

SHOOTING DOWN THE MORE GUNS, LESS CRIME HYPOTHESIS

Ian Ayres* and John J. Donohue III**

Abstract: John Lott and David Mustard have used regression analysis to argue forcefully that “shall-issue” laws (which give citizens an unimpeded right to secure permits for concealed weapons) reduce violent crime. While certain facially plausible statistical models appear to generate this conclusion, more refined analyses of more recent state and county data undermine the more guns, less crime hypothesis. The most robust finding on the state data is that certain property crimes rise with passage of shall-issue laws, although the absence of any clear theory as to why this would be the case tends to undercut any strong conclusions. Estimating more statistically preferred disaggregated models on more complete county data, we show that in most states shall-issue laws have been associated with more crime and that the apparent stimulus to crime tends to be especially strong for those states that adopted in the last decade. While there are substantial concerns about model reliability and robustness, we present estimates based on disaggregated county data models that on net the passage of the law in 24 jurisdictions has increased the annual cost of crime slightly -- somewhere on the order of half a billion dollars. We also provide an illustration of how our jurisdiction-specific regression model has the capacity to generate more nuanced assessments concerning which states might profit from or best harmed by a particular legal intervention.

* William K. Townsend Professor of Law, Yale Law School. ian.ayres@yale.edu

** William H. Neukom Professor of Law, Stanford Law School.
jjd@stanford.edu

We thank John Lott for generously sharing his state and county data sets with us. David Autor, Alan Krueger, Steven Levitt and seminar participants at Harvard, Columbia, and Washington and Lee law schools provided valuable comments. Jennifer Chang, Craig Estes, Melissa Ohsfeldt Landman, David Powell, Matt Spiegelman, Fred Vars, and Nasser Zakariya provided superb research assistance. We also gratefully acknowledge the research support we have received from Yale and Stanford Universities.

Table of Contents

Introduction	3
I. Theoretical Issues Concerning the Effect of Increased Carrying of Concealed Handguns	8
II. Empirical Issues in Estimating the Effect of Increased Carrying of Concealed Handguns	10
A. Introducing Explanatory Variables into the Panel Data Model.....	14
1. Model Specification.....	14
2. Control Variables.....	17
B. Comparing the Results Using the Zheng and Lott Controls	19
C. Sensitivity of the Lott Results to Time Period and Inclusion of Demographic Controls	21
D. Problems With Unequal Years of Data From Early and Late Adopters in the Pre-Post Comparison	24
E. Problems With Endogenous State Adoption	26
III. The Basic County Data Findings Of Lott and Mustard.....	29
A. Lott's County Data Analysis for 1977-92.....	30
1. The Dummy Variable Model.....	30
2. Lott's Trend (or Spline) Model.....	31
3. The Hybrid Model Testing for Main and Trend Effects	32
B. Extending Lott's County Data Through 1997	33
IV. Estimating State-Specific Passage Effects	36
V. Some Interpretations, Speculations, and Methodological Insights	43
Conclusion.....	51
Appendix 1 – Coding the Timing and Status of Shall-issue Laws.....	53

Introduction

In a remarkable paper published in 1997, John Lott and David Mustard managed to set the agenda for much subsequent work on the impact of guns on crime in America by creating a massive data set of crime across all U.S. counties from 1977 through 1992 and amassing a powerful statistical argument that state laws enabling citizens to carry concealed handguns had *reduced* crime.¹ The initial paper was followed a year later by an even more comprehensive and sustained argument to the same effect in a book solely authored by John Lott entitled *More Guns, Less Crime* (now in its second edition). The work by Lott and Mustard has triggered an unusually large set of academic responses, with talented scholars lining up on both sides of the debate.² Indeed, a panel of the National Academy of Sciences has been convened to sort through the now large body of conflicting studies.

But in the world of affairs rather than ideas, it did not take long for the National Rifle Association (NRA) and politicians across the country to seize upon the work of Lott

¹ John R. Lott, Jr. and David B. Mustard, Crime, Deterrence, and Right-to-Carry Concealed Handguns, 26 *Journal of Legal Studies* 1 (1997). A law that allows a citizen to carry a concealed handgun if he or she can demonstrate a need to a government official is a discretionary, or “may-issue,” law. The “shall-issue” laws are designed to eliminate discretion on the part of governmental officials by requiring them to issue a permit to carry concealed handguns unless specific and easily verifiable factors dictate otherwise. Essentially, these “concealed carry laws” enable adults without serious criminal records or identified mental illness to carry concealed handguns in virtually all public places. For a listing of the states that have shall-issue laws (according to two different sets of coding), see Appendix Table 1.

² Among the articles that are supportive of the more guns, less crime thesis are: Bronars, Stephen and John R. Lott, Jr. 1998. “Criminal Deterrence, Geographic Spillovers, and the Right to Carry Concealed Handguns.” *American Economic Review* 88(2): 475-79; Bruce Benson and Brent Mast, “Privately Produced General Deterrence” *Journal of Law and Economics* 44(2): 1-22, October 2001. John R. Lott, Jr. and William M. Landes, “Multiple Victim Public Shootings,” Working Paper, available at Social Science Research Network Electronic Library, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=272929 (Date Posted: June 10, 2001); Carlisle E. Moody “Testing for the Effects of Concealed Weapons Laws: Specification Errors and Robustness” Presented at the Conference on Guns, Crime, and Safety, December 10-11, 1999 at the American Enterprise Institute, Pages 1-17, December 20, 2000; and David Mustard, “The Impact of Gun Laws on Police Deaths” *Journal of Law and Economics* 44(2): October 2001.

Articles that raise doubts about the Lott and Mustard findings include: Dan A. Black and Daniel S. Nagin, “Do Right-To-Carry Law Deter Violent Crime?” *Journal of Legal Studies* 27 (January 1998): 211; Jens Ludwig, “Concealed-Gun-Carrying Laws and Violent Crime: Evidence from State Panel Data,” *International Review of Law and Economics* 18 (1998): 242; Zimring, Franklin and Gordon Hawkins. 1997. “Concealed Handguns: The Counterfeit Deterrent.” *The Responsive Community* 7(2): 46-60. Dezhbakhsh, Hashem and Paul H. Rubin. 1998. “Lives Saved or Lives Lost? The Effects of Concealed-Handgun Laws on Crime.” *American Economic Review* 88(2): 468-74. Duggan, Mark. 2001. “More Guns, More Crime.” *Journal of Political Economy* 109(5): 1086-1114.

and Mustard to oppose efforts at gun control and advance the cause of greater freedom to carry guns. For example, in the same year that the initial article was published, Senator Larry Craig (R-Idaho) introduced The Personal Safety and Community Protection Act, which was designed to facilitate the carrying of concealed firearms by nonresidents of a state who had obtained valid permits to carry such weapons in their home state. Senator Craig argued that the work of John Lott showed that arming the citizenry via laws allowing the carrying of concealed handguns would have a protective effect for the community at large because criminals would find themselves in the line of fire.³ On May 27, 1999, Lott testified before the House Judiciary Committee that the stricter gun regulations proposed by President Clinton would either have no effect or would actually cost lives,⁴ and a number of Republican Congressmen have since included favorable references in their speeches to Lott's work.⁵ Moreover, Lott has also testified in support of concealed gun laws before several state legislatures, including Nebraska (1997), Michigan (1998), Minnesota (1999), Ohio (2002), and Wisconsin (2002).⁶

This past summer, Lott's work was favorably cited in a letter to Attorney General John Ashcroft dated July 8, 2002, signed by eighteen state attorneys general, in support of Ashcroft's decision to interpret the Second Amendment as protecting the right of individuals to bear arms. The letter concluded with the following statement: "As chief law enforcement officers of our respective states, we wish to make one final point that is outside the scope of constitutional analysis. Simply put, your position on the Second Amendment is a sound public policy decision. There is an increasing amount of data available to support the claim that private gun ownership deters crime. That evidence comes both from the United States (particularly as highlighted in the empirical research of John Lott) and from abroad."⁷

Lott has also drawn upon his scholarly work to become a major popular commentator in the wake of various instances of violence and mayhem. After sixteen individuals were killed in a school shooting in Germany in April 2002, Lott attacked the strict gun control measures in Europe claiming that "The problem with such laws is that they take away guns from law-abiding citizens, while would-be criminals ignore them,

³ 143 *Cong Rec S* 5109 (May 23, 1997). Although Congress has not yet adopted this legislation, it was reintroduced by Congressman Cliff Stearns (R-Florida) in 2000 who also specifically cited Lott's work. 146 *Cong Rec H* 2658 (May 9, 2000).

⁴ John R. Lott, "Gun Regulations Can Cost Lives: Statement before the Crime Subcommittee of the House Judiciary Committee," *Federal News Service* (May 27, 1999).

⁵ John Doolittle, "Motion to Instruct Conferees on H.R. 1501, Juvenile Justice Reform Act of 1999," 106th Cong., 1st sess., 145 *Cong Rec H* 8645 (September 24, 1999); Larry Craig, "Gun Control," 106th Cong., 2d sess., 146 *Cong Rec S* 349 (February 7, 2000).

⁶ Of these states, only Michigan has adopted a right to carry law -- it passed the law in January 1, 2001. http://www.micatholicconference.org/pdf/FOCUS_Jan_2001.pdf. The Wisconsin Assembly adopted the proposed law supported by Lott on 2/26/02, but the Senate has not yet passed the bill. <http://www.legis.state.wi.us/2001/data/AB675hst.html>.

⁷ Bill Pryor, Alabama Attorney General, et al., Letter to Attorney General John Ashcroft (July 8, 2002), available at http://www.ago.state.al.us/ag_items.cfm?Item=81.

leaving potential victims defenseless.”⁸ After the 9/11 attacks, Lott similarly argued that “fears of having guns on planes are misplaced. The special, high-velocity handgun ammunition used on planes packs quite a wallop but is designed not to penetrate the aluminum skin of the plane.” Noting that “States that pass concealed handgun laws experience drops in violent crimes, especially in multiple victim shootings—the type of attack most associated with terrorism,” Lott argued that, “The use of guns to stop terrorists shouldn’t be limited to airplanes. We should encourage off-duty police, and responsible citizens, to carry guns in most public places. Cops can’t be everywhere.” Clearly, Lott’s message has been widely heard. This paper will explore whether the message is in fact true.

In addition to their statistical work, Lott and Mustard have also compiled a large body of anecdotal evidence concerning instances where law-abiding citizens have used guns to capture or thwart dangerous criminals. It may well be the case that many advocates of gun control have been inattentive to the possible benefits – in terms of protection or psychological comfort – that have at times been achieved by those lawfully carrying concealed weapons. Among the many anecdotes in Lott’s book designed to emphasize this point, consider the case of Suzanna Hupp, who was sitting in a cafeteria in Killeen, Texas in 1991 when a gunman crashed his car into the restaurant and began shooting the patrons. Although Hupp had a gun in her car which she believes might have been used to stop the killer, both her parents and 21 others died in the massacre. Following the horrific event, she verbally attacked those who had “legislated me out of the right to protect my family members.”⁹ Going on to become a legislator herself, Hupp now always carries a firearm, and she is trying to extend the right to carry concealed handguns (now legal in Texas in part due to her efforts) to the few remaining places where they are banned, such as in churches, university campuses, and public schools.

Of course, we don’t know whether Hupp would have been carrying her gun on her person that day if she had had the legal right to do so (the fact that she carries it now is not dispositive on that question), or if she would have been able to save her parents or others (rather than just become another casualty by staying and fighting instead of escaping out a shattered window to safety, as she did). It is clear, though, that her loss was terrible, and having a concealed carry law in effect at the time *might* have reduced the carnage in that situation (or might have added one or more other victims). But while Lott and Mustard have energetically catalogued the situations in which armed citizens have protected themselves or others, they never acknowledge cases on the other side of the ledger where the presence of guns almost certainly led to killings. For example, the nightmare scenario for those asserting the value of defensive use of guns is not

⁸ Supporters of tighter controls on guns note, however, that “among rich countries, the U.S. far surpassed any other country in firearm-related deaths—with 30,419 reported in 1998, or 11.3 per 100,000 people. That compared with 83 gun-related deaths in Japan and 197 in the United Kingdom during the same period, with rates of .1 and .3 per 100,000, respectively.” Rachel Zimmerman, “Study Finds Violence took 1.6 Million Lives in 2000,” *The Wall Street Journal* (Oct. 3, 2002), at D5.

⁹ Hector Tobar, “Texas Lawmaker’s Answer to Gun Violence: More Guns,” *Los Angeles Times*, p. A7, March 16, 2001.

mentioned: the case of the Japanese exchange student, Yoshihiro Hattori, on his way to a Halloween party in October of 1992 who mistakenly approached the wrong house and was shot to death by the homeowner Rodney Peairs. Peairs, who was later found civilly liable for the boy's wrongful death, was at home with his family when Hattori and his American host mistakenly rang the doorbell in search of the party. Peairs' wife answered and, apparently frightened by the costume, yelled to her husband to get his gun. Peairs shot Hattori dead after warning him to "freeze," a phrase the young man apparently did not understand. A Baton Rouge, Louisiana judge awarded more than \$650,000 in damages to the parents of the Japanese exchange student, saying there was "no justification whatsoever" for the killing of the 16-year-old boy.¹⁰ Although Mr. Peairs was acquitted of manslaughter, there is little doubt that if he hadn't owned a gun, the 16 year-old boy would not have been killed. At the memorial service for her dead son, Mrs. Hattori called the man who killed her boy "a victim of the American gun culture."

Or consider the story of Skip Olson, 58, and his roommate and friend of 25 years, Michael Jurisin, 50. The two Palo Alto, California residents were fighting about rent payments when Jurisin took out a handgun. Olson grabbed Jurisin's gun and shot him in the back of the head on February 17, 1998. Olson was later convicted of second-degree murder. One suspects that if neither man had owned a gun, no one would have been murdered.¹¹ In the end, one must acknowledge there are costs and benefits of either allowing or prohibiting the carrying of handguns, and the task for the scholar is to try to determine which effect dominates.

The existence of a widely cited study based on the statistical analysis of a massive dataset that is invoked in both political and popular circles as an argument against most forms of gun control suggests that careful scrutiny of the empirical evidence is warranted. Lott and Mustard based their analysis on the current (although increasingly questioned) state-of-the-art technique of micro-econometric evaluation – a panel data model with fixed effects.¹² That is, Lott and Mustard began by collecting data over a period of years (1977-1992) for individual states and counties across the United States, and then used panel data regression techniques to estimate the effect of the adoption of shall-issue laws, controlling for an array of social, economic, and demographic factors.¹³ In earlier work,

¹⁰ Adam Nossiter, "Judge Awards Damages In Japanese Youth's Death," *The New York Times*, September 16, 1994, Section A; Page 12; Column 1.

¹¹ Emily Richmond, "Roommate Found Guilty of Second-Degree Murder," *Palo Alto Daily News*, May 27, 1999 at 3.

¹² There are actually two "fixed effects." The first is a dummy variable for each county or state that is designed to reflect any unvarying trait that influences crime in that county or state yet is not captured by any of the other explanatory variables. The second is a set of "year fixed effects," which are dummy variables included for each year of the data set to capture any national influence on crime that is not captured in any of the other explanatory variables, but which might be expected to affect all jurisdictions equally. A full list of the variables included in the regressions (other than year and county dummies) and their summary statistics is included in Appendix Table 2.

¹³ Lott in his book and a variety of other articles also supports his "more guns, less crime" thesis with other types of evidence. For example, Lott collected data at the city and state levels to test whether different sized jurisdictions exhibited a reduction in crime when a shall-issue law was adopted. See Lott (169 & 190). Lott was also able to secure permit data from ten states that he used to test whether counties

we commented on concerns that we had about model reliability based on Lott's analysis of 1977-92 data evaluating the effect of the adoption of shall-issue laws in 10 states.¹⁴ We opined on the potential theoretical and empirical infirmities in that analysis, and noted the value in further study given that more state adoptions and the passage of time would likely either strengthen Lott's case if it were true or weaken it if it were false. Having extended the state data through 1999 and the county data set through 1997, we are now able to test that prediction. We conclude that Lott and Mustard have made an important scholarly contribution in establishing that these laws have not led to the massive bloodbath of death and injury that some of their opponents feared. On the other hand, we find that the statistical evidence that these laws have reduced crime is limited, sporadic, and extraordinarily fragile. Minor changes of specifications can generate wide shifts in the estimated effects of these laws, and some of the most persistent findings -- such as the association of shall-issue laws with increases in (or no effect on) robbery and with substantial increases in various types of property crime -- are not consistent with any plausible theory of deterrence. Indeed, the probabilistic underpinnings of statistical analysis suggest that running regressions for nine different crime categories to see if there is any measurable impact on crime will, by chance alone, frequently generate estimates that on their face are "statistically significant," and it may well be the case that the scattered negative coefficients for various violent crime categories should be thought of as statistical artifacts.¹⁵ While we do not want to overstate the strength of the conclusions that can be drawn from the extremely variable results emerging from the statistical analysis, if anything there is stronger evidence for the conclusion that these laws increase crime than there is for the conclusion that they decrease it.

The remainder of this article is divided into four parts. The first section discusses the theory underlying Lott's empirical project, discussing the ways in which shall-issue laws could dampen -- or increase -- crime. Section II delves into a host of methodological issues a researcher must confront in estimating the impact of law on crime. Using a state data set with several additional years of information, this section shows that while the simplest regression models suggest that crime has tended to increase more when states adopt shall-issue laws, these results vary over different time periods (yielding the

with more concealed handgun permits have larger reductions in crime. Lott (178). In addition, Lott has concluded from other work that counties located next to states passing shall-issue laws experience an increase in crime just as the passing states experience a decrease. Although our disaggregated analysis below will cast doubt on this finding, it should still be noted that any estimate of the overall effect on crime from these laws will be biased toward making them look more beneficial than they are if they cause crime to shift from passing to non-passing states. John Lott & Steven G. Bronars, *Criminal Deterrence, Geographic Spillovers, and Right-to-Carry Concealed Handgun Laws*, 88 *Am. Econ. Rev.* 246 (1988). We do not have these additional types of data and therefore do not analyze them in this article.

¹⁴ Ayres and Donohue, "Nondiscretionary Concealed Weapons Law: A Case Study of Statistics, Standards of Proof, and Public Policy," 1 *American Law and Economics Review* 436 (1999).

¹⁵ As Milton Friedman stated: "I have long had relatively little faith in judging statistical results by formal tests of statistical significance. I believe that it is much more important to base conclusions on a wide range of evidence, coming from different sources over a long period of time." Quoted in Daniel Hamermesh, "The Craft of Labormetrics," 53 *Industrial and Labor Relations Review* 363 (2000).

opposite results before crime began dropping faster in the 1990s in the non-adopting states) and are surprisingly sensitive to inclusion of seemingly extraneous right-hand side control variables. While the Lott and Mustard model estimated on state data gives some support for the view that shall-issue laws can lower murder and rape, better models undermine this conclusion. The most robust findings from the most up-to-date state data is that various property crimes rise with passage, although the absence of a compelling theoretical justification for this result raises concerns about the predictive validity of the models. Section III takes up Lott’s own county data set to assess the extent to which the “more guns, less crime” result persists in less constrained specifications with additional years of data. Section IV explores even less constrained regressions, in particular estimating state specific effects, and finds that the core finding of more guns, less crime is reversed once the statistically preferable state specific regression models are used. Section V discusses a hierarchy of possible conclusions to emerge from our empirical work, and provides an illustration of how state-specific regression models can potentially provide more nuanced policy recommendations across states than are possible with more customary aggregated models.

I. Theoretical Issues Concerning the Effect of Increased Carrying of Concealed Handguns

Given the massive extent of gun ownership in this country, coupled with the fact that the United States is exceptional in only one aspect of its crime problem – its high rate of lethal violence – it might at first appear that guns must be a part of the problem. But over the last decade, a number of scholars have offered theoretical and empirical support for the notion that allowing law-abiding citizens to carry concealed handguns (unobservable to potential criminals) can deter criminal behavior.¹⁶ The theory is that criminals will be willing to arm themselves whether or not this is lawful, so that laws designed to restrain gun ownership and carrying will only serve to protect criminals, who will have a lessened fear of encountering armed resistance to their criminal designs. Allow the law-abiding to carry guns, so the theory goes, and the costs of engaging in criminal activity will rise, thereby dampening the amount of crime.

Lott and Mustard’s empirical project is grounded in the important theoretical insight that hidden precautions by potential victims can generate powerful general deterrence effects. Visible precautions by potential victims may simply tend to displace crime toward victims who take less precaution, while unobservable precautions (silent alarms, gasoline kill switches, Lojack) make potential criminals generally more reluctant to commit crime. Thus, while the conventional wisdom focuses on the danger that more guns pose to the citizenry, the new critique emphasizes the protective effect that spills

¹⁶ Kleck, Gary, *Point Blank: Guns and Violence in America*, A. de Gruyter, NY, 1991; Polsby, Daniel D., “The False Promise of Gun Control”, *The Atlantic Monthly*, Vol. 273, No. 3, pg. 57 (March 1994).

over from those who carry concealed weapons. Because criminals cannot know in advance who is armed with a concealed weapon, their risk goes up in an encounter with any potential victim. Note, then, that even though the open carrying of handguns might only divert criminals from potential victims with guns to those without them, legalizing the concealed carrying of weapons holds out at least the potential of reducing crime rather than merely shifting its incidence.

The first rejoinder to this view is that shall-issue laws allow anyone of a certain age without an officially documented problem of mental health or criminal record to secure a permit to carry a concealed weapon, which is not a particularly exacting standard. A moment's reflection on one's own acquaintances would likely suggest the names of numerous angry or intemperate individuals who could pass the "shall-issue" test even though the prospect of their carrying a concealed weapon would not be likely to enhance one's sense of personal security. Still, Lott and Mustard have a pretty good reply to this point – the number of crimes committed by those who have obtained permits appears to be rather small (although it is doubtless higher than official records would suggest as the identity of the perpetrator in a substantial proportion of crimes is never discovered).

But even if no one securing a concealed carry permit ever used it to commit a crime, there are still a number of avenues by which the passage of a concealed carry law could stimulate crime. First, even if the adoption of a shall-issue law increased the riskiness of criminal activity and thereby dampened the number of criminals, it might also *increase* the number of criminals who decided to carry weapons themselves (by hypothesis, illegally) and also might *increase* the speed at which a criminal decides to shoot or disable potential victims (as the presence of armed victims increases the cost of hesitation once a criminal engagement has been launched). Therefore, the number of murders and aggravated assaults can rise if criminals respond to shall-issue laws by packing more heat and shooting quicker. Arming the citizenry can encourage an arms race leading more criminals to carry even higher-powered weapons, and to discharge them more quickly when threatened.¹⁷

Second, even when no criminal act is initially contemplated, the injection of a gun into an angry dispute, perhaps in lawful defense, might escalate a minor dispute into a criminal homicide or serious wounding.¹⁸ As an earlier president of the Connecticut Chiefs of Police Association once stated, "We are concerned about the increasing availability of handguns and the ease with which a person can get a pistol permit. [A] permit is dangerous in the hands of a neophyte who goes to a bar and shows off his phallic symbol to the boys."¹⁹ Indeed, there was a bit of a scandal in Connecticut in 1977 when it was revealed that Michael O'Brien -- deemed by the federal organized crime strike force special prosecutor as one of the "two most important criminals in the

¹⁷ John Donohue and Steven Levitt, "Guns, Violence, and the Efficiency of Illegal Markets," 88 *American Economic Review* 463 (May 1998).

¹⁸ For an illustration of this, see the text at footnote 11, *supra*.

¹⁹ Lincoln Millstein, "Police Toughen Criteria for Getting Gun Permit," *Hartford Courant*, January 15, 1978, at 1.

Hartford area” and convicted for racketeering, extortion, and gambling -- had obtained a right to carry a concealed weapon with the support of letters of recommendation from certain major political figures in the state.²⁰ This suggests that those who are able to secure handgun permits are not always model citizens, and that at least some criminals find it useful to have the legal right to carry weapons.

Third, with some estimates suggesting that as many as one million or more guns are stolen each year, we know that putting more guns in the hands of the law-abiding population necessarily means that more guns will end up in the hands of criminals.²¹ In fact, with guns being a product that can be easily carried away and quickly sold at a relatively high fraction of initial cost, the presence of more guns can actually serve as a stimulus to burglary and theft. Even if the gun owner had a permit to carry a concealed weapon and would never use it in furtherance of a crime, is it likely that the same can be said for the burglar who steals the gun?

Fourth, allowing citizens to carry concealed weapons imposes burdens on police in that they must ascertain whether the gun is being carried legally. Officers of the Illinois State Police have indicated that their job would be complicated if certain individuals were permitted to carry guns as time would need to be spent on confirming whether some guns were being legally carried.²² As it stands now in Illinois, anyone caught with a gun in public is violating state law and can be immediately brought into custody without the need for further investigation, which the state police believe has been a powerful tool for taking criminals off the streets.²³ Finally, accidental deaths and suicides are obviously aided by the presence of guns, and these costs could conceivably outweigh any benefits of shall-issue laws in reducing crime. Extensive empirical study is needed to assess the relative magnitudes of the likely conflicting effects of a state’s decision to permit citizens to carry concealed weapons.

II. Empirical Issues in Estimating the Effect of Increased Carrying of Concealed Handguns

It is probably useful to begin our empirical investigation by giving a sense of the nature of the difficulties in trying to uncover the ultimate impact on crime of the adoption

²⁰ Andrew Kreig, “Pair Gets 3 to 10 Years in Prison in Racketeering Case,” Hartford Courant, July 6, 1977 at 3.

²¹ Duggan (2001: 1094) reports that 500,000 guns are stolen in the United States every year (others estimate the figure to be as much as 3 times that number).

²² Donohue has served as a consultant to the Illinois State Police on matters relating to a claim of racial profiling.

²³ Lott also identifies another mechanism by which shall-issue laws might induce crime: they may embolden citizens to frequent high-crime areas (or enter areas at night) that would be avoided were they not carrying a gun, thereby increasing possible exposure to criminal acts. While we would lament the increased crime in this situation, the greater willingness to move about would be a benefit associated with the right to carry weapons.

of concealed-carry laws. To this end, look at Figure 1, which shows the pattern of robbery over the period from 1977 through 1999 in four groups of states: the four states that had shall-issue laws prior to 1977, the eight adopters between 1977 and 1989, the 17 adopters between 1990 and 1999, and the 22 states that have never adopted a shall-issue law.²⁴ Robbery is a good place to start our inquiry because it is committed in public more than any other crime, and should be the crime most likely to decline if the Lott and Mustard story of deterrence has any plausibility.²⁵ The first thing one notices in this figure is that (particularly for the top three state groupings) there is a broadly similar pattern for the four groups: robbery rates rose until about 1980 then dipped, began rising yet again starting in roughly the mid-1980s, followed by a decline in the 1990s. This suggests that there are some factors operating across the entire nation that tend to push crime up and down in broad waves lasting from five to ten years. The second point that leaps out from the figure is that the 22 states that have not adopted shall-issue laws have had much higher rates of robbery than states that allow the carrying of concealed handguns, at least until recently (more about this later). Note that this is not what Lott and Mustard mean when they suggest more guns, less crime. They realize, as sophisticated researchers, that in 1977, the 22 never-adopting states had double the robbery rate of the other states for reasons having nothing to do with their lack of shall-issue laws. Indeed, only four of the other 29 states allowed the carrying of concealed handguns at that time. The main story is that robbery occurs more frequently in large, densely populated urban areas. Thus, one could not hope to establish the effect of a shall-issue law by looking only at which states have such laws and which don't at any one point in time – a so-called “cross-section” analysis. Even the most zealous supporters of shall-issue laws should realize that introducing the right to carry concealed handguns could not cut the robbery rates by anything close to one-half – clearly, other factors explain the large differences in the core rates of crime in the different sets of states.

Given this fact, is there anything that can be said about the likely effect of adopting a concealed-carry law on robbery? Certainly nothing definitive, and we will explain why shortly, but it is worth noting that the only group of states to have experienced a substantial drop in robbery over this time period (albeit one punctuated by two sharp upturns) was the 22 states that never adopted shall-issue legislation. These states experienced a whopping drop in robberies of roughly one-third, a drop so large that the never-adopting states went from having by far the highest robbery rates of the four groups in 1977 to ending up in 1999 with the same robbery rate as the eight adopters over

²⁴ As Appendix 1 indicates, there has been some dispute as to whether certain states have a shall-issue law. The Figure 1 graphs are constructed using the shall-issue coding supplied in Jon Vernick and Lisa Hepburn, “Description and Analysis of State and Federal Laws Affecting Firearm Manufacture, Sale, Possession, and Use, 1970-1999.” (Brookings Paper, forthcoming 2002) (hereinafter “the Vernick coding”). Appendix 1 sets forth the Vernick coding (column D) and an alternative coding that is closer to the coding used by Lott and Mustard (column A) that we will refer to as “the Lott coding.”

²⁵ Lott and Mustard have tried to argue that because some robberies are of banks and other commercial entities that are already protected by armed guards, the predicted effect of shall-issue adoption on robbery is uncertain. We are not persuaded by this claim.

the period from 1977-89.²⁶ Indeed, if one were forced into making causal attributions from this graphical data, one might conclude that shall-issue laws tend to increase robbery rates. Similar conclusions could be derived from an examination of the other crime categories depicted in Figures 1b-1f.²⁷ But, there is no need to rely on visual inspection, since the statistical tool of regression can more formally and precisely do what the graphical analysis is trying to do – control for the initially different levels of crime and the common national forces acting on crime to see whether shall-issue adoption has any systematic effect on crime. We ran just such a regression model that controlled only for the average crime rate in the state and the common national influence each year and found that adoption of a shall-issue law was associated with an almost sixteen percent *increase* in robbery.²⁸ Indeed, as Table 1, line 2 shows, running these same parsimonious regressions across all nine crime categories for the period 1977-99 yields results that, with only a single exception, are uniformly statistically significant and positive, suggesting – however naively -- that shall-issue laws *increase* crime.²⁹

Unfortunately, life is not so simple. First, note that the story that was just told was very sensitive to the time frame we examined. If our evaluation had been conducted at a time when data was only available through 1992, which was the case for the initial Lott and Mustard paper, then we would not have seen the dramatic drop in robbery in the non-adopting states. As Table 1, line 1 shows, over the 1977-92 period, the simple regression would have suggested that shall-issue laws were associated with statistically significant *decreases* in murders and aggravated assaults. Second, the Table 1 predictions would be valid estimates of the impact on crime of the adoption of a shall-issue law only if two conditions held: 1) each state had a fixed underlying rate of crime that only changed when some omnipresent national influence of crime pushed that rate up or down by the same amount everywhere or when a shall-issue law was adopted; and 2) the fact

²⁶ A somewhat similar pattern can be seen in Figure 1b for violent crime. While the never-adopters experienced a small *drop* in violent crime, the three other sets of shall-issue states experienced violent crime *increases* over this period, and for the adopters after 1977, those increases were sizable. Indeed, while both sets of post-1977 shall-issue adopters started out with equal levels of violent crime that were substantially below the level in the 22 never-adopting states, by 1999 the never adopters had a violent crime rate about equal to that of the late-adopters and clearly below that of the adopters in the 1977 through 1989 period.

²⁷ Consider the pattern of rape in the four state groups, shown in Table 1d. While the never-adopting states started out with the highest level of rape among the four groups, it ended the period with the lowest level. The same pattern can be found, yet even more dramatically, for property crime.

²⁸ The more technical description of this regression model would be a panel data model – that is, one combining crime data over a period of time with cross-section data for 50 states and the District of Columbia – controlling only for “state and year fixed effects.” See discussion at footnote 12, *supra*.

²⁹ The sole exception is murder, which has a small negative but wholly statistically insignificant coefficient. The other crime rate increases estimated in Table 1, line 2 range from a low of 3.9 percent for aggravated assault to a high of 23.2 percent for auto theft. In order to conserve space, our tables will only report coefficients of interest relating to the impact of the shall-issue law. But the interested reader can find the complete regression output for all the regressions in this paper (as well as the underlying STATA do files and data sets for independent verification) on the internet at www.yale.law.edu/ayres/.

and timing of adoption of a shall-issue law was not influenced by the current trend or level of crime in a state (thereby approximating a random treatment effect).³⁰

Figure 1 provides some visual support for part of the first condition in that one can discern the presence of strong influences on crime that seem to operate across all states.³¹ But the second part of the first condition -- the claim that the only additional statewide influence on crime beyond the fixed state effect and the national year effect is the adoption of a shall-issue law -- is obviously heroic. Some of the major measurable influences on crime, such as economic conditions, poverty rates, police activities, rates of incarceration, or demographics, may vary across states. For example, some states may decide to get tougher on crime by increasing their rates of incarceration faster than other states. If such factors *both* change in different ways across states *and* influence crime significantly, then the claim that the only changes in crime are caused by uniform national influences or the shall-issue legislation is undermined. As we will see, more complete regression models will be used below to try to control for these other plausible influences on crime, and to the extent that this effort is successful the regressions will (hopefully) succeed where the reliance on graphs could fail. But note that the correctness of the regression results will be imperiled if there are influences – such as, say, the criminogenic influence of the introduction of crack cocaine – that impact differentially across states and are not controlled for in the regression model, yet are correlated with the adoption of a shall-issue law.

One can get a glimpse of how this problem of omitted variable bias might be operating in this context by examining Figure 1c, which shows the pattern of murder rates for the four state groupings. Note that murder rates spiked into the early 1990s in the two top groups (which either never adopted a shall-issue law or adopted one in the 1990s) but did not show the same increase in the group of states that adopted in the 1977-1989 period. The Table 1 regressions would attribute the greater run-up in murder starting in the late 1980s to the failure of the top two groups to adopt shall-issue laws – hence Table 1, line 1 found the shall-issue laws to be associated with a substantial 8.3 percent *decrease* in murder over the period 1977-92. But an alternative explanation is that the crack cocaine problem drove up crime in certain highly urban areas in states that were reluctant to enact a shall-issue statute while largely bypassing more rural states that were adopting such laws. In this case, the regression would identify a relationship between higher crime and the failure to adopt a shall-issue law when the real cause would have been the influence of crack, for which the regression failed to control.

³⁰ As those familiar with tests of the effectiveness of medicinal drugs or medical interventions know, the gold standard for ascertaining the effect of a treatment is to randomize individuals into treatment and control groups and observe if there are statistically significant differences in outcomes across the two groups. In the medical context, the benefits of this approach would be lost if individuals were free to choose whether to take the treatment. Similarly, in our evaluation of the impact of a legal change, the regression results can be misleading when legislatures are free to enact or reject the shall-issue law instead of having the law randomly assigned to different states. This issue, referred to as the problem of endogenous state adoption, is addressed at greater length in Section II.E, *infra*.

³¹ The pattern is very clear for robbery and violent crime – Figures 1a and 1b.

The second issue relevant to estimating the effect of a shall-issue law is whether the law can be thought of as being adopted as a random treatment – that is, the law is an exogenous event – rather than being part of the endogenous system that influences crime and the various responses to crime. In other words, if spikes in crime lead to demands for the right to carry handguns, which turns out to be the case, then adoptions of shall-issue laws are not exogenous but are endogenous to the variable that we are trying to explain, which is crime itself. In this event, the regression might attribute subsequent drops in crime, even if merely reflecting a return to more normal rates of crime, to the passage of the law.³² This endogeneity problem is a vexing one in all studies examining the impact of a legal change, and perhaps particularly so when examining a law’s impact on crime (because politician’s like to be seen as “solving” crime problems by passing laws). We will also try to address this issue – admittedly with only limited success – later in this article.

A. Introducing Explanatory Variables into the Panel Data Model

Putting aside the endogeneity problem for now, we begin our exploration of the panel data regressions with the finding that shall-issue laws seem to increase crime – although, as Table 1 revealed, we might have thought differently about the impact on murder and aggravated assault if we had had to rely on data ending in 1992 rather than 1999. But before we can repose confidence in conclusions from graphical evidence or the parsimonious regressions of Table 1, we must probe whether introducing controls for other possible influences on crime might reverse this initial conclusion. For this to happen, the omitted variables would have to be correlated with both the passage of shall-issue laws as well as with higher crime. Before turning to these results, we must first address a few details about the nature of the regression models (specification) and the explanatory (or control) variables.

1. Model Specification

Although the regression model employed by Lott and Mustard is for a panel data set with both state (or county) and year fixed effects, even within this structure, a number of decisions need to be made concerning how one models the impact of the law. In trying to determine the impact of the passage of shall-issue laws, one would ordinarily begin by asking whether crime was on average higher or lower after the law went into effect, controlling for all the explanatory variables that are thought to impact on crime in the state. This is the so-called “dummy variable” model, which posits that the law will have a fixed (once-and-for-all) percentage impact on crime -- that is, the law will raise or lower crime by, say, 5 percent. Lott and Mustard begin their analysis with this dummy

³² The “return to more normal rates of crime” after a sudden upturn is referred to as regression to the mean.

variable model, but also explore a second specification in which they attempt to estimate whether the passage of the law will cause a break in the time path of crime, causing it to tip up or down depending on whether the law raises or lowers crime. In conducting this trend (or “spline”) analysis, Lott and Mustard estimate what the average linear time trend of crime is before the law passes, and then probe whether this trend changes after passage.³³

The Dummy Variable Model: If Lott and Mustard were correct that the passage of a shall-issue law reduces crime, one might see a sudden and persistent drop in crime that would be captured by a post-passage dummy in the panel data regression. Since typically the dependent variable in these regressions will be the natural log of the crime rate, the coefficient on the post-passage dummy variable can be interpreted as the percentage change in crime associated with the adoption of the law.³⁴ Lott interprets negative coefficients on the post-passage dummy variable to imply that prospective criminals anticipate the dangers they would face in trying to prey upon a more armed population and that they would reduce their criminality in response. Of course, one way that prospective criminals could reduce any increased personal risk to them from such a law would be to continue their life of crime in a more hospitable jurisdiction (presumably one without a shall-issue law) to pursue their criminal objectives.³⁵

The Lott Spline Model: Again assuming Lott and Mustard are correct that the law reduces crime, one might observe crime falling as individuals applied to the relevant state officials to secure the right to carry concealed weapons and then in fact began carrying them. Since this process might unfold gradually, one might expect to see a gradual and continuing decrease in crime – at least until the increase in the number of citizens carrying lawful concealed weapons came to an end. In this model (the trend or “Lott spline model”), then, a time trend would emerge after passage reflecting a dampening effect on crime that grew as the number of concealed handguns being carried increased.

In response to criticisms about the robustness of the dummy variable (or “static”) specification results,³⁶ Lott has correctly noted that, as a theoretical matter, shall-issue laws could still dampen the trend of crime without showing any effect in the simple

³³ Instead of reporting regression coefficients, with regard to his linear specification, Lott only reports the change in the before and after linear trends (see Lott’s original Table 4.8), and with regard to quadratic terms in the time trend, he graphs the before and after quadratic effects (see Lott’s original Figures 4.5-4.9).

³⁴ The Table 1 estimates are from just such a model (albeit controlling only for state and year fixed effects) and can thus be interpreted in this manner.

³⁵ Note that the simple panel data results cannot distinguish between an effect of a shall-issue law that reduces crime overall versus one which only shifts it to another jurisdiction. Obviously, crime transfers are much less desirable from a policy perspective than crime reductions. Lott & Bronars, *supra* note 13, claim to have found such geographic substitution, but our disaggregated analysis below casts doubt on their findings.

³⁶ See Ayres, Ian, and Donohue, John, “Nondiscretionary Concealed Weapons Law: A Case Study of Statistics, Standards of Proof, and Public Policy,” 1 *American Law and Economics Review* 436 (1999).

dummy regressions if crime were to follow an “inverted V” pattern. For example, Figure 2 depicts a case in which the crime rate increased for five years in a state before the law’s passage and then symmetrically declined in the five years after the law’s passage. In this situation, the static regression would estimate no systematic impact of the law (as depicted by the constant estimated horizontal lines before and after passage) – even though there is a large change in the before and after slopes.³⁷

Lott argues that the “inverted V” theory is borne out in the data, and, indeed, the graphs in his book of the trends in crime for the states adopting shall-issue laws appear to show that the linear trend in robbery and other crime categories were increasing prior to passage of the laws and fell thereafter. But to sort out whether the dummy variable model, the linear trend model, or any other model can support Lott’s story that shall-issue laws *caused* a downward shift in the trend of violent crime, we examine Lott’s estimated static and trend effects using more data (extending Lott’s data set through 1999 for our state data analysis and through 1997 for our county data analysis) and less constrained econometric specifications.

The Hybrid (or Main Effect Plus Trend) Model: While Lott and Mustard employ both of the preceding models, a third more general model is actually a hybrid of the preceding two in that it allows a post-passage dummy to capture the main effect of the law but also allows the law to change the linear trend in crime for adopting states. This model could be important if an “announcement effect” initially scares some criminals into fearing possible victim or bystander retaliation, but that the ultimate effect is that more guns leads to more serious criminal acts – perhaps as fist fights end with someone dead or seriously injured instead of with a bloodied nose.³⁸ Under this scenario, one might even see an initial drop in crime followed by a subsequent turn around as the number of concealed guns being carried and crime increase in tandem. Note that, although Lott does not employ this model,³⁹ it can be used to test whether one or both of the first two models is appropriate.⁴⁰

The hybrid model will generate two estimated effects that could either be reinforcing (both the dummy and trend have the same sign), or in conflict in that one effect is positive and the other is negative. We will later see empirical evidence

³⁷ Id. at 445. Note that Figure 2 defines the “true impact” of the law by the difference between the pre-passage linear trend and the observed post-passage time path of crime. Note the need for assumptions about the unobserved counterfactual – what would have happened had the law not been passed. “The Search for Truth: In Appreciation of James J. Heckman,” 27 *Law and Social Inquiry* 23 (2002).

³⁸ Zimring and Hawkins first characterized the announcement effect and the increased-risk effect in Franklin Zimring and Gordon Hawkins, “Concealed Handguns: The Counterfeit Deterrent,” *The Responsive Community* 7(2) (Spring 1997): 46-60.

³⁹ Lott, however, did use the hybrid specification in his analysis of geographic substitution. See Lott & Bronars, *supra* note 13 at 250.

⁴⁰ If the estimated coefficient on the post-passage dummy were virtually zero, one would reject the first model, and if the estimated coefficient on the time trend were virtually zero, one would reject the second model. If they were both virtually zero, one would conclude that the law had no effect on crime.

suggesting that regression results indicating the main effect of the law would be pernicious while the trend effect is benign is likely generated by some model misspecification rather than evidence that the law actually generated this pattern.⁴¹ Lott does suggest a way in which a pernicious main effect could be followed by a benign long-term trend effect, but this argument is unconvincing. In discussing his findings that public shootings increase for a few years after passage of nondiscretionary handgun laws, Lott suggests that there might have been a type of temporal substitution: people planning such shootings might “do them sooner than they otherwise would have, before too many citizens acquire concealed-handgun permits” (p. 102). But we find temporal substitution to be an unlikely explanation for the aggregate behavior of violent criminals.⁴²

The opposite pattern – a negative dummy and a positive time trend -- could occur in a number of ways. For example, this pattern would emerge if the announcement effect dampens crime more powerfully than any effect initially generated by the actual higher risk to the potential lawbreakers of confronting an armed citizen, but over time this effect is overwhelmed by the stimulus to crime that greater gun prevalence generates.⁴³ Similarly, this pattern would be seen if an initial announcement effect led to a drop in crime, followed by a return to previous levels of crime as the salience of the new law recedes from the consciousness of criminals. The Lott spline model, in contrast, essentially posits no announcement effect and assumes a gradually growing risk to lawbreakers generates a reduction in crime (the Lott thesis) or that the gradual pernicious influence of more guns in the hands of hotheads or criminals leads to more shootings and opportunistic criminal acts. Of course, if both influences are operating – an announcement effect dampening or increasing crime and an increased risk effect (reducing crime if the risk is to lawbreakers and increasing it if the risk is to the public) – then the Lott spline model would be mis-specified.

2. Control Variables

Lott tries to control for an array of measurable factors – listed in Appendix Table 2 -- that might influence crime rates, such as the arrest rate for a particular crime category in a state or county, the level of income, and various demographic measures. Since many

⁴¹ To foreshadow the later empirical evidence, the problem is revealed in Figure 3a, in which one sees that the effect of the shall-issue law is completely absent until after the tenth year (when many late-adopting states drop out of the analysis). At that point, the regression perceives that crime has dropped sharply (although this is a pure selection effect). The result is that the regression line has to tilt sharply down, and this downward tilt is accommodated through an ostensible jump upwards in the main effect.

⁴² We can conceive of a more plausible potential theoretical mechanism for generating a positive main effect and a negative trend: in the wake of passage of a shall-issue law the increased gun carrying generates more harmful, offensive uses of these weapons, but that over time criminals are deterred by the potential defensive uses of guns.

⁴³ As noted above in footnote 21, a very large number of guns are stolen each year, so if shall-issue laws encourage more lawful gun ownership, they probably lead to an increase in the number of guns that are stolen. Note that these thefts could undo any initial benign effect of the law in a way that leads to a negative post-passage dummy (as crime initially dropped) and a positive time trend (in light of the increased arming of criminals).

of these variables make little difference to the analysis it is not worth arguing about whether they should be included or not.⁴⁴ A potentially problematic variable employed by Lott and Mustard is the arrest rate, which is used to capture any changing deterrence that might be resulting from alterations in the intensity or effectiveness of police or prosecutorial resources. Reliance on the arrest rate (measured as the ratio of total arrests for any particular crime to the total number of occurrences of that crime) can lead to a large amount of data being excluded because of the realization of the dependent variable, which is always problematic.⁴⁵ Accordingly, we examined the sensitivity of the Lott results to using the state incarceration rate as an alternative and, in some ways, preferable measure of the likelihood of punishment.⁴⁶ To dampen problems of endogeneity, we used the lagged value of the incarceration rate. But whether one uses arrest rate or incarceration rate one must acknowledge the possible problem that instead of being a truly independent variable, either of these measures will in fact be influenced by the trends in crime. Therefore, we also estimated models in which both of these explanatory variables are omitted as not being truly exogenous controls.⁴⁷

⁴⁴ A minor variable that attempts to measure the amount of income maintenance per capita given to those over age 65 under certain social welfare programs is problematic in that the reason for inclusion is uncertain and attempts to replicate and extend this data beyond 1992 to 1997 were unsuccessful. Since its inclusion or omission matters little to the analysis before 1992, we just excluded it in all reported regressions.

⁴⁵ Mustard has argued that this exclusion based on the undefined arrest rate in zero crime counties is actually a good thing: “Omitting arrest rates may generate a truncation problem because many counties with zero crime rates will be included in the regression. By construction it is impossible for a Shall-Issue law to reduce crime in a county that has no crime, no matter how effective is the law” Comments at Brookings Conference. We disagree with this point. If a county has no crime, then adopting a shall-issue law can have no crime-reducing benefit. If one is interested in ascertaining what the average impact on crime is from the adoption of the law, one needs to include all jurisdictions, including those that derive no benefit. To take the extreme case, assume that the vast majority of counties had no crime, and a few had substantial crime that dropped by 10 percent when a shall-issue law was adopted (while by assumption there was no benefit when such laws were adopted in the zero crime areas). Lott and Mustard would argue that shall-issue laws cause crime to fall by 10 percent, even though the true drop in crime in all the adopting areas would only be a miniscule fraction of that figure.

⁴⁶ The disadvantage of the incarceration rate data is that it is only available on a statewide (as opposed to countywide) basis, and therefore may be less useful in the county data analysis. In part for this reason, we primarily rely on the incarceration rate when conducting analyses on state data and the arrest rate when conducting analyses on county data.

⁴⁷ In general, as indicated in our discussions of Tables 5 and 6 below, the only explanatory variables that tended to have a large impact on the estimated effect of shall-issue laws were the demographic variables.

B. Comparing the Results Using the Zheng and Lott Controls

Our parsimonious Table 1 regressions for the period 1977-1999 confirmed what the graphs of Figure 1 had shown -- that shall-issue laws were associated with statistically significant higher crime rates in eight of the nine crime categories. We can now enrich the simple model by adding controls for other factors that might be influencing crime over this period. There is a vast array of different variables that might be entered into a crime equation, and we begin by using the specification employed by a gifted researcher at Stanford, named Wentong Zheng, who has been working with state crime data to ascertain the impact of state lotteries on crime. Zheng generously shared his data set and programs and coded in the states that adopted shall-issue laws. The resulting regressions control for the following state-wide attributes: population, the percentage of black population; the percentage of metropolitan population; the percentage of population aged 15 to 17, 18 to 24, and 25 to 34 respectively; per capita income; unemployment rate; the percentage of population under the poverty line; number of prisoners in state prisons (once-lagged); number of sworn police officers (once-lagged); and per capita alcohol consumption.⁴⁸ While one might argue that some different explanatory variables should be included or that the included one's should be treated differently, the advantage of using this specification is that it was adopted by a different researcher for a different crime paper and therefore it is immune to the argument that the explanatory variables were chosen to influence the estimated effect of shall-issue laws.

The results of this analysis are shown in Table 2. Let's begin with line 1, which is the "dummy variable model," in which the regression estimates the change in crime from the pre-passage to the post-passage period while controlling for the influence of Zheng's explanatory variables. Comparing the results in line 1 to the simple regressions of Table 1, line 2, we see that the general story that, over the 1977-99 time period, shall-issue laws are associated with *higher* rates of crime comes through, although less powerfully than before and primarily for robbery and property crimes.⁴⁹ While before, 8 of nine estimates were positive and statistically significant, now seven of nine are positive, of which only five are statistically significant. The two negative effects -- for murder and rape -- are statistically insignificant.

Lines 2 and 3 of Table 2a provide alternative methods for estimating the effect of the shall-issue laws by testing whether there is any break in the trend of crime (line 2) or allowing both a change in the level as well as a change in the trend in crime (line 3). The line 2 results essentially wipe out any effect, leaving only one statistically significant variable -- robbery, which switched in sign from positive to negative in going from line 1

⁴⁸ Summary statistics for these variables are presented in Appendix Table 2.

⁴⁹ The results shown in Table 2 use the Vernick coding. Wentong Zheng also ran similar regressions in which he used the Lott coding for the period of his original data set (1977-92) and the Vernick coding for 1993-99. This different coding led to similar results, except that murder became significant and negative.

to line 2. When both effects are estimated in the line 3 hybrid model, they become somewhat anomalous, showing a positive main effect in most of the crime categories, but some negative trend effects.⁵⁰

Table 2b adopts the same estimation approach as in Table 2a, with one change. Since the panel data model is essentially a comparison of the period before and after adoption of the shall-issue law, there is some danger that if crime is trending up (or down), the Table 2a regressions might erroneously attribute a pre-existing trend to the adoption of a law. Consequently, the Table 2b results control for any pre-existing state trends. The results are similar to those in Table 2a, but the anomalous pattern of strong positive main effects and negative trend effects in the hybrid model is eliminated. Importantly, the robbery effect comes through as positive and statistically significant in both the dummy variable model and the hybrid model when one controls for state fixed trends.

So while our simplest Table 1 regressions (with controls for only state and year fixed-effects) suggested that shall-issue laws substantially increase crime, moving to a full regression model in Table 2, we see that the results become less stark – adding Wentong Zheng’s set of explanatory variables has generally maintained the story that robbery and property crime rates are higher with adoption of shall-issue laws, while the results for the four other violent crimes are more mixed across the five different models shown. If one adopts the view that the results should be robust across both Tables 2a and 2b before they can be accepted, one would probably conclude that shall-issue adoption *may* be associated with lower amounts of rape – the hybrid results with state trends are ambiguous – but there is no consistent evidence of lower crime in any other crime category. At the same time, there is more robust statistical evidence of higher property crime, although the theoretical rationale for such an effect is obscure.

It is now worth exploring whether the Lott and Mustard models (which differ from Zheng’s only in terms of the explanatory variables used) will generate similar results. Therefore, we ran our version of the Lott model (which includes a host of additional right-hand side variables, using incarceration rates instead of arrest rates) on this same expanded state data set covering the years 1977-1999. Somewhat surprisingly, the results (shown in Table 3) appear far stronger for the Lott and Mustard thesis than those depicted in Table 2.⁵¹ For example, while the dummy variable model of Table 2 showed 5 statistically significant positive variables (and none that were negative), the same model in Table 3 has two positive and two negative estimates that are statistically significant. Indeed, in the hybrid results of Table 3 almost all the positive and significant coefficients from Table 2 disappear. Again, using the standard of robustness across models, the strongest results are that rape goes down and larceny goes up. Despite the

⁵⁰ We previously indicated that such a pattern was likely to be a sign of model misspecification. See TAN 41.

⁵¹ To ensure that the different shall-issue law coding schemes mentioned in footnote 24 are not driving the results, Appendix Table 3 replicates the Table 3 results using the Lott shall-issue law coding. The results are substantially similar.

anomaly of the initial positive effect followed by a downward trend in the hybrid model of Table 3b, the Lott model also suggests that murder falls with shall-issue adoption. While Table 3 provides the first support for some drops in violent crime (for rape and murder), it still gives no consistent evidence of any decline in robbery, which is the one area in which one would most expect the deterrent impact proposed by Lott and Mustard to show up if it were in fact important.

The question merits investigation: why do the results of the Lott model (Table 3) support the Lott thesis more than the results of the Zheng model (Table 2)? The reason turns out to be somewhat surprising. As Appendix Table 2 documents, the two models include some different explanatory variables that one might think could have important implications – for example, Lott controls for population density and transfer payments to the poor, while Zheng controls for police, poverty, unemployment, and alcohol consumption. But these differences in the substantive controls turn out to be largely unimportant. Interestingly, as we will discuss in the next section, what drives the entire difference between Tables 2 and 3 is that Lott controls for a far richer array of demographic variables. Indeed, the array is so rich as to make multicollinearity a serious issue. Specifically, while Zheng’s model controls for percent black and three age groupings, Lott has 36 separate demographic percentages, breaking down each of three different race categories – black, white, and neither black nor white – and both sexes into six separate age categories from age 10 up. The sensitivity of the results to the inclusion or exclusion of an array of highly collinear demographic variables serves as a cautionary tale to those who conduct or rely upon panel data models of crime. Probably no one examining either Wentong Zheng’s own work or that of Lott and Mustard would suspect that conclusions reached from their models would be sensitive to these seemingly second-order demographic controls.

C. Sensitivity of the Lott Results to Time Period and Inclusion of Demographic Controls

The results so far suggest that both a sparse model with no controls other than state and year fixed effects, as well as a reasonable specification developed for another context (thereby eliminating the possibility of data mining to achieve a desired result) undermine the Lott and Mustard more guns, less crime hypothesis. At the same time, though, their own specification generates some superficially supportive results in that murder and rape seem to be lower after passage. We say “superficially supportive” because a number of factors, including the failure of the model to show a drop in robbery, casts doubt on the causal story that they advance.

The first problem with the Lott story – as suggested in various graphs of Figure 1 – is that during the 1990s crime in non-shall-issue states fell far more than in shall-issue states. Indeed, when we limit our focus to the 13 states that adopted shall-issue laws between 1991 and 1999 as shown in Table 4, we see that the Lott story does not hold up

during this period. The only robust finding we can see in Table 4 across both panels is that three of the four property crime categories are higher with shall-issue law adoption.⁵²

We have already alluded to the second problem with the Lott story: the results are incredibly sensitive to the inclusion of various seemingly unimportant demographic controls. The first line of Table 5a shows the results from estimating a dummy variable model using all of Lott's explanatory variables (again replacing arrest rate with the incarceration rate), except his large array of demographic controls. Here the results look very much like Table 1, line 2 in showing that shall-issue laws are associated with (statistically significantly) *higher* rates of crime in every category but murder. In other words, the similarity of the Table 5a, line 1 results and the Table 1, line 2 results suggests that the addition of the incarceration rate, state population and density, the controls for transfer payments, and real per capita state income essentially had no effect on the story depicted in the various graphs of Figure 1 (that shall-issue laws were associated with more crime). Each successive line of Table 5a augments the model estimated in line 1 by adding additional demographic controls. By the time the full array of demographic controls is entered, five of the eight strong positive estimates of line 1 have flipped sign (although only violent crime and rape have become statistically significant in the opposite direction). Even the crime categories, such as property, auto theft, and larceny, which retain their positive signs, still drop very sharply in size. For example, the estimated increase in auto theft from adopting a shall-issue law is 24.1 percent in line 1 of Table 5a, but only 4.5 percent in line 9.

What explains the enormous sensitivity of Lott's results to the inclusion of minor demographic controls? Table 5b gives insight into this question by replicating the analysis of Table 5a, while controlling for fixed state trends. Now we see much closer agreement with the results as we move down the nine regressions listed. Apparently, then, Lott and Mustard's 36 demographic variables mimic time trends in crime that we can control for directly with our controls for state trends.

Assume, though, that Lott's decision to include 36 demographic controls is excessive and that the appropriate number of these highly collinear variables is some smaller number. For example, since men commit over 90 percent of the violent crime, perhaps it would be enough to control only for the percent of white males in each of six separate age groupings above 10 as well as the percentage of black males in each of these six age categories, for a total of 12 demographic controls. One could follow Lott in including additional controls for the percentage of white or of black females in each of these six age groupings for an additional 12 more controls, but each black male category is already highly correlated with each corresponding black female category and the same correlation applies for whites as well. One could go even further in following Lott and put in the percentage of the population that is neither white nor black in 12 other age-sex categories, but again it would not be clear a priori why one would do that.

⁵² We also replicated this table using the alternative coding of the shall-issue laws shown in Appendix Table 1, column A. The resulting Appendix Table 4 yields quite similar results, except that rape is robustly negative in the two panels.

Table 6 mimics the Lott and Mustard model of Table 3, but limits the demographic controls from Lott's 36 to the six age categories above age 10 for white males and the same six age categories for black males – a total of 12 controls. Now it is hard to find *any* crime category that seems to have a robustly lower crime rate across the various models of panel A and panel B. Larceny and property crime seem to be robustly higher with shall-issue adoption and rape may be lower (although the results including state trends are mixed for this crime). Overall, the results seem much more supportive of a positive impact on crime of shall-issue laws than a negative one. Perhaps one can argue that it is important to include all of the other 24 highly collinear demographic variables to return to the (already weak) results of Table 3, but results that are sensitive to inclusion or exclusion of these marginal control variables must be treated with skepticism.

Since the Figure 1 graphs, the Table 1 regressions, the Table 5 regressions with the non-demographic controls, and the Table 6 regressions with 12 of Lott's demographic controls are all showing a largely *positive* relationship between shall-issue law adoption and crime over the 1977-99 period, it becomes important to ascertain what about the demographic controls is capturing some true phenomenon that would reverse our assessment of the crime effects generated by shall-issue laws. Ordinarily, if the inclusion of an explanatory variable has such a dramatic effect on one's estimate, one would explore whether the estimated effect of that explanatory variable makes sense. In this case, though, the 36 demographic controls are so highly collinear that it is impossible for the regression to provide meaningful results for any given demographic control.⁵³ Thus, all we are left with is a set of estimated effects on the demographic controls resembling output from a random number generator, and the knowledge that for some unknown reason these demographic controls matter considerably. We can think of no reason to prefer the regressions in Table 3, which might be deemed to provide weak support to Lott

⁵³ For example, if one regresses the percent black age 10 to 19 on the other demographic controls, the R-squared value is .9999. Consequently, there is a very severe multicollinearity problem here that makes it impossible to interpret any of the coefficients on the individual demographic controls (although ordinarily one wouldn't expect this to invalidate the estimated shall-issue effects). Indeed, when one looks at the estimated effects of a change in the demographic variables on violent crime from the Table 3 regressions, one finds utterly bizarre results. For example, an increase of one percentage point in the percentage of black males aged 30-39 would be expected to almost double the violent crime rate, while a similar increase in the percentage of black males aged 40-49 would lead to a drop in violent crime of 60 percent. Similarly, increasing the percentage of black males aged 50-64 would cause violent crime to jump by 145 percent, but increasing the percentage of black males over age 65 would lead to a 78 percent decline in violent crime. These nonsense results prevent us from understanding why the demographic controls can influence the estimates of shall-issue adoption so strongly.

and Mustard, to those in Table 6, which are devastating to the more guns, less crime thesis.⁵⁴

While virtually the only evidence shown so far that offers even weak support for the Lott and Mustard thesis is Table 3, there is a statistical argument that even these results inaccurately over-state their statistical significance. A recent comment by Willard Manning of the University of Chicago suggests that when, as here, one is estimating nine different regressions, one should make adjustment for the fact that multiple comparisons are being made in assessing statistical significance:

“Failure to correct for multiple comparisons causes the true significance level to be much less than the nominal level would suggest. If there are seven comparisons, then a nominal five percent standard applied to each is actually more like a thirty percent standard. The former is usually considered statistically significant if it is met, while the latter is considered statistically insignificant, and not noteworthy unless one is looking for a null finding.”⁵⁵

Correcting for the presence of these multiple comparisons and for autocorrelation in crime across years, Manning found that the estimates of the effect of shall-issue laws “would have t statistics that are about 50-60 percent of their reported values. With such a correction, the results [estimating the effect of a shall-issue law] appear to be statistically significantly different from zero about as often as one would expect if they occurred at random.” Again, this is not a ringing endorsement of the more guns, less crime hypothesis. If one were to apply such a correction to the Table 3 results, there would simply be no evidence of any crime-reducing benefit associated with the adoption of shall-issue laws.

D. Problems With Unequal Years of Data From Early and Late Adopters in the Pre-Post Comparison

Thus far our regressions have attempted to generate an estimate for the impact of shall-issue laws that are aggregated across all adopting states. Essentially, we are implementing a difference-in-difference design, in which we compare how crime changes over time in states that do not have shall-issue laws (or that had such laws prior to our data period) to how crime changes in states that adopt shall-issue laws during our data

⁵⁴ In response to the claim that the unpredictable estimated impacts of the demographic controls suggest model mis-specification, Lott explains that “What I have tried to do is ‘overcontrol’ for all possible demographic factors to make sure that any effects attributed to the right-to-carry laws are not arising because I have accidentally left out some other factor.” Lott (2000: 144). Ironically, the effects that Lott attributes to these laws occur not because he has left out any demographic variable, but primarily because he has put them in.

⁵⁵ Willard Manning Comment, Brookings Conference on Guns.

period. An important aspect of this investigation is to select a ‘before’ and ‘after’ period for pre-post comparisons, and thus far we have used the entire 1977 – 1999 panel to calculate the aggregated pre-post difference for shall-issue states. This can create problems, though, because shall-issue laws were adopted as early as 1985 and as late as 1996, which implies that for some states, observations from 19 years before adoption to 18 years after adoption would be used to form a pre-post contrast. This can be potentially problematic when a very selected sample at the two ends of the overall time span may be tilting the regression line unduly. Table 7 shows the periods of time for which pre- and post-passage data is available for the 25 adopting jurisdictions (in the Vernick coding). It should be immediately apparent that while all 25 states are used in identifying the effect on crime over the period from 8 years prior to three years after passage, only three states are entering into the estimate of crime 19 years before passage Texas, Kentucky, and South Carolina. Thus, Texas will disproportionately influence the population weighted pre-passage because it will be counted in 11 more yearly observations than the Maine observations are counted in the pre-passage period. Conversely, in the post-passage period, Florida, Georgia, and Pennsylvania, all fairly populous states, will influence the post-passage estimates more highly both by virtue of their substantial populations and because they will have from 10-12 yearly observations that will be averaged into the post-passage effect compared to the mere three years of observations appearing for Kentucky, Louisiana, and South Carolina.

One crude way to show that this factor is important is to simply “dummy out” the periods more than eight years prior to or more than three years after the passage of a shall-issue law.⁵⁶ This approach essentially will enable the estimates of the aggregated effect of shall-issue laws on crime to come primarily from the twelve-year span that begins 8 years prior to passage and extends through the year of passage to the three years after passage. Table 8 shows the results of this exercise. The bottom line seems to be that rape may be lower with shall-issue laws (the results are somewhat mixed across the two panels), but larceny and burglary are robustly larger. Indeed, if one accepts the estimates from the hybrid model controlling for state trends, shall-issue laws would be deemed (at least by the third year) to have *increased* violent crime, aggravated assault, robbery, property crime, burglary, and larceny. There is one negative estimate (rape), but its effect is not statistically significant in the third year and it turns positive in the fourth.⁵⁷

To underscore the message that the observations from the adopting states at the two ends of the time span – either well before passage or well after passage – are causing mischief when estimating a single aggregated impact of the shall-issue laws, it may be

⁵⁶ By creating individual dummy variables for the stated years before and after adoption, we are ensuring that our overall estimates of the impact of the law will not be influenced by observations from those years.

⁵⁷ Appendix Table 5 replicates the Table 8 results the Lott shall-issue law coding. The results are similar, but with two changes: the Appendix Table 5 results weaken the Lott and Mustard thesis since robbery is *positive* and statistically significant across both panels (as opposed to just in Panel B of Table 8), but they strengthen it in that rape is more robustly negative.

helpful to illustrate this point graphically. Figures 3a-i plot the estimated annual effect on crime for every year from 19 years prior to adoption to 17 years after adoption.⁵⁸ Figure 3b shows the results for murder and it should immediately be apparent that the period from 8 years prior to 3 years post-passage evidences relatively little movement in this crime category (as the Table 8 regression results would suggest). However, outside this time frame (at either end), one sees large swings in the estimated effect of shall-issue law on murder. Of course, the thought that shall-issue laws caused crime to drop by almost 25 percent in the thirteenth year after passage and then caused it to increase by 3 percent in the fourteenth year is obviously untenable. These wild swings are caused not by any true impact of shall-issue laws but by the selective dropping out from the individual year estimates of states that adopted the law more recently, leaving only the shrinking number of earliest adopters to identify the particular annual impact. Whether one looks at murder, aggravated assault, rape, or robbery, the final year estimate at fourteen years after passage always swings wildly – sometimes up, sometimes down – and the earliest pre-passage estimates also show erratic movements (note the jump in the first year for aggravated assault). Another way to get at this same problem of unbalanced panel data is simply to undertake a disaggregated analysis – which estimates different shall-issue impacts for individual states. We will undertake just this analysis when we later turn to analyze the county data, but first we will address an issue that we previously set aside.

E. Problems With Endogenous State Adoption

A major concern of any analysis on the effect of shall-issue laws on crime is the possibility that crime correlates with shall adoption in a non-causal manner. Indeed, if we examine the graphs of Figure 3, one can see that, in four of the five violent crime categories and for burglary, even before adopting states passed their shall-issue legislation, crime was substantially higher than the regression model would have otherwise predicted (given the full array of explanatory variables).⁵⁹ This raises concerns about the reliability of the regression model, because the higher crime rate in the pre-passage years may suggest that the laws were passed because crime was higher. That would imply that shall-issue law adoption is, in fact, endogenous to the crime rate. Lott is aware of this problem and indeed confirms it in his book, noting that shall-issue laws

⁵⁸ In other words, rather than simply having a single post-passage dummy as in our customary dummy variable model, we now define a separate dummy variable for each individual year before or after adoption of a shall-issue law, and then plot these individual estimates in the various graphs of Figure 3.

⁵⁹ Indeed, the evidence concerning violent crime presented in Figure 3a suggests that violent crime was 30 to 40 percent higher than one would have otherwise expected given the full array of control variables. On the one hand, this may suggest that shall-issue laws are adopted when violent crime is high, thereby complicating efforts to determine the effect of the law (since a return to more usual crime rates after passage may erroneously be attributed to the law). On the other, this pattern may suggest that the crime model is not working very well prior to passage for these states, and therefore we should be concerned that it will not work well after passage either.

“have so far been adopted by relatively low-crime states in which the crime rate is rising” (Lott 2000). He attempts to use the appropriate two-stage least squares (2SLS) technique to address the problem, and he states that his “results differ from all my previous estimates in one important respect: nondiscretionary concealed-handgun laws are associated with large, significant declines in all nine categories.”⁶⁰

As Willard Manning and others have emphasized, the problem is that the statistical demands for the successful use of the 2SLS approach are exacting, and it is unlikely that Lott has satisfied them since his instruments cannot be considered convincingly exogenous.⁶¹ Lott creates his instrumental variable by regressing the presence of a shall-issue law on violent and property crime rates and the change in those rates, percent of state population in the National Rifle Association (NRA), percent of state population voting for Republican presidential candidate, percent of blacks and whites in state population, total state population, dummies for the South, Northeast and Midwest, and year dummies (Lott 2000).

The manner in which he performed his 2SLS analysis suggests that he considers his main instruments to be the violent and property crime rates and the change in those crime rates. However, these variables are obviously endogenous and cannot serve as unbiased instruments. His NRA and Republican presidential election variables are the most compelling, but these too are at least potentially endogenous. These variables correlate with factors such as an area’s wealth and criminal tendencies, which in turn affect crime directly. We attempted to replicate Lott’s 2SLS results using state-level data to determine how much of an effect the first-stage crime rate variables were having on the final results.⁶²

⁶⁰ Lott (2000) at 118.

⁶¹ As Manning notes, “the major requirements for the instrumental variables in the linear model to yield consistent estimates of the effect of the endogenous explanatory variable on the outcome of interest are: (1) the instruments are correlated with the endogenous explanatory variable(s); (2) the instruments do not conceptually belong in the equation of interest nor are they proxies for variables which should be in the equation of interest but are omitted from the specification; (3) the instrument is uncorrelated with the error term in the equation of interest; and (4) the instruments are not weak in the sense of Staiger and Stock (1997) or Bound et al. (1995)” Manning Comment.

⁶² While it is unclear exactly how Lott performed these regressions, it appears that he instrumented the shall-issue law dummy using only the variables mentioned above, deviating from standard 2SLS procedure by failing to include all of the exogenous variables used in the second stage of the regressions. If he did include all the exogenous variables, then one would question introducing the regional dummy variables and the racial demographic percentages into the first-stage equation given the existing state dummy variables and extensive demographic controls. However, since excluding the exogenous variables from the first-stage regression can possibly bias the estimates, we avoided replicating this potential inaccuracy. Thus, there are a few differences between our regressions and Lott’s. First, because we were unable to obtain NRA membership data, we were forced to leave this variable out. Fortunately, as we were able to roughly approximate Lott’s results, this omission proved not to be a substantial problem. Second,

Table 9, line 1 presents our replication of Lott's 2SLS estimates for the 1977-1992 time period.⁶³ According to this method, shall-issue law adoption reduces violent crimes by 61 percent, murders by 43 percent, rapes by 20 percent, robberies by 51 percent, and aggravated assaults by 64 percent.⁶⁴ While these numbers should seem implausibly large, we were gratified to see that we had basically succeeded in replicating Lott's 2SLS results. While the excessively large predicted effects have led a number of researchers to dismiss these results immediately, we go on to show that the results are entirely driven by the inclusion of the inappropriate instruments -- the logs of violent and property crime. Line 2 of Table 9 presents the results of the identical 2SLS regressions of line 1, except excluding the logs of crime rates as instruments. Notice that some of the most extreme negative coefficients from line 1 drop drastically. For example, the coefficient estimated for violent crime drops from -.94 to -.09, and that for aggravated assault drops from -1.18 to -.12. The dramatic changes in the 2SLS estimates when the most flawed instruments are dropped suggests, as we suggested above, that this endogenous set of instruments was severely biasing the results. We then went one step further and performed another set of regressions that excluded the change in the crime rates as instruments, thereby leaving only the matrix of political variables as the remaining instrument. As shown in line 3 of Table 9, this approach generated estimates that shall-issue laws would strongly dampen murder, robbery, and aggravated assault.

The next three lines of Table 9 replicate these three 2SLS regressions for the 1977-1997 period. While the Lott approach of line 4 continues to show large drops in violent crime, rape and aggravated assault, the line 4 numbers are much less favorable to the Lott and Mustard thesis than the line 1 estimates for murder, robbery, and all four property crimes. But while the drops in violent crime and aggravated assault resulting from the shall-issue laws that are estimated in line 4 are once again vastly too large to be plausible -- the second and third lines that drop clearly inappropriate instruments alter the results dramatically. Indeed, when the flawed instruments are dropped in lines 5 and 6, the estimated effect on crime of shall-issue law adoption is never significantly different

instead of including the percentage of blacks and whites in the state population and the region dummies, we used the array of demographic variables and state dummies that Lott used in the second stage of his 2SLS regressions. Accordingly, our first stage consisted of regressing the shall-issue law dummy on the vector of demographic variables, the log of violent and property crime rates, the changes in the log of violent and property crime rates, and a vector of presidential election variables. This final vector was created in the exact way Lott created his: interacting the percentage of state population voting Republican with dummy variables for a four-year period around the election year. Since our effort is at replication, we also used the Lott coding of the shall-issue laws.

⁶³ We use Lott coding of the shall-issue laws in an effort to replicate the Lott and Mustard approach as closely as possible.

⁶⁴ The precise estimated percentage effect on crime is given by the formula $e^B - 1$ where B is estimated in Table 9. For values under about .2 in absolute value, B is a fairly good approximation of the proportionate change.

from zero. Accordingly, the evidence from 2SLS estimation, in our view, offers no support for the more guns, less crime hypothesis.⁶⁵

III. The Basic County Data Findings Of Lott and Mustard

Up to this point we have focused exclusively on analyzing state data, which enabled us to use more current crime data – up through 1999. Lott and Mustard decided, however, to rest their analysis most heavily on county crime data rather than state crime data (although they do present some state data results as well). The advantages of this approach are that (1) with over 3000 counties in the country as opposed to only 50 states, there is far more data with a county data set than with a state data set, which, *ceteris paribus*, should improve the precision of the estimates; and (2) the county fixed effects will explain a great deal of the fixed cross-sectional variation in crime across the country, thereby diminishing the inevitable problem of omitting some appropriate, but possibly unavailable, time-invariant explanatory variables.

There are also some disadvantages to using county data. First, the intervention of interest is generally a statewide phenomenon.⁶⁶ Using county data under these circumstances will likely exaggerate the amount of independent data available to the researcher, thereby possibly creating the appearance of statistical significance when in fact none exists.⁶⁷ Second, many of the explanatory variables are only measured on the state level and thus again the county data analysis may be giving a false sense of precision. Third, Lott uses arrest rates (the ratio of arrests to crime in a county) as an explanatory variable, which leads to many counties being dropped from the analysis. This occurs because of missing arrest data in some cases, and also because the arrest rate is undefined for any county that experiences no crime in a particular category in a particular year (since the rate would have zero in the denominator in such cases). Thus, a substantial number of counties are thrown out of the Lott analysis by virtue of the realization of the dependent variable (if it is zero in a given year, that county is dropped

⁶⁵ We are not claiming that the instruments used in our line 5 or 6 2SLS estimates meet the criteria for valid instruments set forth in footnote 60, *supra*. Indeed, the political variable serving as the instrument in lines 3 and 6 is quite weak, only raising the R-squared value in a first-stage regression by .006 (from .815 to .821). We are confident, however, that the other instruments used by Lott that are dropped in the third regression for each time period in Table 9 profoundly violate these standards.

⁶⁶ But not uniformly. Pennsylvania initially excluded Philadelphia from its 1989 shall-issue law. In 1995, the law was extended to include Philadelphia.

⁶⁷ Indeed, when we reran both Lott's and our own specifications clustering on the state level we found dramatically lower statistical significance for the estimated impacts. The cluster procedure allows the regression to relate the variability across counties within a state for more refined estimation of the variance-covariance matrix. See *Stata User's Guide* (Release 6, 258-9). However, to be both conservative and consistent with Lott, all the regressions reported in this paper do not cluster by state.

from the analysis), which can potentially bias the results of the regression estimation.⁶⁸ Fourth, Maltz and Targonski consider the quality of UCR county-level data to be so poor that they dismiss Lott's work on that basis alone (at least if the data extends beyond 1992).⁶⁹

By using our state data, we were able to extend our analysis through 1999, while our county data set is only updated through 1997. But because there are some advantages to using county data and in a sense to take on Lott and Mustard on their own ground, we turn for the rest of this article to an analysis of county data.

A. Lott's County Data Analysis for 1977-92

1. The Dummy Variable Model

We begin by presenting Lott's initial regression results for county level data over the period from 1977-92. Line 1 of Table 10 shows the predicted effect on nine crime categories using the dummy variable model. A quick examination of the line 1 results reveals 1) four of the five categories of violent crime (the exception is robbery) have negative and statistically significant coefficients, suggesting that shall-issue laws reduce these types of violent crime by 4-7 percent; and 2) all four property crimes have positive and statistically significant coefficients, suggesting that the laws increase property crime by 2-9 percent. Lott accepts the regression results at face value and concludes that the passage of these laws causes criminals to shift from committing violent crime to committing property crime where, he argues, they are less likely to be shot since the victim is frequently not present when the crime occurs. Thus, we see violent crime decreasing by 3.5 percent and murders falling by over twice that percentage, while property crime rises by over 5 percent. As Ayres and Donohue (1999) stressed, however, the fact that robbery is not dampened by the adoption of a shall-issue law constitutes a major theoretical problem for Lott's interpretation of the results of the dummy variable model. If there is to be the type of substitution away from violent crime that Lott predicts, one would expect that the new law would induce potential robbers to avoid confronting victims and shift to more stealthy property crime. Yet, as Table 10 reveals, we see no evidence of this effect. Hence, the dummy variable model undermines a key prediction that Lott offers to explain the line 1 regression results for the period 1977-92.

⁶⁸ The percentage of dropped observations (because of either missing numerators or zero denominators) vary for the individual crime categories – from a low of about 9% for auto-theft to a high of about 48% for murder. 57% of county/year observations have at least one of the arrest rates undefined or missing.

⁶⁹ Maltz, Michael, and Targonski, Joseph, "A Note on the Use of County-Level UCR Data," 2001. Although in this portion of the article we focus on county data for the period 1977-97, we generally find, with some noted exceptions, broadly similar results when we either use the state data set or confine the county results to the period not beyond 1992. We also note that to the extent – however unlikely -- that the post-1992 break in the county crime data series to which Maltz and Targonski refer is relatively uniform across all counties, the year effects would control for this change. See footnote 73, *infra*.

In Table 4.1 of his book, Lott presents a version of the line 1 robbery regression showing that shall-issue laws reduce robbery by 2.2 percent, which is indicated to be statistically significant at the .10 level (considered marginally significant). But Donohue and Ayres (1999) reveal that this –2.2 percent figure is an error that results from a miscoding of the effect of the shall-issue laws. The problem was that, instead of following his articulated strategy of assuming that the effect of the law would emerge in the first year after passage, Lott coded the shall-issue law in that fashion only for Florida and Georgia, with all other states being coded so that the effect of the law begins in the year of passage. Correcting this error to adhere consistently to the articulated Lott protocol wipes out the size and significance of the estimated effect on robbery.⁷⁰

2. Lott's Trend (or Spline) Model

Lott responds to the point that shall-issue laws seem not to dampen robberies in the dummy variable model by arguing that a model that captures the change in the linear trend of crime – the Lott spline model -- provides a better picture of the effect of the passage of a shall-issue law in this case. The only numbers that Lott reports in his book concerning his trend analysis are found in a single row of numbers representing the difference between the before-passage linear trend and after-passage linear trend for the states that passed shall-issue laws (appearing in the book's Table 4.8). Lott's regressions include year effect dummies, so the pre-and post-passage trend coefficients would capture linear movements in crime in the 10 passing states apart from the general movements in crime for the nation as a whole (which would be captured by the general year dummies). Lott's message in his Table 4.8 is that a trend analysis shows that shall-issue laws lower all crime categories – both violent and property – and in all cases but one (larceny) the reduction is statistically significant.

Line 2 of Table 10 follows Lott's lead in testing whether there is a break in the linear crime trend in the year of passage.⁷¹ We use a spline specification that is

⁷⁰ Ayres and Donohue (1999) replicate Lott precisely with the coding error and then show how the correction eliminates the robbery effect. The line 1 regressions in Table 10 of this paper are identical to Lott's Table 4.1 results with three exceptions, which are maintained in all the regressions presented here: 1) the coding error is corrected, 2) standard errors are corrected to adjust for heteroscedasticity, and 3) one explanatory variable – the measure of the real per capita income maintenance, SSI and other, for those over 65 – was dropped. One can compare the results in Table 1 of Ayres and Donohue (1999) with those of Table 10 here to see that the only one of these changes that influences the basic story is the correction for the coding error. As noted in Section II(D) above, the explanatory variable of real per capita income maintenance for the elderly has been omitted in all our regressions.

⁷¹ Panel A2 of Appendix Table 6 reproduces the methodology (although not the reported results) of the Lott trend analysis and reports both the before and after linear time trends (as well as their difference, which is taken as the estimated effect of the law). The discrepancies between Lott's Table 4.8 and Appendix Table 6.A2 result from two different Lott errors. First, Lott has informed us that he mistakenly wrote down the "Shall Trend After" coefficient for violent crimes instead of reporting the difference between the before and after coefficients. Second, Lott has erred in his reporting of the

somewhat easier to interpret (than Lott's difference in before and after trend test) because the statistical significance of any induced change in trend can be ascertained directly from the regression output (the t-statistic for the post-passage linear trend coefficient). Even though he did not use a spline himself, we call this the Lott-spline model because it is equivalent to the linear trend model that Lott did employ in which he estimates pre- and post-passage trends and then subtracts them.

Note that the story in line 2 of Table 10 is changed in a number of respects from that of line 1 (the dummy variable model). Instead of all violent crime (but robbery) falling and property crime rising, line 2 suggests that shall-issue laws have no effect on property crime (or overall violent crime and aggravated assault), but dampen murder, rape, and the heretofore unaffected robbery. Lott's discussion of the impact of shall-issue laws causing criminals to shift from committing violent to committing property crime is no longer central if the Lott spline analysis (regression 2 in Table 10) is the appropriate estimation approach.

3. The Hybrid Model Testing for Main and Trend Effects

The different results between the dummy and trend models suggests the advisability of employing the more general hybrid model, which will enable us to test whether either of the more constrained models is statistically preferable. Consequently, we estimate regression 3 in Table 10, which is a less constrained specification than either the dummy variable or the linear specification because it allows (and tests) for the existence of both a once-and-for-all announcement effect as well as a changed (linear) trend effect.⁷² But for the four violent crime categories, we see a pattern that is potentially problematic – the main effect of the shall-issue laws is positive but over time this effect gets overwhelmed as the linear trend turns crime down. In other words, according to the hybrid model, in the year after passage the main effect of the shall-issue

statistical significance of these effects, perhaps because he did not correct the standard errors for the presence of heteroscedasticity in the panel data. Lott reported that the change in trend owing to the shall-issue law for violent crime, aggravated assault, and auto theft was statistically significant at the level shown in the top row of Appendix Table 6. But Appendix Table 6.A2 shows, however, that the differences for these three crime categories were all insignificant.

This Appendix Table 6.A2 replication of Lott still ostensibly finds that the law generates a statistically significant reduction in the time trends (at, at least, the 5% significance level) in 5 of the 9 crime categories tested. But Lott's regressions also incorrectly identify the passage date of three jurisdictions that adopted shall-issue laws, which makes the laws look more effective than they are. Lott coded the enactment dates in Oregon, Pennsylvania, and Virginia earlier than was proper. Appendix Table 1 shows discrepancies Lott's shall-issue coding in his own dummy variable and trend models (columns B and C). The corrected numbers are presented in Appendix Table 6.A3, which shows that the shall-issue laws statistically significantly reduce crime in only three of the nine categories (murder, rape and robbery).

⁷² Contrary to the results of regression 1, the hybrid model confirms the prediction of regression 2 that the shall-issue laws have virtually no effect on any of the property crime categories (although we will soon see that this finding breaks down when we extend the analysis through 1997 in regression 6 of Table 10).

law is a 6.7 percent increase in violent crime, which is dampened by the 2 percent drop associated with the negative trend variable, for a net effect of 4.7 percent higher crime. After 3.5 years, the conflicting effects cancel out at which point crime begins to fall. Our antennae are raised when we see this particular result of a positive main effect and a negative trend effect since there are empirical reasons to believe that this observed pattern is not truly reflecting an initial acceleration in crime followed by a drop, but model mis-specification resulting from the selective dropping of states from the post-passage period in a way that tilts the trend regression line downward inappropriately.⁷³ This may suggest that the even the most general form of the three crime models is still mis-specified and hence that its results are unreliable.

B. Extending Lott's County Data Through 1997

Lott's initial analysis using 1977-92 data captured the period in which only 10 states newly adopted shall-issue statutes, and therefore Lott's regression results should be taken as the predicted effect of the adoption of the law in these ten states. Between 1992 and 1996, however, 14 more jurisdictions (13 states and Philadelphia) adopted the law, and therefore one might hope to gain more accurate results by extending the period over which the effect of the law is estimated. Regressions 4-6 in Table 10 simply repeat the models of regressions 1-3 but now estimating them over the longer time period from 1977-1997 (and thus measure the effect of adoption of the law in 23 states and the city of Philadelphia).⁷⁴ Comparing (the dummy variable model in) lines 1 and 4 of Table 10, we see that adding more years of data weakens Lott's story. Importantly, violent crime is no longer negative, so the basic story that the prospect of meeting armed resistance shifts criminals from violent crime to property crime is undermined. Lott might respond that murders fall by nearly 8 percent and rape by over 3 percent, as murderers and rapists shifted over to committing property crime, thereby raising its prevalence by 8 percent.

⁷³ As noted above, if the results had been flipped with the main effect dampening crime and the time trend suggesting a longer term increase, one could interpret those results in a straightforward manner: the announcement of the law scared potential criminals, thereby dampening crime initially, but as more guns got out on the street and/or as the fear subsided, crime ultimately turned up.

⁷⁴ In the second edition of his book, Lott analyzes 4 additional years of data that allow him to test the effects of shall-issue laws in 13 additional states. But he only reports results for this data set from tests of the trend specification. It is important to emphasize that combining these later years of data with the original data set is potentially problematic. The 1994 codebook for the (NACJD) crime data that both Lott and we use explicitly notes under a major heading, "Break in Series," that describes a new imputation procedure it will use from 1994 on and cautions:

These changes will result in a break in series from previous UCR county-level files. Consequently data from earlier year files should not be compared to data from 1994 and subsequent years

Michael D. Maltz & Joseph Targonski, A Note on the Use of County-Level UCR Data (unpublished manuscript 2000). If the break in series caused a uniform jump up or down in crime that applied to all jurisdictions, then our year dummies would control for this problem. Unfortunately, it is generally unlikely that errors in crime data would be uniform across the country (or even random across the country), so the break in the series is a concern. See *supra* note x (discussing newer data).

But the suggestion that this pattern could be explained by the changed behavior of would-be murderers and rapists is not compelling.⁷⁵

Indeed, the idea that a thwarted rapist would decide to switch to property crime because rape had become more dangerous (to the perpetrator) seems rather fanciful. The issue is important for the following reason. The dummy variable model regression on the full 1977-97 period gives very strong significant results on a number of crimes – murder and rape being negative and property crime, auto theft, and larceny being positive. Yet if the theoretical explanation for this substitution pattern in crime is flawed, then the fact that we cannot believe the regression finding of a large jump in property crime as an effect of the shall-issue law suggests we should not believe the accompanying regression finding of a substantial drop in murders and rapes. Again, the possibility of model misspecification seems to be a serious concern.

Interestingly, while adding five years of data weakens Lott's story based on the dummy model (line 1 versus line 4), it actually strengthens his story using the spline analysis (compare lines 2 and 5 in Table 10). Thus, we see in line 5 that every crime except property crime is both negative and significant. Moreover, in contrast to both dummy variable models, the Lott spline estimated effect for robbery for both time periods is negative and significant – an almost indispensable finding if the Lott deterrence story is in fact true.

The added five years of data also has a considerable impact on the estimates generated by the hybrid model. For Lott's initial time period, the hybrid model (line 3 of Table 10) basically suggested that shall-issue laws were not affecting property crime or robbery but were having the potentially problematic conflicting effects on four other violent crimes. For the full time period, however, the potentially troubling pattern previously observed for violent crime also shows up for property crime (line 6 of Table 10). For example, in the year after passage the line 6 regressions indicate that shall-issue laws increase murders by over 3 percent (the main effect of 6.9 percent less the trend effect of 3.5 percent). Rather than seeing shifting from violent to property crime, we observe similar patterns for a number of crimes in which early increases in crime are followed by subsequent drops after a number of years. The primary conclusion that emerges from these six regressions is that there is a considerable degree of instability in the predicted effects of shall-issue laws as one adds five extra years of data or switches among the two models presented by Lott or the third hybrid model which includes both a dummy variable and a time trend effect.

⁷⁵ Consider the case of Florida -- one of the states that statistically is most conducive to the Lott story in that murders fell after the passage of a shall-issue law in 1987. Assume that every murder and rape that would have occurred but for the shall-issue law would have been committed by a different individual (that is, there were no multiple offenses for either of these crimes). This would imply that a total of 282 individuals did not commit a murder or rape (106 fewer murders, 176 fewer rapes) because of the law but that they each committed over 240 property crimes instead, thereby elevating property crime by 68,590. If a number of these rapes and murders represented multiple crimes by the same individual, then the number of property crimes that would have had to have been committed per former rapist/murderer to generate this large property crime drop would rise commensurately. It seems unlikely that the shall-issue law could explain an increase in the amount of property of this magnitude.

Because the hybrid model nests both the dummy variable and the spline models, it is possible to test whether the data rejects the implicit constraints imposed by these more-restrictive specifications. Specifically, we can test the spline specification's implicit assumption that there is no announcement effect by looking to see whether the dummy variable coefficient in the hybrid model is statistically different than zero. The data strongly rejects the spline specification in 6 of the 9 hybrid regressions (in specification 6). The dummy variable model analogously assumes that there is no spline effect. This assumption is also rejected in 8 of the 9 hybrid regressions.⁷⁶

In sum, the foundation of the Lott thesis essentially is captured in regressions 1 (dummy variable model) and 2 (spline model) of Table 10. While these results are not identical to those presented in Lott's book, these regressions are probably more authoritative because some apparent coding errors by Lott have been corrected. The results are not as stable as one might like, but if one were to examine only those two regressions, the evidence would tend to support Lott's thesis. Obviously, the analyst's task would be easiest if the regressions generated by three different models (dummy, spline, hybrid), for three different time periods (1977-92 and 1977-97) all conveyed essentially the same picture. Unfortunately, they do not. Importantly, both the dummy variable and spline models are essentially rejected by the data by virtue of the large and statistically significant positive effects on both terms in the hybrid models (lines 3 and 6) – particularly for the full data set. But the hybrid model's prediction of initial jumps in crime followed by subsequent declines in response to the adoption of a shall-issue law raises our concern about model mis-specification as states that had adopted shall-issue laws close to the end of our data period drop from the estimates of the late post-passage effects.

C. Replicating Table 10 While Controlling for State Trends

We have previously discussed how one must take care to avoid generating statistically significant results that are really the spurious consequence of pre-existing trends (for example, where an upward trend in crime is confused for a post-passage increase in crime). Table 11 replicates the Table 10 results while controlling for these state trends. The first thing to note as we compare the Table 10 and 11 results for the dummy variable model in 1977-99 is that without controlling for state trends, the shall-issue laws were seen to increase crime in three property crime categories and decrease in two violent crime categories, while after adding such controls the only statistically

⁷⁶ It is interesting to compare the results of Table 3 (which uses state data for 1977-99 and the incarceration rate instead of the arrest rate) with those of Tables 10 and 11 (which use county data for 1977-97 and the arrest rate). Beginning with the dummy variable model and no state trends (Tables 3 and 10), the results are similar in each, although in the county data one sees stronger effects of dropping violent crime and increasing property crime. For the dummy model with state trends (Tables 3 and 11), though, the state results show drops in rape and aggravated assault and increases in auto theft and larceny, while the county results show only an *increase* in aggravated assault. When the hybrid model is used with state trends, the county results are clearly stronger for the Lott and Mustard thesis than the state data results – unless one is troubled by the positive main effects and negative trends found for the county data.

significant effect is that aggravated assault increases. The hybrid results for the same period tend to show positive main effects followed by negative trend effects, which raises the specification concerns we expressed earlier in our discussion of the end of period drops in estimated crime effects seen in the graphs of Figure 3. If we take that pattern as accurate, though, then after two to three years shall-issue laws would elevate murders, rapes, robberies, auto theft, and burglaries, and auto theft (while increasing aggravated assaults and larceny).

IV. Estimating State-Specific Passage Effects

On the surface, the panel data regressions in Table 10 (and Table 11 for the hybrid model) appear to establish a *prima facie* case that shall-issue laws reduce crime (or, at least in the dummy variable model, reduce violent crime while increasing property crime). But all of the regressions presented so far have estimated an aggregated effect for the laws across all adopting states. It is well known, however, that aggregation can at times lead to misleading conclusions in statistical studies.⁷⁷ For example, the model would be mis-specified if one tried to estimate a uniform effect from the shall-issue law while the law had systematically different effects across states. Moreover, as we have noted, the dangers of estimating a single aggregated effect are particularly acute in this case because a state that adopts a shall-issue law early in the data period will contribute fully to the estimated post-passage effect, while a state that adopts near the end of the period will have little weight. Since we know that the late adopters tended to experience crime increases, the aggregated analysis will give less weight to these states in estimating the overall effects of shall-issue laws.⁷⁸ Thus, what might look like a changing effect over time from the passage of the law may simply be a compositional effect as certain states drop out of the analysis.

One way to avoid these aggregation and compositional biases is to change the specification to estimate a state-specific effect for each state that adopts a shall-issue law.⁷⁹ In other words, we include in our regression for each crime category a separate post-passage dummy for each adopting state (as opposed to a single post-passage dummy pertaining to all adopting states). Building on our previous dummy variable model with

⁷⁷ Simpson's Paradox, also sometimes referred to as aggregation bias, is just one such example. See Bickel, Hammel, and O'Connell, "Sex Bias in Graduate Admissions: Data From Berkeley," 187 *Science* 398 (1975) (showing that while aggregate data suggested bias against female graduate applicants, the disaggregated data by department reversed this conclusion).

⁷⁸ If we run the Table 10 regression on just the years 1991 to 1997, we find statistically significant increases in all nine crime categories.

⁷⁹ Black & Nagin, *supra* note 2, were the first to run a disaggregated regression estimating state specific impacts of the shall-issue law. However, their analysis was limited to Lott's initial data set, which only allowed them to test for the impacts on ten passing states. They reported substantial heterogeneity in the law's impact for a sample limited to large counties, but their results differ from ours in that (for their limited data set) the state specific impacts were more beneficial, but the impacts tended to be less statistically significant. They also limited their analysis to the dummy specification – which may mis-estimate the true impact of the law (because of Lott's inverted V argument or some other misspecification). For example, in Table 12, we include state fixed trends as an added control.

state fixed trends (Table 11), we now use the full 1977-1997 county data set to estimate the effect on 9 crime categories for 24 jurisdictions that adopted shall-issue laws – a total of 216 estimates. Table 12 presents all of these estimates for all 9 crime categories, while Figures 4 - 6 graphically depict the results for violent crime, murder, and property crime.⁸⁰ These state-specific “dummy variable” coefficients represent an even less constrained specification than we saw in line 4 of Table 11 in that they separately estimate the impact of shall-issue laws for each of the 24 jurisdictions that passed the laws between 1977 and 1997. The Table 12 results reject the more constrained specifications of the aggregate regressions, which implicitly assumed that the impact of the shall-issue law was constant across jurisdictions. More importantly, the state specific estimates frequently undercut Lott’s “more guns, less crime” thesis.

For every crime type there are more states where shall-issues laws produce a positive and statistically significant coefficient than states that produce a negative and statistically significant coefficient. For example, as Figure 4 reveals, while there are three states that experience a statistically significant drop in violent crime upon passage of the law, there are five states that experience a statistically significant increase. Overall, there are almost twice as many jurisdictions (15) that have an estimated increase in violent crime as those that have an estimated decrease (9).⁸¹

The second column of Table 12 and Figure 5 both reveal the even more dramatic disparity for murder: there are 8 states with a statistically significant *increase* in murder, while only 4 states exhibit a statistically significant decrease. Of the 12 jurisdictions that experienced a statistically significant effect on property crime from the passage of a shall-issue law, Figure 6 shows that 8 experienced an *increase* in crime.

Stepping back, we see that out of the 216 estimated impacts in Table 12 (24 jurisdictions by 9 crime categories), 68 exhibited statistically significant increases in crime while only 27 exhibited statistically significant decreases. Overall, Table 6 shows 141 increases in crime versus only 75 decreases. The striking implication from this disaggregated analysis is that, according to this model, shall-issue laws increased crime in substantially more jurisdictions than it decreased crime.⁸²

How can it be that the disaggregated analysis in Table 12 is suggestive of crime increases, while the aggregate model (Table 11, line 4) suggests statistical decreases in crime? First, note that weighting by population in the regression gives far greater influence in the regression to large states and that Texas and Florida (the two largest

⁸⁰ As noted above, Philadelphia is treated as a separate jurisdiction, because the law became effective in the city of Philadelphia at a different time than for the rest of Pennsylvania. For convenience, we will still refer to state-specific estimates in referring to these 24 jurisdictions.

⁸¹ Figure 4 also shows the estimated effect for the corresponding aggregated model (analogous to line 4 of Table 11), which is 1.5 percent (albeit statistically insignificant).

⁸² The same story comes through if one uses the Table 12 analysis on state data. In that event, there are only 207 effects (9 regressions x 23 jurisdictions) because we don't treat Philadelphia separately in the state data regressions. We find that there are 37 significant increases in crime versus 9 significant decreases. Overall, 122 of the 207 effects were positive versus 85 negative.

states) had large and statistically significant estimated drops in violent crime after they passed shall-issue laws.⁸³ As Table 10 indicated, the estimated aggregated effect on murder in the dummy variable model is a drop in crime of 7.7 percent. Running the aggregated regression without weighting by population lowers the estimated effect on murder from -7.7 percent to -5.1 percent. Hence, weighting clearly increases the apparent murder reducing capacity of shall-issue laws in the aggregated dummy variable model, but it is not the entire story.

Second, as we have seen, the fact that a state adopts a shall-issue law earlier means that it will have a greater impact in the estimation of any post-passage dummy in the aggregated analysis. Thus, imagine a scenario under which only two states (with equal populations) adopt shall-issue laws -- one in 1987 and another in 1996. Assume the effect in the two states is exactly opposite, in the early adopter crime *drops* by 10 percent in the first year after passage and stays at that lower level through 1997, while in the late adopter crime *increases* by 10 percent and will stay that way for ten years. In the disaggregated analysis, one will see equal and opposite impacts, suggesting no overall net effect on crime. This is also what the aggregated dummy variable analysis would show if the laws had been adopted at the same time. But the later adoption in the second state means that its impact will be diminished when the aggregated dummy variable model is estimated. Indeed, the aggregated effect in this hypothetical will be a drop in crime of 9 percent because the 10 years of a crime drop of 10 percent will be averaged with the 1 year of the crime increase of 10 percent.⁸⁴ As it turns out, two (large) early passing states (Florida and Georgia) experienced drops in murder -- thus inordinately dragging down

⁸³ Even though one can interpret the coefficients on the individual state post-passage dummies as the percentage effect of the law on crime, one has to at least entertain the prospect that these estimates are picking up other changes in the states in question that happened to coincide with the passage of the shall-issue laws. This could happen for any state, but one that has been singled out on this ground is Florida because of the influx of roughly 125,000 largely male, uneducated, and young refugees from Cuba from 1980 through 1981 in the Mariel boatlift, which swelled crime in Florida in the early 1980s, presumably followed by at least some crime decline once the refugees had been fully integrated into the community (or removed from it). David Card, *The Impact of the Mariel Boatlift on the Miami Labor Market*, 43 *Industrial and Labor Relations Review* 245, No. 2 January 1990. Any such crime decline occurring after 1987 from this factor would be captured in the Lott analysis as the result of shall-issue law. Moreover, even if the effect of the Mariel boatlift had been completely dissipated by the time of passage as Lott has argued, the very sharp spike in crime that can be seen for Florida in Figure 1(a) of Ayres and Donohue (1999) reveals that the pre-passage fixed effect for Florida would be artificially elevated, biasing downward any estimated effect on the post-passage dummy. See Lott, "the Concealed Handgun Debate," 27 *Journal of Legal Studies* 221, 232 (1998).

Of course, one cannot simply select the states that seemed to do well under the law for further evaluation, but it does suggest that some examination of whether there were any identifiable factors influencing crime in any of the states that appear to have large crime changes -- whether positive or negative -- at around the time of the adoption of a shall-issue law might be worthwhile.

⁸⁴ Lott includes graphs in his second edition showing the distorting impacts of unbalanced data sets in estimating the impact of the law. Lott (216).

the estimated aggregate impact. But when we decouple the impact of the law on individual jurisdictions, a much different picture emerges.⁸⁵

Lott might respond that these jurisdiction-specific dummy effects could understate the true impact of the law because his “inverted V” concern might operate on an individual state by state basis. While this specific concern is dampened somewhat by the inclusion of state-specific trends in our regressions, there is value in exploring whether the hybrid analysis is superior to the dummy variable model for the disaggregated analysis as it was for the aggregated analysis. Accordingly, we employed a disaggregated version of the hybrid specification, which estimates for each jurisdiction both an intercept effect and a trend effect. While only 30% of the estimated state-specific spline effects were statistically different than zero, we were able to reject in each of the 9 crime type regressions the hypothesis that the 24 disaggregated spline effects were jointly equal to zero. Thus, the regressions suggest that the implicit constraints of the disaggregated static model are once again too restrictive.

While we report the raw coefficients of these hybrid regressions in an appendix,⁸⁶ Table 13 reports the net 5-year impact of the law annualized in order to facilitate comparison with the static model.⁸⁷ Turning to the substance of the disaggregated hybrid specification captured in Table 12, we find a remarkably consistent pattern to that presented in Table 12. Just as in the static model, for the disaggregated hybrid

⁸⁵ Our disaggregated results also substantially weaken the power of Lott and Bronar’s geographic substitution result. Lott and Bronars use an aggregate specification to show that passage of the law caused crime to decrease in the passing states but increase in adjoining states – because, they argue, of geographic substitution. Lott & Bronars, *supra* note 13. But this purported spillover result could simply be a byproduct of aggregation bias. Our disaggregated analysis demonstrates that passage of the law was likely associated with increases in many metro areas and with decreases in others. The Lott and Bronars story would only be true if crime fell on the shall-issue side of the metro border and rose on the no-shall-issue side, but nothing in their aggregated analysis would ensure this was the case. If the spillover regression were re-estimated on a more disaggregated basis, we predict that most metro areas would show similar movements in crime in both the areas that were covered and uncovered by the law, which would be the exact opposite of the Lott and Bronars hypothesis of crime falling on one side of the border (in response to the shall-issue law) and rising on the other.

⁸⁶ See Appendix Table 7.

⁸⁷ To calculate the 5-year impact of the shall-issue law under the hybrid specification it is necessary to add together the impacts of the intercept and trend terms for individual years and then sum the yearly impacts. For example, the predicted impact of a law for individual years is:

Year 1:	$1 * \text{beta}(\text{shall dummy in state X}) + 1 * \text{beta}(\text{spline trend in state X})$
Year 2:	$1 * \text{beta}(\text{shall dummy in state X}) + 2 * \text{beta}(\text{spline trend in state X})$
Year 3:	$1 * \text{beta}(\text{shall dummy in state X}) + 3 * \text{beta}(\text{spline trend in state X})$
Year 4:	$1 * \text{beta}(\text{shall dummy in state X}) + 4 * \text{beta}(\text{spline trend in state X})$
Year 5:	$1 * \text{beta}(\text{shall dummy in state X}) + 5 * \text{beta}(\text{spline trend in state X})$

where $\text{beta}(\text{shall dummy})$ and $\text{beta}(\text{spline trend})$ represent the estimated coefficients on the intercept and trend variables. Summing these individual year impacts together, we were able to calculate a net annualized 5-year impact as:

$$\text{beta}(\text{shall dummy in state X}) + 3 * \text{beta}(\text{spline trend in state X}).$$

We also tested whether this linear combination of regression coefficients was statistically different than zero and report the results of this testing above in Table 13.

specification more states experienced statistically significant increases in crime after the shall-issue law than experienced statistically significant decreases in crime. Overall, 17 of the 24 states report a net increase in violent crime, and 21 out of 24 showed an increase in assault. Only one state (Florida) showed a statistically significant drop in violent crime, while 6 states showed statistically significant increases in violent crime. Similarly, while 8 states experienced a statistically significant increase in assault, not one experienced a statistically significant decrease. In fact, as before, every crime category reports more increases than decreases in crime. For example, there are 6 states showing a statistical increase in murder while only 4 report a statistical decrease. And the disparity is even greater for rape with 4 states displaying a statistical increase, while only 1 state reports a statistical decrease. Overall 150 (of the 216) tests indicate that the shall-issue law increased crime and 59 of these estimates were statistically significant (at a 5% level), while only 17 states report a statistically significant decrease – a ratio of more than 3 to 1.⁸⁸

Indeed, for the clear majority of states for *all crime types*, shall-issue laws are associated with *increases* in crime, and the statistically significant impacts are more than twice as likely to exhibit increases in crime. While the story of murder or robbery dropping can be found in the aggregated analysis with the linear spline model, it is purely an artifact of the happenstance of early adoption that weights a few large states most heavily.⁸⁹ If one takes the population-weighted average effect for the 24 passing jurisdictions in Table 12, shall-issue laws are associated with *more* crime in all 9 crime categories. For table 13, this is true for 8 of the 9 categories, the sole exception being a -.5 percent weighted average for murder, which is statistically insignificantly different from zero.

We take these disaggregated (state-specific) hybrid regressions to be our most definitive results for the county based data. In a sense, our analysis of this county data has been an exercise in testing and rejecting a series of progressively less constrained specifications. We began by rejecting the simple aggregate dummy variable and spline

⁸⁸ These results were qualitatively unaffected when we instead calculated the 4-year and the 6-year annualized impact of the law. We also estimated the Table 13 results for state data, which generated 31 positive and statistically significant 5-yr annualized effects versus 18 negative and statistically significant effects. Overall, there were three more positive effects than negative effects, 105 vs. 102.

⁸⁹ The disaggregated analysis is also amenable to the same kinds of test of internal theoretical consistency that we undertook earlier with respect to the aggregate analysis. Looking again at property crimes and robbery, we see in Table 13 on a disaggregated basis that 16 (of 24) states experienced an increase in property crime. And that 14 of those 16 states also experienced an increase in violent crime of which 5 were statistically significant increases in violent crime. Of the other two states, that experienced an increase in property crime but a decrease in violent crime, in only one was the decrease statistically significant. Once again, the strong conclusion is that shall-issue laws are associated with *increased* property crime (whether looking at a disaggregated or aggregated analysis), yet we have no plausible story to back up this effect. Table 13 shows neither a general shift from violent to property crime, nor a more nuanced shift from robbery to property crime, as none of the 16 states that have estimated increases in property crime reported statistically significant decreases in robbery.

models in favor of the aggregate hybrid specification. We next rejected the constraint that the law had the same impact on early and late passing jurisdictions. We then rejected the decision to exclude state-specific trends. And finally we rejected the disaggregated dummy variable specification. The disaggregated hybrid model that we have finally settled on allows the data to reveal a variety of different impacts of the law – allowing separate intercept and trend effects for each of the 24 passing jurisdictions. And while we might have concerns that estimating this many impacts would rob the regressions of statistical significance (as we eat up degrees of freedom), we still find that over one-third of the state/crime type tests (76 out of 216) are statistically significant.⁹⁰

But in interpreting the results, it is important not to forget the lessons of Part II – which showed that the results of the state data regressions were incredibly contingent on an array of factors, such as the inclusion or omission of rather innocuous right-hand side controls. We found a similar fragility with regard to these county data regressions. Indeed, as we were double checking our results we discovered that we had omitted one of Lott’s original controls -- Percentage of County Population that Was Not Black or White But Was Male and Aged 20-29. And we were surprised to find that adding an extremely innocuous demographic variable decreased our estimate of the dollar impact that the law’s passage had on crime (that we are about to report) by more than 20%. The non-robustness of the results with respect to these attenuated demographic controls is a legitimate reason to discount or at least question this entire enterprise. But because the disaggregated county data still yields some jurisdictional pockets in support of Lott’s “More Guns, Less Crime” hypothesis, we soldier on to assess a possible net impact of the laws.

To get a better handle on the net impact of the law on all crime, we have estimated the dollar impact of the law on particular crime categories – using the same

⁹⁰ It is possible of course to estimate even less constrained specifications that admit the possibility of higher order impacts. And indeed, we estimated a disaggregated quadratic hybrid that is identical to the disaggregated hybrid discussed above, but which includes a pre-passage quadratic term and a post-passage quadratic spline term. Estimating this quadratic hybrid specification allowed us to test (1) whether the implicit restrictions of the (linear) hybrid are rejected by the less constrained specification and (2) whether the results of the (linear) hybrid were robust to the less constrained specification. We found that the (linear) hybrid’s implicit assumption of no quadratic post passage effect was not decisively rejected in that only 49 of the 216 coefficients were statistically different than zero (although the quadratic spline effects were jointly different than zero in eight of the nine regressions). But the basic results of the (linear) hybrid analysis discussed in the text remain unaltered: calculating the net annualized 5-year impact, we continued to find that the vast majority of the statistically significant impacts were positive (48 vs.20).

Lott has sharply criticized Black and Nagin’s decision to estimate state-specific quadratic time trends in an aggregated dummy variable model. See Lott (2009) & Black & Nagin, *supra* note 2. We agree that such a specification can understate the law’s impact if the impact comes primarily through a kink or bend in the time trend. But Lott’s criticism is not relevant to our quadratic hybrid specification, which allows the law’s impact to come through an intercept effect, a linear spline effect, or a quadratic spline effect.

dollar value per crime that Lott used in his initial study.⁹¹ Table 14 reports the annualized dollar impact of crime for each of the 24 jurisdictions and for each of the 9 crime categories – as well as various aggregations of these amounts. 18 of the 24 states have estimated increased dollar harms, as shown in the “Total” column, which is also depicted in Figure 7. Viewed in aggregate the Table suggests that the net annual impact of the law was to increase the dollar harm of crime by approximately \$1 billion. This represents a “harm weighted” annual increase in crime of 2% that amounts to an annual burden of \$9.63 on the average citizen in the passing states.

But the Table also reveals substantial variation in the impact of the law. In Texas and Florida alone, the law is estimated to have *reduced* the annual cost of crime by more than \$3 billion, while in Louisiana and Tennessee the law is estimated to have *increased* the annual cost of crime by \$1.7 billion. In the 24 adopting jurisdictions that we examined, the mean impact of the law was a \$41 million increase in crime, but the standard deviation was more than fourteen times this amount (\$613 million).⁹²

These dollar figures price all of the estimated impacts of the law (reported above in Table 13) regardless of their statistical significance. An alternative way to estimate the aggregate impact of the law is to put a zero dollar value on all the impacts that are not sufficiently statistically significant. Table 15 reports the aggregate dollar impact of the law for alternative levels of significance – for both the dummy variable specification (originally reported in Table 13) and our preferred hybrid specification (of Table 14). The table shows that the law continues to display a pernicious dollar effect even when we limit our focus to 5-year impacts that were statistically significant at the 10% or 5% level. At these levels, the static specification estimates a net annual impact on the order of half a billion dollars, while the less-restrictive hybrid specification estimates an increased cost ranging between \$3 and \$524 million.⁹³

We take these results to be generally devastating to Lott’s “More Guns, Less Crime” hypothesis. Estimating a less constrained specification with more data, we find that more jurisdictions experience an increase in crime than a decrease and that the statistically significant increases outpace the statistically significant decreases by more than a two to one rate in county data (and still by a substantial amount in state data).

⁹¹ In 1997 dollars, the harm of the following crimes was assumed to be: murder, \$3,092,804; rape, \$91,522; aggravated assault, \$25,247; robbery, \$8416; auto theft, \$3,892; burglary, \$1472; and larceny, \$389.

⁹² One can see from Figure 5 that the two states showing the biggest dollar impact on crime are Texas (a crime decline of over \$2 billion per year) and Louisiana (a crime increase of over \$1 billion) both adopted the law late in the data period, which implies that there is relatively little post-passage data with which to estimate these figures. It is quite likely that with more years of data, one would see the estimated effects for those two states move closer to the mean.

⁹³ Only if we restrict attention to 5-year impacts that were statistically significant at the 1% level do we find a net benefit of the law – with an estimated reduction in crime in the hybrid specification of \$784 million. But here it is important to note again that the result is driven by just 2 states – Texas and Florida (with an estimated combined benefit of more than \$3 billion) – which more than offset the estimated crime-increasing impact of Louisiana and Tennessee (approximately \$1.7 billion).

V. Some Interpretations, Speculations, and Methodological Insights

Our effort to find the statistically most appropriate model with which to assess the impact on crime of shall-issue laws has involved an extended odyssey as our testing, on a more comprehensive county data set, has constantly pushed us towards more disaggregated and less restrictive models than the more aggregated and highly constrained models employed by Lott and Mustard. The lesson has been a sobering one in that the facially plausible models relied upon by Lott and Mustard that we present in Table 10 could well encourage a researcher or policymaker into believing that shall-issue laws reduce crime. Yet when we reached the end of the journey with our more complete data, the use of a statistically superior model that estimates jurisdiction-specific effects while estimating both main and trend effects (our “hybrid”) and controlling for state fixed effects reduces the initial conclusion to ashes. The best, albeit admittedly imperfect, evidence suggests that, for the majority of states, shall-issue laws are associated with higher levels of crime.

It is important, though, to be clear about the degree of confidence that we can repose in any particular interpretation of the evidence. In the end, we are left with a hierarchy of 3 conclusions that we will discuss in turn below.

1. There remains no robust, credible statistical evidence that the adoption of shall-issue laws will generally lower crime, and indeed the best, albeit admittedly imperfect, statistical evidence to be presented thus far points in the opposite direction, that the adoption of shall-issue laws will generally increase crime.

We believe that a fair evaluation of all the state and county evidence that we have presented offers virtually no basis for believing that shall-issue laws reduce crime. While particular regressions at times predict drops in certain violent crimes following adoption, the overall regression evidence predicts increases in other crime categories, most frequently among the many types of property crime. Indeed, there is abundant reason to be concerned that the regressions are not performing particularly well, and plausible adjustments to the aggregated regressions – whether by using alternative model specifications or by trying to control for selection effects of very late or very early adopters in the state analysis or by introducing state trends in the county data – largely eviscerate the more guns, less crime hypothesis. In particular, state-specific estimates generated on the 1977-97 county data are more supportive of the view that shall-issue laws increase crime than that they decrease it.

While Lott and Mustard have tried to offer other types of evidence that can bolster their core statistical findings, if the foundation falls, the entire edifice will crumble. We believe we have shown the foundation has collapsed. Whether further advances in statistical modeling or additional years of data analyzing more state adoptions (or repeals) of shall-issue laws will be able to resurrect the structure remains to be seen.

We hasten to add, though, that showing that superior statistical modeling on more complete data reverses the Lott and Mustard conclusion does not necessarily resolve the debate, since “better” doesn’t always imply “good enough.” The non-robustness of the estimates with regard to innocuous changes in specifications provides grounds for rejecting the authoritativeness of any of the results. Therefore, another plausible conclusion from the evidence that we have presented is:

2. While the best evidence suggests shall-issue laws generally tend to increase crime, there is still too much uncertainty to make strong claims about their effects.

The dramatic reversal in findings in moving from Table 10 to Table 13 certainly reveals that many conclusions about the impact of shall-issue laws will be dependent on the particular statistical model that is employed. Some will be convinced that our disaggregated model is superior and therefore the conclusions of the state-specific analysis can be accepted. More cautious analysts will be concerned that the problems we have highlighted of data accuracy, model mis-specification, endogeneity, and lack of robustness (such as the sensitivity of the state data results to exclusions of minor demographic controls) are too severe to confidently assert whether shall-issue laws dampen crime, increase crime, or have no overall effect on crime. We share these concerns, especially since the theoretical argument for one of the most robust findings – that the laws increase property crime -- is not particularly compelling.⁹⁴ Thus, a plausible interpretation of the existing evidence would be that even if evidence can be mustered that shall-issue laws generally appear to increase crime, the uncertainty about whether the statistical models are working properly makes it difficult to make any strong claim about the impact, other than to say that it is not so huge that it can overwhelm any defects in the model.

At the end of the day, then, it is still possible that shall-issue laws have no effect – positive or negative – on crime. This is particularly so if one credits Willard Manning’s suggested correction for the presence of these multiple comparisons and for autocorrelation in crime across years. In addition, as discussed above, one reason states like Florida and Texas – which both showed large crime decreases after adoption -- may have an estimated negative impact is simply that they passed the law in response to crime increases and as crime reverted of its own accord to its normal levels, the regression inappropriately attributed the mean reversion to the passage of the law.

Earlier in this paper, we touched upon an alternative reason why even the best regression results may not be believable – the huge omitted variable problem represented by the upturn in crime following the advent of crack into certain urban areas in the mid to late 1980s. If states adopting shall-issue laws in the 1980s had no major crack problem, while non-adopting states were more likely to have a crack problem, then that fact alone

⁹⁴ The figures we provide show property crime rising as much in response to shall-issue laws as violent crime does (with virtually no evidence of the shifting from violent crime to property crime that was initially posited by Lott and Mustard).

could make it appear that shall-issue laws reduced crime, because crime would rise relatively faster in the non-adopting states (by virtue of the crack problem). In his book, Lott reported that the states adopting shall-issue laws tend to be Republican and have high NRA membership and low crime rates. That doesn't sound like the sort of place where one finds the worst problem with crack. Without a plausible control for the criminogenic influence of crack, every crime regression is subject to a charge of inadequacy owing to the omission of a potentially important explanatory variable that is correlated with the adoption of shall-issue laws.⁹⁵

Mis-specifications of this type make it difficult for the researcher to distinguish between the possibility that the law has no effect and the possibility that any effect of the law is beyond the current ability of researchers to identify. Ironically, however, either a "no effect" or "don't know the effect" assessment might be enlisted to argue normatively for adoption of the shall-issue law. While the first empirical assessment would tend to lead consequentialists to oppose the law,⁹⁶ the second assessment of (no effect or) ignorance might provide a libertarian ground to support such legislation. There are, however, many reasons why reasonable people who embrace conclusion 2 might nevertheless oppose the statute. Non-consequentialists and expressivists may oppose the law notwithstanding lack of evidence that the law increases crime. Risk aversion or a concern that the law moves society away from a more global maximum of altogether fewer weapons might also ground opposition. Finally, if the realization that thousands of citizens were carrying around concealed handguns generated fear or apprehension in the community, one might oppose shall-issue laws even if one could not prove that they increased crime. In a world where NRA members have bumper stickers stating "Keep Honking, I'm Reloading," the costs of intimidation of law-abiding citizens may become intolerable.⁹⁷

⁹⁵ In his second edition, Lott admirably included region-specific fixed year effects for five regions to allow for more heterogeneous year effects for different parts of the country. Lott (170). But Lott's less constrained approach would not be sufficient to solve the crack problem (outlined above) which may very well have played out within the regions.

⁹⁶ Lott, himself, is clearly in the consequentialist camp. See Lott (21) ("the ultimate test: does it save lives?")

⁹⁷ David Mustard, "The Impact of Gun Laws on Police Deaths" *Journal of Law and Economics* 44(2): October 2001, cites a 1999 incident in Phoenix in which a concerned citizen with concealed gun permit came to the assistance of a dying officer and helped capture three fleeing drug suspects. The citizen returned fire by shooting 14 shots into the car of the criminals, which "he thought was pretty good since he shot quite a few of them with his off hand, hanging out the window." While the apprehension of dangerous felons was an obvious benefit of the citizen's intervention, it is not hard to imagine how such behavior could have ended badly for some innocent bystanders. Simply because this episode ended happily doesn't mean that we should encourage such efforts. Perhaps most importantly, the message on the citizen's car demonstrates one of the costs of shall issue laws that Lott and Mustard have never acknowledged. The citizen's license plate holder states: "Keep honking, I'm reloading." It is a very high

Some may feel that just as many were willing to make policy choices on the basis of the initial Lott and Mustard study, there is even stronger reason to rely on our less constrained models estimated on more years of county data showing that after passage twice as many jurisdictions experienced a statistical increase as experienced a statistical decline in crime. Those who are willing to repose confidence in such results have to consider whether the disaggregated evidence might support a view that some states would benefit from shall-issue law adoption even though most would not. We will now evaluate the normative implications of this heterogeneity assessment. While the first and second assessments militate toward across the board policies, the heterogeneity assessment might counsel toward a more nuanced, piece-meal adoption of the statute.

3. We should simply accept the 24 different jurisdiction-specific estimates, and conclude that shall-issue laws increase crime in most states but reduced it in other states.

This is the most speculative of the three conclusions we discuss in that it violates Milton Friedman's admonition against accepting statistically significant findings too readily. On the face of the evidence presented in Table 13, there are a number of problems with accepting that crime fell with the adoption of shall-issue laws in even a limited number of states. There is only one state that shows a statistically significant decrease in at least two of the five violent crime categories without showing a statistically significant increase in another violent crime category (Georgia, showing drops in rape and robbery). Florida shows substantial drops in violent crime and murder, but a statistically significant increase in aggravated assaults (all against the background of the potentially confounding influence of the Mariel boatlift), while Oregon shows huge drops in murder and robbery, but a huge increase in aggravated assaults. Only two other states show a statistically significant drop in any other violent crime category (Texas and Montana saw enormous drops in murder). In other words, it is rare (only 5 out of 24) to see any statistically significant evidence of declines in any violent crime category from the adoption of a shall-issue law. Certainly, there is nothing to give one confidence that an overall drop in violent crime is likely to be spawned anywhere by the adoption of a shall-issue law. Indeed, the rare and seemingly haphazard pattern of statistically significant drops across isolated violent crime categories makes one think of these drops as more random noise than estimates upon which much confidence can be reposed that real effects have been identified. This is underscored by the fact that finding 8 statistically significant drops in crime across the 120 estimates (24 jurisdictions times 5 violent crime categories) in Table 13 looks to be only modestly more than one would expect from a purely random process.

Illustrating A Methodology for Deriving State-Specific Predictions

One interesting consequence of conclusion (3), however, is that it invites the researcher to investigate whether there may be some particular attributes about the small

social cost if law abiding citizens have to worry that if someone with a concealed weapon takes offense at their lawful (even if annoying) actions, they might get shot.

number of states for which crime drops were estimated that diminish the generally harmful effects of shall-issue laws enough so that they may actually dampen crime. If one overcomes the concern that the infrequent instances of crime reductions are merely random artifacts and concludes that there is valid information captured in the harm-weighted estimate of the total effect on crime for a state, we can investigate whether there is any pattern that explains which states show overall crime increases and which show decreases. Unfortunately, we have very little data to make this assessment. The county fixed-effects model which we have followed Lott in using to test for the law's impact only allows us to test for the impact of changes in the law within particular counties.⁹⁸ But now we are called upon to assess the determinants of the law's impact across the passing states. In essence, we have just 24 observations (23 passing states plus Philadelphia) on which to try to disentangle what caused the law to increase crime in some jurisdiction and decrease it in others.

Since we are now taking the output of potentially fragile regressions that were used to estimate the impact on crime from adopting a shall-issue law in 24 jurisdictions, and then trying to run yet another regression based on this output having only 24 observations, we are now moving very far out on a branch. Still, there is value in at least detailing a methodology for deriving state-specific predictions of the effect of a law even for those that have not adopted the legislation. Accordingly, we regressed various measures of the impact of shall-issue laws on a variety of state characteristics. But before reporting the results, let us emphasize that with just 24 observations we are certainly pushing the limits of data. The results that we report are not nearly as robust as our prior findings to the inclusion or exclusion of other variables.⁹⁹ Indeed, the data's resistance to explanation can be taken as yet more evidence that, notwithstanding our best efforts, the underlying model (the disaggregated hybrid version) is still mis-specified.

With these caveats about the substantial limits in our data and the residual concerns with mis-specification, we proceed to discuss the impact regressions of Table 16. The dependent variable in the first row regression is the harm weighted percentage impact of the hybrid regression (reported above in Table 14), which ranges from a 25.7 percent *drop* in crime in Texas to a 35.7 percent *increase* in crime in Nevada.¹⁰⁰ We use

⁹⁸ Lott attempted to assess what characteristics were associated with larger or smaller impacts of the law by interacting the law dummy variable with various demographic characteristics of the county (its density, its income, etc.). But this interaction specification in a fixed-effects model only allowed Lott to assess the impact of, say, changes of density *within* particular counties that had passed the law. Since, on average, there is very little within-county variation in density over a twenty-year period, this specification can tell us very little about impact of different densities across different counties or states.

⁹⁹ And the fact that we ran a number of alternative specifications that we did not report suggests a "pre-testing" effect that should lead us to discount the nominal levels of significance reported in the regression.

¹⁰⁰ The dispersion in the estimated effects of the 24 shall-issue laws is quite wide, which in itself shows that there is a considerable degree of noise in the estimate of any particular jurisdiction. As Black and Nagin comment in finding substantial dispersion in state-specific estimates for the 10 states adopting shall-issue laws during the 1977-92 period: "Widely varying estimates such as these are classic evidence that, even beyond

the following 6 jurisdiction characteristics as explanatory variables: the year the law was adopted, population, log of violent crime rate, density (measured in population per square mile), and two regional dummies (South and West).¹⁰¹ Looking at the estimated coefficients, we see that later passing states are predicted to have a more deleterious impact – with each additional year adding more than 1.6 percentage points to the estimated impact (prob. = .07). This is huge effect (a law passing 10 years later would be expected to have a 16 percentage point higher crime impact), and is qualitatively consistent with the aggregate results showing crime increases in the 1990s for later passing jurisdictions.¹⁰²

The harm-weighted estimates that we use as regressors have the advantage of aggregating the individual crime category impacts in a natural way that weights the individual crime estimates commensurate with their underlying importance. But a disadvantage of the harm-weighted estimates is that they ignore the varying significance of the individual crime impact estimates. Returning to Figure 5 (which depicts the disaggregated dummy variable model estimates for murder from Table 12), one can readily identify the positive correlation between the estimated impact of the law on the murder rate and the year in which the law was adopted. We see that 13 of the final 14 passing states had estimated increases in murder (and 7 of these were statistically significant), while only 4 out of the first 10 passing states posted estimated increases in murder (and only 1 of these was statistically significant).¹⁰³ More generally, Table 14 shows that 11 of the last 13 passing jurisdictions experienced *increases* in violent crime.

If one accepts the accuracy of the jurisdiction-specific estimates, one might interpret this cohort result as a kind of “Peter Principal” applied to law. As a pop theory of job advancement, the “Peter Principal” asserts that employees are promoted to jobs requiring successively higher skills until they reach a job where they are relatively incompetent. Analogously, the data indicates that the law has performed less well –

the assumption of homogeneous impacts across states, the model is misspecified.” Dan A. Black & Daniel S. Nagin, “Do Right-To-Carry Law Deter Violent Crime?” 27 *Journal of Legal Studies* 209, 214 (1998).

¹⁰¹ There were no Midwest states that passed the law during this twenty-year period (and the Northeast region was the excluded attribute).

¹⁰² Query, though, whether this result tells us that the later adopters can expect more pernicious outcomes from adopting shall-issue laws or whether it is reflecting the sharp drop in crime in the post-crack world for the non-adopting states. The year of adoption is clearly a proxy for something – but it is uncertain whether it is a proxy for a feature of the state that will influence the impact of a shall-issue law (as the regression posits) or a proxy for the time period during which sharp declines from previous crack-induced crime hikes occurred in selected non-adopting states.

¹⁰³ In his second edition, Lott attributes the diminishing beneficial impact of the law in later passing jurisdiction to heightened fees and training requirements that were imposed on permit applicants in the later passing states. But while the imposition of greater obstacles could have explained a diminished beneficial trend, it does not explain why we find that later passing states generally experienced increases in crime. It also does not explain why Texas (which Lott notes requires 10 hours of training and charges the highest fee in the sample) was one of the great outliers in generating a beneficial impact. We did not have available the fee and training data that Lott used, and hence did not control for these attributes in the regressions reported in Tables 10 and 11.

rising to the level of its incompetence – as successive states have chosen to adopt it. On the other hand, the temporal pattern that states that adopted shall-issue laws in the late 1980s did better while those adopting in the 1990s did worse may simply reflect the influence of a time-varying factor (the crack trade?) that caused sharp rises in crime for many states in the late 1980s and then greater than average price declines in the 1990s.

The regressions, however, do at least suggest some non-temporal traits that are associated with the estimated impact. Passing jurisdictions with larger populations had more beneficial impacts and this correlation was strongly significant, while less dense jurisdictions and jurisdictions starting with higher base levels of violent crime had more pernicious impacts (and these effects were significant at the 10% level).¹⁰⁴ The negative association between state population and the estimated dollar impacts (estimated in Table 14) is dramatically illustrated in Figure 5 – with the two largest states (Texas and Florida) having by far the largest harm-weighted dollar drops in crime. This should not be surprising, though, since a given percentage change in crime will have a bigger dollar impact in a large (or higher-crime) state. Finally, western states tended to be associated with better crime outcomes, and this effect was statistically significant.¹⁰⁵ While there are always some “cultural” rationales that we could offer ex post for these effects, we should emphasize that these results are suspect as they are the byproduct of reduced form regressions rather than growing out of a priori theory.

B. The Resulting State-Specific Predictions

Having derived the regression estimates of the factors that influence the impact of the law, we are at least conceptually able to use the results of these regressions to predict out of sample the expected impact of the law on the jurisdictions that either have never passed a shall-issue statute or had already passed the statute before our data began in 1977.¹⁰⁶ To continue our methodological roadmap for how to generate these estimates (for cases in which the regression results were both stronger and were based on a stronger foundation), Appendix Table 8 reports the predicted impact (as well as the standard deviation of the prediction) of passing the law today in all 52 jurisdictions (including Philadelphia and the District of Columbia) -- given the jurisdiction’s underlying characteristics. While this predictive process is based on extremely limited data, it may be useful to illustrate the possibility of more nuanced state-specific policy recommendations in contrast to the across the board recommendations to adopt or repeal that are the only possible product of the type of aggregated analysis that Lott championed.

As one might suspect the predictions produced heterogeneous results – which we break into four categories:

¹⁰⁴ While potentially interesting, it is not evident that any theory could support this empirical finding.

¹⁰⁵ In the remaining columns of Table 10, we replicate the regression of the first column by regressing alternative measures for the impact of crime coming from the murder category and from the static regressions. One sees that the patterns of sign and statistical significance are generally consistent.

¹⁰⁶ Lott undertook an analogous prediction procedure when he uses a first-stage regression run on ten states to predict how many permits would be issued in other states. Lott (176).

Jurisdictions in which adoption is strongly counterindicated. We found 31 jurisdictions in which adoption of the law predicted an annual percentage increase in the harm-weighted measure of crime that was more than twice the prediction's estimated standard deviation. 15 of these jurisdictions had already adopted the law and thus would need to repeal the law to avoid the deleterious effect, while 16 of these states would merely need to refrain adopting the law in the future.

Jurisdictions in which adoption is strongly indicated. We found just two states (Texas and California) in which adoption of the law predicted an annual percentage *decrease* in the law that was more than twice the prediction's estimated standard deviation. One of these jurisdictions (California) would need to adopt the law to secure its benefits, while the other (Texas) would merely need to retain the law that it had previously adopted.¹⁰⁷

Jurisdictions in which adoption is weakly counterindicated. We identified 14 jurisdictions in which adoption of the law predicted annual percentage increases in crime, but were not statistically significant at the 5% level. Risk-aversion would probably counsel against passing (or retaining) shall-issue laws in these jurisdictions. Under this normative reckoning, 11 of these jurisdictions had previously passed the law and thus would be targets for repeal, while 3 should simply continue to refrain from adopting the law.

Jurisdictions in which adoption is weakly indicated. Finally, we identified 5 jurisdictions in which the predicted impact of the law on a harm-weighted measure of crime was beneficial but not statistically so. A libertarian would argue that these five states – which all previously passed the law – should retain them. However it should be noted that in 4 of these 5 states (Virginia, West Virginia, Pennsylvania and Montana) the shall issue law was estimated in Table 1h to have a deleterious (but statistically insignificant) impact on crime – so that the more direct measure of impact would weakly contra-indicate passage of the statute.¹⁰⁸

On net, this analysis suggests the following legislative (in)action: 6 states should retain their current shall-issue laws; 1 state (California) should adopt the law for the first time; 26 states should repeal the law; and 19 states should continue to refrain from

¹⁰⁷ While crime did drop as the Texas shall-issue law went into effect on January 1, 1996, California has never adopted a shall-issue law, so the prediction is based on our weakly predictive model with 24 data points. But while California is shown by these tentative impact regressions to be a possible candidate for enactment of a shall-issue law, its experience of showing dramatic crime declines in the 1990s, both absolutely and vis a vis shall-issue states, suggests another reason for caution before accepting the predictions of these regressions.

¹⁰⁸ Conversely, in two states (Georgia and Oregon) the law was predicted (in Appendix Table 8) to have a deleterious impact while the more direct estimate (from Table 14) was beneficial (but statistically insignificant). Accordingly these two states might be deemed to be reclassified as weakly indicated for passage (although a risk averse analyst might think that a state showing either a deleterious predicted or actual estimate was contraindicated). These discrepancies between the predicted effects from the Table 16 regressions versus the more direct estimated effects from the full regressions that generated Table 14 underscores again the fragility of these results.

adopting.¹⁰⁹ For all of the reasons set forth above, this analysis can only be taken as suggestive of the type of nuanced policy recommendations that are possible from statistical analysis of state-specific estimates of the impact of the law. Nonetheless, we think it can provide a useful blueprint for researchers who have rich enough data and well-specified state-specific regression results to be able to make more finely grained policy recommendations than are possible from the typical aggregated analysis.

Conclusion

Judge Richard Posner has recently criticized moral philosophy for failing to persuade on any contentious issue.¹¹⁰ But a similar criticism might be made of quantitative empiricism. Readers might tend to accept only those quantitative analyses that resonate with their prior normative beliefs. Indeed, Judge Stephen Reinhardt famously proclaimed at a Yale seminar that social science had never affected his judicial decisionmaking. And Donald Braman and Dan Kahan have recently called upon econometricians like Lott and us to put away our statistical packages. In a piece, provocatively titled, “More Statistics, Less Persuasion,” Braman and Kahan argue that rather than “quantifying the impact of gun control laws on crime,” academics “should dedicate themselves to constructing a new expressive idiom that will allow citizens to debate the cultural issues that divide them.”¹¹¹

We disagree. A body of empiricism can over time disentangle thorny issues of causation and lead toward consensus. We view this paper as playing a role in this process (not ending the conversation). On net, we believe that Lott and Mustard’s efforts made an important contribution to the literature. They asked the initial question, amassed an important new panel data set, and then energetically and creatively analyzed it (and indeed their data set, which we know from experience was quite costly to construct, has been used by many researchers to explore this and other questions about crime). Unfortunately, their results have not withstood the test of time. When we added five years of county data and seven years of state data, allowing us to test an additional 14 jurisdictions that adopted shall-issue laws, the previous Lott and Mustard findings proved not to be robust. Importantly, we show that the Lott and Mustard results collapse when the more complete county data is subjected to less constrained jurisdiction-specific specifications or when the more complete state data is tweaked in plausible ways. No longer can any plausible case be made on statistical grounds that shall-issue laws are likely to reduce crime for all or even most states. How much farther one can go in arguing that shall-issue laws likely increase crime across the board or have heterogeneous

¹⁰⁹ Lott’s own analysis suggests that issuing an unlimited number of permits may produce a pernicious impact on crime. In his second edition, Lott estimated, in an aggregate specification, that when concealed handgun permits exceeded a maximum percentage of the population, additional permits would be predicted to increase crime. Lott (178-80). If Lott is right on this point, a state that passed the law under the foregoing normative account might be advised to cap the maximum number of permits issued.

¹¹⁰ Richard A. Posner, *The Problematics of Moral and Legal Theory* (1999).

¹¹¹ Donald Braman & Dan M. Kahan, *More Statistics, Less Persuasion: A Cultural Theory of Gun-Risk Perceptions at 1* (Working Paper 2002).

effects across states (albeit most commonly pernicious) will be matters about which various analysts will differ. We conclude with Learned Hand's admonition that an academic must "keep an open mind to every disconcerting fact, [and] an open ear to the cold voice of doubt.... You may not carry a sword beneath a scholar's gown."¹¹²

¹¹² L. HAND, *THE SPIRIT OF LIBERTY* 138 (1952), from 83 *Harvard Law Review* 366 December, 1969

Appendix 1 – Coding the Timing and Status of Shall-issue Laws

The data set that Lott and Mustard initially constructed covered the years from 1977 – 1992. Because of their fixed effects estimation technique, their analysis is able to measure the effect of the law only for those states that changed their legal status over this period. Hence, the coding of any state as either a shall-issue or non-shall-issue state will not influence the estimated effect on crime as long as the legal status persisted over the entire sample time period.¹¹³ Perhaps surprisingly, there are conflicts among the supporters and opponents of gun control legislation about whether various jurisdictions even have a shall-issue law or not. For example, the National Rifle Association characterizes Alabama and Connecticut as having shall-issue laws while the Brady Campaign to Prevent Gun Violence treat both states as having a more discretionary system of providing permits to carry concealed weapons (a so-called “may issue” law).¹¹⁴ Since neither of these laws officially changes status after 1977 (although query whether administrative enforcement patterns as well as citizen behavior concerning the purchase and carrying of handguns may have changed over this period), this dispute will have no bearing on the estimated effect of shall-issue laws. Nonetheless, there are numerous disagreements among different scholars about the timing of adoption of shall-issue laws that can influence these estimates.¹¹⁵ For purposes of replication, we generally tried to

¹¹³ Since the data from these states will influence the year effects and the estimated coefficients for the various explanatory variables, their inclusion in the analysis – as opposed to the coding of their shall-issue laws -- will have an indirect influence on the estimated effect of shall-issue laws.

¹¹⁴ See <http://www.handguncontrol.org/b-main.htm>. The NRA’s description of the relevant Alabama law suggests that demonstrating need is in fact a requirement for obtaining a concealed-carry permit (<http://www.nraila.org/research/riflaws.html>), which would seemingly support the coding advanced by the Brady Campaign. On the other hand, it might well be the case that the law that on the surface seems to be a “may issue” law was always, or came to be, administered as a shall-issue law. Indeed, between 1985 and the present, every Southern state from Texas to Virginia – with the single exception of Alabama - - adopted a shall-issue law, and it is possible that this lone exception to a universal Southern trend of adoption reflects the NRA’s recognition that in practice Alabama had indeed become a shall-issue state in practice even if it had not originally been one by virtue of statutory language.

The case of Connecticut is more complicated, and since no state bordering on Connecticut has enacted the law, one cannot draw inferences about the enforcement of the law from geography as we suggested for Alabama. In January 1978, there was debate over whether local police chiefs (who first need to approve any application to carry a pistol before state-wide approval can be sought) were being too stringent in rejecting the applications. At that point, nearly 50,000 Connecticut residents held state pistol permits, up from 27,628 in 1973. Apparently, concerns by the police about who was getting gun permits led local police chiefs to begin setting their own ground rules for determining who should get permits. Lincoln Millstein, “Police Toughen Criteria for Getting Gun Permit,” *Hartford Courant*, January 15, 1978, at 1.

¹¹⁵ While conceding that there are different interpretations of which states have shall-issue laws, Lott and Mustard indicate in footnote 32 that they follow the shall-issue law classification found in Clayton E. Cramer & David B. Kopel, “Shall-issue”: The New Wave of Concealed Handgun Permit Laws, 62 *Tenn. L. Rev.* 679, 691 (Spring 1995). Lott and Mustard cite two states in particular—Maine and Virginia—as potentially not “true” shall-issue states, though they state that their results are not affected by either

use shall-issue codings that were closer to those employed by Lott and Mustard. In light of our need to update the data beyond what Lott and Mustard's coding sources provided, we relied heavily on the Vernick coding as this was a very recent and thus up-to-date coding of shall-issue laws by independent researchers. Appendix Table 1 sets forth the list of states that have shall-issue laws under the coding of Lott and Vernick (and underscores some discrepancies in the Lott coding).

Choices not only have to be made about identifying when and if a state adopts a shall-issue law, but how to begin modeling its effect. Lott states that he assumes that the effect of the shall-issue law would emerge in the first year after the law takes effect. However, Lott coded the shall-issue law dummy in that fashion only for Florida and Georgia, with all other states being coded so that the effect of the law begins in the year of passage.¹¹⁶ Appendix Table 1 shows the passage dates of the various shall-issue laws that we employ when attempting to adhere to the Lott coding, and our differences with Lott and Mustard, and their own inconsistencies across models. The coding used by Vernick, which we also employ in our regressions, is set forth in column D. Our post-passage dummy, as well as our post-passage trends, are coded to begin in the year following the passage dates indicated in either column A or D of Appendix Table 1, depending on whether we are trying to follow Lott or Vernick coding.

Note that there is imprecision in these dates both because the statutes are not entirely clear about the precise legal status and because adoption of a shall-issue law does not perfectly equate with the actual enforcement of the law either within the state or over time, since enforcement could be quite different county-by-county and year-by-year. This problem will exist whenever one must characterize imprecisely defined statutes into sharply delineated discrete categories so that one can say when a state changes from one category to another. The difficulty is greatest when the statutory language of the shall-issue law invokes the command "shall" but then includes inherently discretionary criteria (such as a requirement for the "good moral character" of the permit recipient), or where the law says "may," but some counties or issuing authorities make it quite easy to obtain a permit. In either of these circumstances, there may be important differences between the law in practice and the law on the books, yet none of those nuances will be captured in our statistical analysis.¹¹⁷

redefining or dropping these states altogether (see Lott and Mustard, footnotes 33, 34, 35 and 49). Indeed, other scholars provide different dates of passage of the shall-issue laws for these two states than those offered by Cramer and Kopel. See Jon Vernick and Lisa Hepburn, "Description and Analysis of State and Federal Laws Affecting Firearm Manufacture, Sale, Possession, and Use, 1970-1999. (Brookings Paper, forthcoming 2002), see Table 5.

¹¹⁶ Ayres & Donohue (1999 at 449 note 21). Lott made different coding choices in his linear trend analysis, coding the enactment dates in Oregon, Pennsylvania, Virginia, and Philadelphia earlier than was proper.

¹¹⁷ See Jon Vernick and Lisa Hepburn, "Description and Analysis of State and Federal Laws Affecting Firearm Manufacture, Sale, Possession, and Use, 1970-1999. (Brookings Paper, forthcoming 2002). While there are enough classification discrepancies among the different authors that have tried to determine the presence of shall-issue laws that it becomes burdensome to probe all of the possible permutations, our efforts suggested that the aggregated results (which are weighted by state population,

thereby sharply reducing the impact of small states on the analysis) were not highly sensitive to these classification issues. Of course, any estimated effects of the shall-issue laws for individual states will be far more sensitive to the classification issues for these small states.