrelation problems.[6]

Rather than focusing on the correct standard errors, we can also calculate how often the standard procedures would have falsely rejected the null hypothesis of no effect. In 1,000 trials using placebo laws on the murder rate, we found 282 out of 1,000 instances in which the null hypothesis would have been rejected at the greater than 5% level! Of these, 180 trials would have produced a coefficient ranging from 4.3% to 19.9% and averaging 9.1%. The remaining 102 cases where the null hypothesis would have been falsely rejected had coefficients on murder ranging from -4.6% to -17.4% and averaging -9.0%. Thus, it is clear that the standard procedure can easily lead to false inferences when not all the assumptions of the standard model are met.

As the true distribution from which the shall-issue laws are drawn is unknown, it is necessary to make some arbitrary decisions concerning the domain from which the placebo laws are drawn. Our first

---

[4] As should be expected given that these are placebo laws, we found the mean estimated effects to be equal to zero within two significant digits; that is, a typical mean was 0.00x.

[5] Results are not identical to those in Lott and Mustard because we use an expanded dataset constructed by Ayres and Donohue, which is very similar to that of Lott and Mustard but with the "correction" of some minor errors and an expanded number of years. Correction is placed in quotes because the differences are debatable (see Lott et al.).

[6] Ayres and Donohue (2003) discuss the impact of clustering, but do not focus on it and do not correct for serial correlation. Moody (2001) explicitly attempts to deal with serial correlation, but uses parametric techniques that Bertrand et al. find are less accurate than the placebo technique.

set of corrected standard errors comes from 10 randomly chosen states in random years post–1984. To check for robustness, we looked at other possible ranges from which to draw placebo laws. For example, in another run, we created *n* placebo laws for each post–1984 year, where *n* is the number of states that actually passed a shall-issue law in that year. In 1989, for example, we randomly chose three placebo states because in that year three states actually did pass a shall-issue law. The results are very similar to those found earlier. As before, in no case does the coefficient value fall outside the 95% critical values (these results are very similar to those found earlier and are not shown).

Another robustness test is to use the five years of additional data that have become available since Lott and Mustard's original paper. During these five years (1992–1997), 14 more states adopted shall-issue laws.[7] Does the additional data reduce the true uncertainty enough to draw firm conclusions? Not really. As before, the standard errors are much larger when a correct accounting is made of serial correlation and other factors (although, as expected, the standard errors do shrink with more data). None of the negative coefficients is statistically significant at the 5% level once the standard errors have been corrected. On the other hand, the positive coefficients on property crime, auto theft, and larceny all remain statistically significant at the 5% level even after correction. One interpretation is that we are more certain that shall-issue laws *raise* nonviolent crime rates than we are that shall-issue laws lower violent crime, such as murder, rape, aggravated assault, and robbery. This interpretation would be naïve, however, because it ignores the cross-equation restrictions implicit in the Lott model. Incorporating these restrictions casts the shall-issue laws in a more favorable light.

Lott and Mustard also estimate their model using before and after trends. Passage of a shall-issue gun law will not cause the population of potential victims to arm themselves overnight. An immediate before and after comparison, therefore, could easily miss important changes in the trend rates of crime. The final section of Table 1 presents the trend results from the original model. The coefficient on murder, for example, indicates that after shall-issue gun laws were passed the trend rate for murders fell by 4.3% compared to before the passage of the shall-issue law. As before, we simulated this model 1,000 times and generated a placebo PDF used to generate true standard errors and critical values. The standard errors are considerably larger than the original model suggests: up to twice as large in some cases. It is apparent that the use of trends advantageously absorbs some of the serial correlation discussed earlier, so the relative increase in standard errors is not as large as with the dummy model. Importantly, the coefficient on murder remains outside the failure-to-reject region. Thus, even with the revised standard errors the trend model indicates that shall-issue laws cause a large and significant drop in the murder trend rate.

5. Cross-Equation Restrictions

The Lott–Mustard theory of shall-issue gun laws is that they deter crime by making it more likely for a criminal to encounter an armed victim. Assuming this theory is correct, shall-issue laws should lower crimes against persons, but should have no direct effect on crimes against property, such as auto theft, burglary, or larceny. Indirectly, shall-issue laws could increase crimes against property if there is a substitution effect away from crimes against persons (i.e., if criminals take more care to avoid victims when victims have a higher probability of being armed, but continue to commit other types of crime). From this perspective, the negative coefficients on crimes against persons and the positive coefficients on crimes against property are potentially very informative. How likely is it that such a pattern would be produced by chance?

We can estimate the probability that a pattern is produced by chance by counting the number of times the pattern appears in data generated from randomly assigned placebo laws. In the 1977–1992 data generated by randomly choosing placebo states post–1985, we find only a .2% probability (i.e., two out of 1,000) of finding negative coefficients on crimes against persons and positive coefficients on crimes against property. Note that this is a very strong test since we look only for the negative/positive pattern.

---

[7] Since 1997 three additional states, Michigan, New Mexico, and Colorado, have passed shall-issue laws, although the New Mexico Supreme Court prevented the law's implementation in that state.

For example, in the placebo data there are no examples of the pattern occurring when the coefficient on murder is as large or larger than -7.3% (i.e., we estimate that the probability of this occurring by chance is less than one in 1,000). We find that the probability of the pattern occurring by chance is similarly low using the set of placebo laws generated from the same years and random states. Using the longer dataset, we find only seven out of 1,000 randomly generated cases fitting the negative/positive pattern. Only four fit the pattern with a coefficient on murder less than or equal to -7.8%, and zero fit the pattern with coefficients on murder less than or equal to -7.8% and on property greater than or equal to 7.6%.

6. Conclusion

Accounting for serial correlation, grouped data, and potential non-normalities can often be difficult especially when the exact form of the problem is not known. An alternative approach is to estimate standard errors from an empirical PDF generated via placebo laws. Using this approach, we reexamined Lott and Mustard's important and controversial study on shall-issue gun laws. We find that that corrected standard errors are dramatically larger than the reported standard errors. Using a dummy model to estimate the effects, we can rarely reject the null hypothesis of no effect once the correct standard errors are used. In this sense, the true uncertainty about the effect of shall-issue laws is larger than has previously been estimated. Standard errors are also larger when using a trend model, but in this specification the data continue to reject a null hypothesis of no effect on the murder rate trend. The Lott–Mustard theory, moreover, contains a little remarked upon cross-equation restriction. Shall-issue laws should reduce crimes against persons, but increase crimes against property. Using the placebo approach, we estimate the probability that these findings could occur by chance. Although individual coefficients are difficult to pin down, the negative/positive pattern of results is very rare in the placebo data. Surprisingly, therefore, we conclude that there is considerable support for the hypothesis that shall-issue laws cause criminals to substitute away from crimes against persons and towards crimes against property.

Table 1: The Estimated Impact of Shall-Issue Gun Laws on Crime, County Data

	Violent Crime	Murder	Rape	Aggravated Assault	Robbery	Property Crime	Auto Theft	Burglary	Larceny
Dummy Model (1977–1992)									
Coefficient	-3.5%*	-7.3%*	-4.8%*	-5.3%*	-0.1%	5.2%*	8.9%*	2.3%*	5.9%*
Original Standard Errors (robust)	1.2%	2.5%	1.5%	1.6%	1.9%	1.1%	2.0%	1.1%	1.9%
Placebo Standard Errors (post-85, random states)	4.9%	6.4%	5.6%	6.6%	7.5%	5.1%	6.5%	5.7%	5.7%
Critical Values/Failure to Reject Region	-9.2–10.7	-13.5–12.5	-12.3–11	-11.7–14.3	-15.4–14.2	-10.2–9.8	-12.0–13.9	-11.9–9.6	-11.1–10.4
Standard Errors Clustered on State (robust)	3.5%	5.5%	4.3%	4.8%	4.6%	2.6%	4.4%	3.5%	2.9%
Placebo Standard Errors (same years, random states)	4.6%	6.2%	5.1%	6.0%	6.7%	4.5%	6.1%	4.9%	5.3%
Dummy Model (1977–1997)									
Coefficient	0.2%	-7.8%*	-2.9%*	-0.1%	-0.4%	7.6%*	10.8%*	1.5%	9.6%*
Standard Errors (robust)	1.1%	1.7%	1.1%	1.3%	1.3%	0.8%	1.5%	0.9%	1.2%
Placebo Standard Errors (post-85, random states)	3.7%	5.4%	4.0%	5.2%	5.0%	3.6%	5.2%	3.9%	4.2%
Critical Values/Failure to Reject Region	-6.6–7.7	-11.0–10.2	-8.0–7.7	-9.8–10.1	-9.8–9.2	-7.4–6.5	-9.9–10.7	-7.8–7.4	-8.5–8.1
Trend Model									
Coefficient	-0.4%	-4.7%*	-1.7%*	0.5%	-1.9%**	0.1%	0.1%	-0.4%	0.8%
Original Standard Errors (robust)	0.5%	1.1%	0.6%	0.7%	0.8%	0.7%	0.9%	0.5%	1.4%
Placebo Standard Errors (post-85, random states)	1.1%	1.6%	1.0%	1.4%	1.4%	.9%	1.4%	1.0%	1.1%
Critical Values/Failure to Reject Region	-1.9–2.4	-2.9–3.1	-2.1–2.2	-2.5–2.9	-2.8–2.8	-1.7–1.8	-2.8–2.9	-2.1–2.1	-1.9–2.1

* Significant at the 1% level.

References

Ayres, Ian and John Donohue (1999) "Nondiscretionlary Concealed Weapons Law: A Case Study of Statistics, Standards of Proof and Public Policy," *American Law and Economic Review* 1:436–470.

— — and — — (2003) "Shooting Down the More Guns, Less Crime Hypothesis," *Stanford Law Review* 55(4): 1193-1312.

Black, Dan and Daniel Nagin (1998) "Do Right-To-Carry Laws Deter Violent Crime?" *Journal of Legal Studies,* 27:209–219.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2002) "How Much Should We Trust Difference-In-Difference Estimates?" NBER Working Paper 8841.

Bound, John, David A. Jaeger, and Regina M. Baker (1995) "Problems with Instrumental Variables Estimation When the Correlation Between the Instrument and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association* 90 (430): 443–450.

Card, David (1990) "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 43:245–257.

Dezhbakhsh, Hashem and Paul H. Rubin. (1998) "The Economics of Gun Control Lives Saved or Lives Lost? The Effects of Concealed-Handgun Laws on Crime," *The American Economic Review*, 88(2):468–474.

Donald, Stephen and Kevin Lang (2001) "Inference with Difference in Difference and Other Panel Data," Working paper.

Duggan, Mark (2001) "More Guns, More Crime," *Journal of Political Economy* 109(5):1086–1114.

Grogger, Jeff and Michael Willis (2000) "The Emergence of Crack Cocaine and the Rise of Urban Crime Rates," *Review of Economics and Statistics* 82(4):519–529.

Lott, John and David Mustard (1997) "Crime, Deterrence, and Right-To-Carry Concealed Handguns," *Journal of Legal Studies,* 26:1–68.

Lott, John (2000) *More Guns, Less Crime: Understanding Crime and Gun Control Laws* Second Edition, University of Chicago Press, Chicago.

Lott, John, Florenz Plassmann and John Whitley (2002) "Confirming More Guns, Less Crime," forthcoming *Stanford Law Review.*

Maltz, Michael and Joseph Targonski (2002) "A note on the use of county level UCR data," *Journal of Quantitative Criminology* 18(3):297–318.

Meyer, Bruce (1995) "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics* 13(2):151–161.

Moody, Carlisle (2001) "Testing for the Effects of Concealed Weapons Laws: Specification Errors and Robustness," *Journal of Law and Economics*, XLIV:799–813.

Moulton, Brent (1990) "An Illustration of the Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economics and Statistics,* 72:334–338.

Mustard, David (2002) "Re-examining Criminal Behavior: The Importance of Omitted Variable Bias," forthcoming *Review of Economics and Statistics*.

Plassmann, Florenz and T. Nicolaus Tideman (2001) Does the Right to Carry Concealed Handguns Deter Countable Crimes? Only a Count Analysis Can Say," *Journal of Law and Economics,* XLIV: 771–798.