Fall 2011

The Impact of Right-to-Carry Laws and the NRC Report: Lessons for the Empirical Evaluation of Law and Policy

John J. Donohue, *Stanford Law School*

The Impact of Right-to-Carry Laws and the NRC Report: Lessons for the Empirical Evaluation of Law and Policy

Abhay Aneja and John J. Donohue III, *Stanford University*, and Alexandria Zhang, *Johns Hopkins University*

Send correspondence to: John J. Donohue III, School of Law, Stanford University, 559 Nathan Abbott Way, Stanford, CA 94305, USA; Fax: 650-723-4669; E-mail: donohue@law.stanford.edu.

For over a decade, there has been a spirited academic debate over the impact on crime of laws that grant citizens the presumptive right to carry concealed handguns in public—so-called right-to-carry (RTC) laws. In 2005, the National Research Council (NRC) offered a critical evaluation of the ''more guns, less crime'' hypothesis using county-level crime data for the period 1977–2000. Seventeen of the eighteen NRC panel members essentially concluded that the existing research was inadequate to conclude that RTC laws increased or decreased crime. The final member of the panel, though, concluded that the NRC's panel data regressions supported the conclusion that RTC laws decreased murder. We evaluate the NRC evidence and show that, unfortunately, the regression estimates presented in the report appear to be incorrect. We improve and expand on the report's county data analysis by analyzing an additional six years of county data as well as state panel data for the period 1977–2006. While we have considerable sympathy with the NRC's majority view about the difficulty of drawing conclusions from simple panel data models, we disagree with the NRC report's judgment that cluster adjustments to correct for serial correlation are not needed. Our randomization tests show that without such adjustments, the Type 1 error soars to 40–70%. In addition, the conclusion of the dissenting panel member that RTC laws reduce murder has no statistical support. Finally, our article highlights some important questions to

consider when using panel data methods to resolve questions of law and policy effectiveness. Although we agree with the NRC's cautious conclusion regarding the effects of RTC laws, we buttress this conclusion by showing how sensitive the estimated impact of RTC laws is to different data periods, the use of state versus county data, particular specifications, and the decision to control for state trends. Overall, the most consistent, albeit not uniform, finding to emerge from both the state and the county panel data models conducted over the entire 1977–2006 period with and without state trends and using three different models is that aggravated assault rises when RTC laws are adopted. For every other crime category, there is little or no indication of any consistent RTC impact on crime. It will be worth exploring whether other methodological approaches and/or additional years of data will confirm the results of this panel data analysis. (*JEL* K49, K00, C52)

1. Introduction

The debate on the impact of "shall-issue" or "right-to-carry" (RTC) concealed handgun laws on crime—which has now raged on for over a decade—demonstrates one of the many difficulties and pitfalls that await those who try to use observational data to estimate the effects of controversial laws.[1] John Lott and David Mustard initiated the "more guns, less crime" (MGLC) discussion with their widely cited 1997 article arguing that the adoption of RTC laws has played a major role in reducing violent crime. However, as Ayres and Donohue (2003b) note, Lott and Mustard's period of analysis ended just before the extraordinary crime drop of the 1990s. They concluded that extending Lott and Mustard's data set beyond 1992 undermined the MGLC hypothesis. Other studies have raised further doubts about the claimed benefits of RTC laws (e.g., see Black and Nagin, 1998; Ludwig, 1998).

But even as the empirical support for the Lott-Mustard thesis was weakening, its political impact was growing. Legislators continued to cite this work in support of their votes on behalf of RTC laws, and the MGLC claim has been invoked often in support of ensuring a personal right to have handguns under the Second Amendment. In the face of this scholarly and political ferment, in 2003, the National Research Council (NRC) convened a committee of top experts in criminology, statistics, and economics. Its purpose was to evaluate the existing data in hopes of reconciling the various methodologies and

---

1.   The term "RTC laws" is used interchangeably with "shall-issue laws" in the guns and crime literature.

findings concerning the relationship between firearms and violence, of which the impact of RTC laws was a single, but important, issue. With so much talent on board, it seemed reasonable to expect that the committee would reach a decisive conclusion on this topic, and put the debate to rest.

The bulk of the NRC report on firearms, which was finally issued in 2005, was uncontroversial. The chapter on RTC laws, however, proved to be extremely contentious. Citing the extreme sensitivity of point estimates to various panel data model specifications, the NRC report failed to narrow the domain of uncertainty about the effects of RTC laws. Indeed, it may have broadened it. However, while the NRC report concluded there was no reliable statistical support for the MGLC hypothesis, the vote was not unanimous. One dissenting committee member argued that the committee's own estimates revealed that RTC laws did in fact reduce the rate of murder. Conversely, a different member went even further than the majority's opinion by doubting that *any* econometric evaluation could illuminate the impact of RTC laws.

Given the prestige of the committee and the conflicting assessments of both the substantive issue of RTC laws' impact and the suitability of empirical methods for evaluating such laws, a reassessment of the NRC's report would be useful for researchers seeking to estimate the impact of other legal and policy interventions. Our systematic review of the NRC's evidence—its approach and findings—also provides important lessons on the perils of using traditional observational methods to elucidate the impact of legislation. To be clear, our intent is not to provide what the NRC panel could not—that is, the final word on how RTC laws impact crime. Rather, we show how fragile panel data evidence can be, and how a number of issues must be carefully considered when relying on these methods to study politically and socially explosive topics with direct policy implications.

The outline of this article is as follows. Section 2 offers background on the debate over RTC laws, and Section 3 describes relevant aspects of the NRC report in depth. Section 4 enumerates the critical flaws of the key results in the NRC report. Sections 5 and 6 explore two key econometric issues where the NRC panel may have erred—whether to control for state-specific trends and whether to adjust standard errors to account for serial or within-group correlation. Section 7 extends the analysis through 2006, and Section 8 offers improvements to the NRC model by revising the regression specification in accordance with past research on crime. Section 9 discusses the issue of whether the impact of RTC laws can be better estimated using

county- or state-level data. Section 10 delves further into three issues in this debate that merit special attention: the problem of omitted variable bias in assessing the impact of RTC laws (and in particular, the difficult-to-measure effect of the crack epidemic), the plausibly endogenous adoption of RTC legislation, and the relatively untouched issue of how RTC laws affect gun violence in particular. Section 11 offers concluding comments on the current state of the research on RTC laws, the difficulties in ascertaining the causal effects of legal interventions, and the dangers that exist when policy makers can simply pick their preferred study from among a wide array of conflicting estimates.

2. Background on the Debate

In a widely discussed 1997 article, "Crime, Deterrence, and Right-to-Carry Concealed Handguns," John Lott and David Mustard (1997) argued, based on a panel data analysis, that RTC laws were a primary driving force behind falling rates of violent crime. Lott and Mustard used county-level crime data (including county and year fixed effects, as well as a set of control variables) to estimate the impact of RTC laws on crime rates over the time period 1977–92. In essence, Lott and Mustard's empirical approach was designed to identify the effect of RTC laws on crime in the ten states that adopted them during this time period. Using a standard difference-in-difference model, the change in crime in the ten RTC states is compared with the change in crime in non-RTC states. The implicit assumption is that the controls included in the regression will explain other movements in crime across states, and the remaining differences in crime levels can be attributed to the presence or absence of the RTC laws.

Lott and Mustard estimated two distinct difference-in-difference-type models to test the impact of RTC laws: a dummy variable model and a trend, or "spline," model[2]. The "dummy model" tests whether the average crime level in the pre-passage period is statistically different from the post-passage crime level (after controlling for other factors). The "spline model" measures whether crime *trends* are altered by the adoption of RTC laws. Lott and Mustard noted that the

---

2. In the "dummy model," RTC laws are modeled as a dummy variable that takes on a value of 1 in the first full year after passage and retains that value thereafter (since no state has repealed its RTC law once adopted). In the "trend model," RTC laws are modeled as a spline variable indicating the number of years post-passage.

spline approach would be superior if the intervention caused a reversal in a rising crime rate. Such a reversal could be obscured in a dummy variable model that only estimates the average change in crime between the pre- and post-passage periods. An effective RTC law might show no effect in the dummy model if the rise in the pre-passage crime rate and the fall in the post-passage rate were to leave the average "before" and "after" crime levels the same.

In both regression models, Lott and Mustard included only a single other criminal justice explanatory variable—county-level arrest rates—plus controls for county population, population density, income, and thirty-six(!) categories of demographic composition. As we will discuss shortly, we believe that many criminological researchers would be concerned about the absence of important explanatory factors such as the incarceration rate and the level of police force.

Lott and Mustard's results seemed to support the contention that laws allowing the carry of concealed handguns lead to less crime. Their estimates suggested that murder, rape, aggravated assault, and overall violent crime fell by 4–7% following the passage of RTC laws. In contrast, property crime rates (auto theft, burglary, and larceny) were estimated to have increased by 2–9%. Lott and Mustard thus concluded that criminals respond to RTC laws by substituting violent crime with property crime to reduce the risk that they would be shot (since, according to them, victims are more often absent during the commission of a property crime). They also found that the MGLC contention was strengthened by the trend analysis, which ostensibly suggested significant decreases in murder, rape, and robbery (but no significant increases in property crime).

From this evidence, Lott and Mustard (1997) concluded that permissive gun-carrying laws deter violent crimes more effectively than any other crime reduction policy: "concealed handguns are the most cost-effective method of reducing crime thus far analyzed by economists, providing a higher return than increased law enforcement or incarceration, other private security devices, or social programs like early education." They went even further by claiming that had the remaining non-RTC states enacted such legislation, over 1,400 murders and 4,100 rapes would have been avoided nationwide, and that each new handgun permit would reduce victim losses by up to $5,000.

2.1. The Far-Reaching Impact of MGLC

The first "MGLC" article and Lott's subsequent research (and pro-gun advocacy) have had a major impact in the policy realm. Over the past decade,

politicians as well as interest groups such as the National Rifle Association have continually trumpeted the results of this empirical study to oppose gun control efforts and promote less restrictive gun-carrying laws. Lott relied on his own research to advocate for the passage of state-level concealed-carry gun laws, testifying on the purported safety benefits of RTC laws in front of several state legislatures, including Nebraska, Michigan, Minnesota, Ohio, and Wisconsin (Ayres and Donohue, 2003b).

The impact of the Lott-Mustard article can also be seen at the federal level. In 1997, ex-Senator Larry Craig (R-Idaho) introduced the Personal Safety and Community Protection Act with Lott's research as supporting evidence. This bill was designed to allow state nonresidents with valid handgun permits in their home state to possess concealed firearms (former football athlete Plaxico Burress sought to invoke this defense when he accidentally shot himself in a Manhattan nightclub with a gun for which he had obtained a Florida permit). According to Craig, Lott's work confirmed that positive externalities of gun carrying would result in two ways: by affording protection for law-abiding citizens during criminal acts and by deterring potential criminals from ever committing offenses for fear of encountering an armed response.[3] Clearly, Lott's work has provided academic cover for policy makers and advocates seeking to justify the view—on public safety grounds—that the Second Amendment confers a private right to possess handguns.

2.2. Questioning MGLC

Immediately after the publication of the Lott–Mustard article, scholars started raising serious questions about the theoretical and empirical validity of the MGLC hypothesis. For example, Zimring and Hawkins (1997) claimed that the comparison of crime between RTC and non-RTC states is inherently misleading because of factors such as poverty, drugs, and gang activity, which vary significantly across gun-friendly and non-gun-friendly

---

3. 143 CONG. REC. S5109 (daily ed. May 23, 1997) (statement of Sen. Craig). The bill was again introduced in 2000 by Congressman Cliff Stearns (R-Florida), who also cited Lott's work. 146 CONG. REC. H2658 (daily ed. May 9, 2000) (statement of Rep. Stearns). Indeed, this proposed legislation, now derisively referred to as "Plaxico's Law," is a perennial favorite of the NRA and frequently introduced by supportive members of Congress (Collins, 2009).

states (and are often difficult to quantify). To the extent that the relatively better crime performance seen in shall-issue states during the late 1980s and early 1990s was the product of these other factors, researchers may be obtaining biased impact estimates. Underscoring this point, Ayres and Donohue (2003b) pointed out that crime rose across the board from 1985 to 1992, and most dramatically in non-RTC states. Since the Lott-Mustard data set ended in 1992, it could not capture the most dramatic reversal in crime in American history. Figures 1–7 depict the trends of violent and property crimes over the period 1970–2007. For each of the seven crimes, the fifty states (plus DC) fall into four groupings: non-RTC states, states that adopted RTC laws over the period 1985–88 ("early adopters"), those that adopted RTC laws over the period 1989–91 ("mid-adopters"), and those that adopted RTC laws over the period 1994–96 ("late adopters"). The crime rate shown for each group is a within-group average, weighted by population. The figures corroborate Ayres and Donohue's point: crime rates declined sharply across the board beginning in 1992. In fact, there was a steady *upward* trend in crime rates in the years leading up to 1992, most distinctly for rape and aggravated assault. Moreover, the average crime rates in non-RTC states seemed to have dropped even more drastically than those in RTC states, which suggests that crime-reducing factors other than RTC laws were at work.

Ayres and Donohue (2003b) also recommended the use of a more general model, referred to as the "hybrid model," which essentially combined the dummy variable and spline models, to measure the immediate and long-run
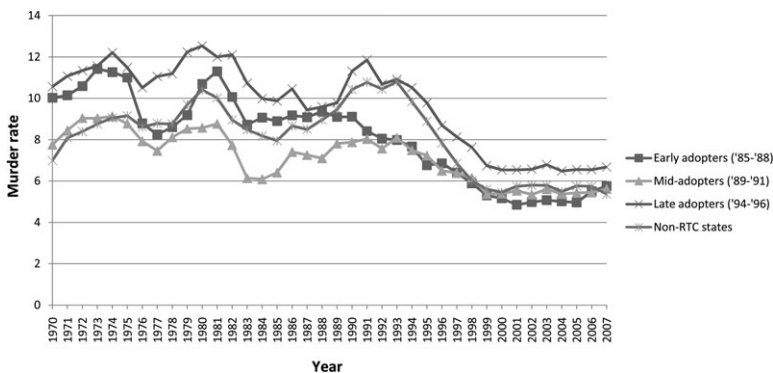
Figure 1. Murder Trends in RTC versus Non-RTC States—Weighted Average of Murder Rates per 100,000 Residents (1970–2007).
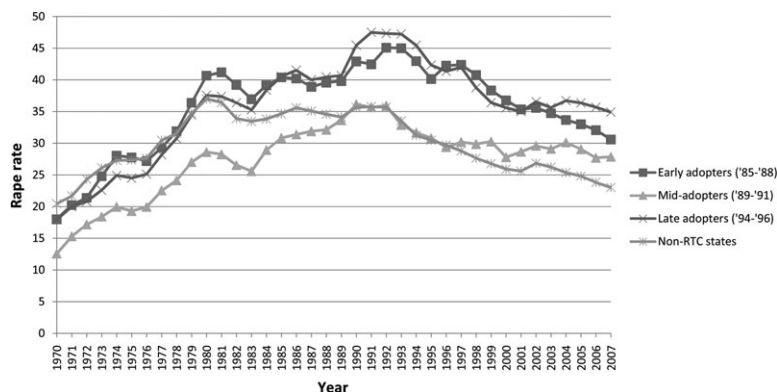
Figure 2. Rape Trends in RTC versus Non-RTC States—Weighted Average of Rape Rates per 100,000 Residents (1970–2007).

impact of RTC laws on crime. Since the hybrid model nests both the dummy and spline models, one can estimate the hybrid and generate either of the other models as a special case (depending on what the data show). This exercise seemed to weaken the MGLC claim. Their analysis of the county data set from 1977–1997 using the Lott-Mustard specification (revised to measure state-specific effects) indicated that RTC laws in aggregate *raised* total crime costs by as much as $524 million.

Just as Lott had identified a potential problem with the dummy model (it might understate a true effect if crime followed either a V-shaped or an

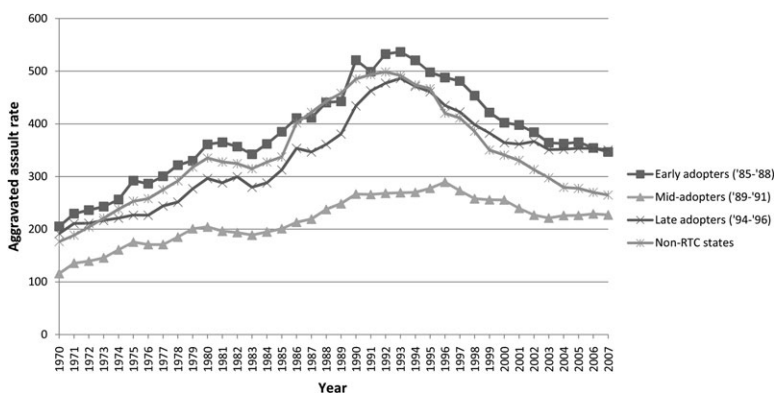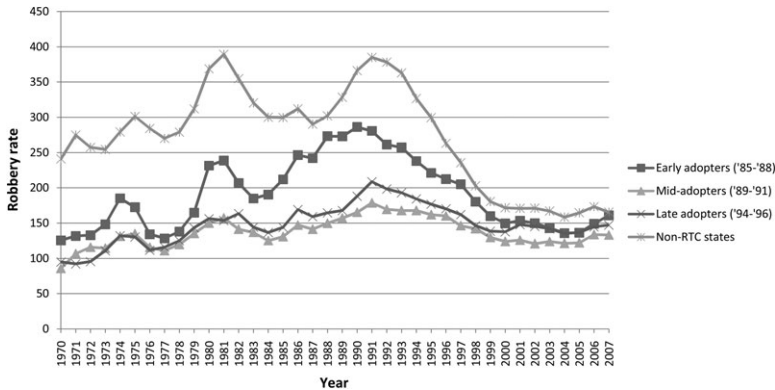**Figure 3.** Assault Trends in RTC versus Non-RTC States—Weighted Average of Assault Rates per 100,000 Residents (1970–2007).

Figure 4. Robbery Trends in RTC versus Non-RTC States—Weighted Average of Robbery Rates per 100,000 Residents (1970–2007).

inverted V-shaped pattern), there is a potential problem with models (such as the spline and the hybrid models) that estimate a post-passage linear trend. Early adopters of RTC laws have a far more pronounced impact on the trend estimates of RTC laws than later adopters since there may only be a few years of post-passage data available for a state that adopts RTC laws close to the end of the data period. If those early adopters were unrepresentative of low-crime states, then the final years of the spline estimate would suggest a dramatic drop in crime, not because crime had in fact fallen in adopting states but because the more representative states had dropped out of the es-timate (since there would be no post-passage data after, say, three years for
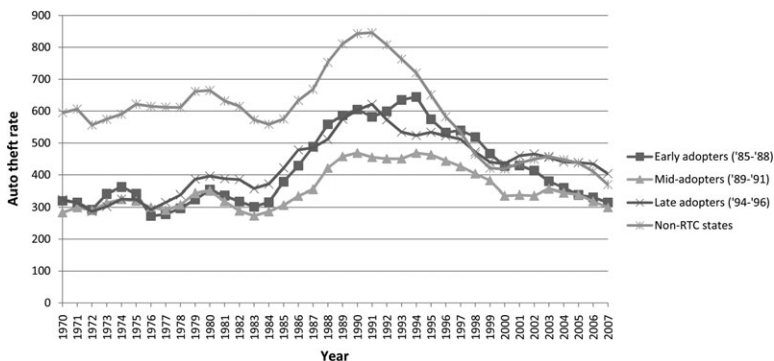

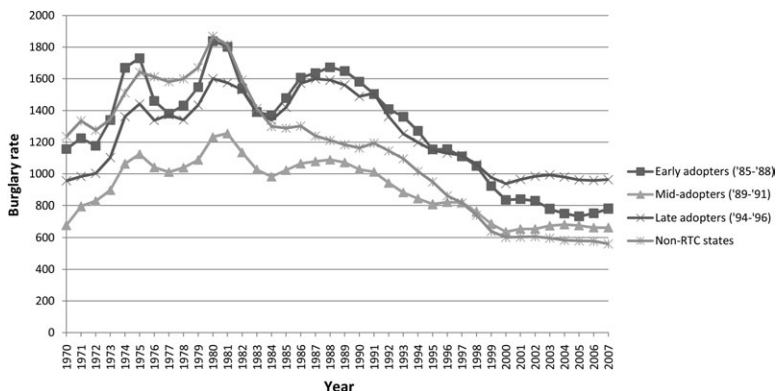**Figure 5.** Auto Theft Trends in RTC versus Non-RTC States—Weighted Average of Auto Theft Rates per 100,000 Residents (1970–2007).

Figure 6. Burglary Trends in RTC versus Non-RTC States—Weighted Average of Burglary Rates per 100,000 Residents (1970–2007).

a state that had adopted the RTC law only three years earlier, but there would be such data for Maine, Indiana, and North Dakota, which were the earliest RTC adopters). We recognize that each model has limitations, and present the results of all three in our tables below.[4]

3. Findings of the NRC

The sharply conflicting academic assessments of RTC laws specifically and the impact of firearms more generally, not to mention the heightened political salience of gun issues, prompted the NRC to impanel a committee of experts to critically review the entire range of research on the relationships between guns and violence. The blue-chip committee, which included prominent scholars such as sociologist Charles Wellford (the committee chair), political scientist James Q. Wilson, and economists Joel Horowitz, Joel Waldfogel, and Steven Levitt, issued its wide-ranging report in 2005.

While the members of the panel agreed on the major issues discussed in eight of the nine chapters of the NRC report, the single chapter devoted to

---

4. We note that in the latest version of his book, Lott (2010) criticizes the hybrid model, but he fails to appreciate that the problem with the hybrid model—and with the spline model he prefers—is that they both yield estimates that are inappropriately tilted down as the more representative states drop out of the later years, which drive the post-passage trend estimates. An apples-and-apples comparison that included the identical states to estimate the post-passage trend would not suggest a negative slope. This is clear in Figure 1 and Table 1 of Ayres and Donohue (2003b).

Figure 7. Larceny Trends in RTC versus Non-RTC States—Weighted Average of Larceny Rates per 100,000 Residents (1970–2007).

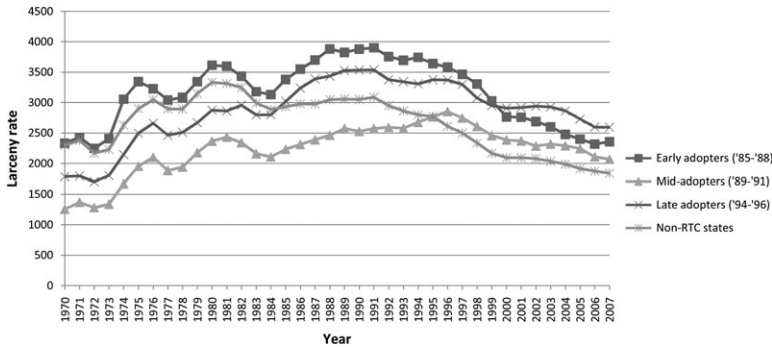exploring the causal effects of RTC laws on crime proved to be quite contentious. After reviewing the existing (and conflicting) literature and undertaking their own evaluation of Lott's county-level crime data, seventeen of the eighteen committee members concluded that the data provided no reliable and robust support for the Lott-Mustard contention. In fact, they believed the data could not support any policy-relevant conclusion. In addition, they claimed they could not estimate the true impact of these laws on crime because (1) the empirical results were imprecise and highly sensitive to changes in model specification and (2) the estimates were not robust when the data period was extended eight years beyond the original analysis (through 2000), a period during which a large number of states adopted the law.

One can get an inkling of the NRC majority's concern about model sensitivity by examining Table 2a (which we will discuss in detail in Section 4.2), which reports estimates from the NRC report on the impact of RTC laws on seven crimes. The estimates are based on the Lott and Mustard (1997) dummy and spline models and county data for the period 1997–2000. The vastly different results produced by the two models gave the majority considerable pause. For example, if one believed the dummy model, then RTC laws considerably *increased* aggravated assault and robbery, while the spline model suggested RTC laws *decreased* the rate of both of these crimes.

The tension created by conflicting estimates was epitomized by the intra-panel dissention, as two members of the committee wrote separately on the NRC's evaluation of RTC laws. One sought to refute the majority's skepticism, and one sought to reinforce it. Noted political scientist James Q. Wilson offered the lone dissent to the committee's report, claiming that Lott and

Mustard's MGLC finding actually held up under the panel's reanalysis. Specifically, Wilson rejected the majority's interpretation of the regression estimates seen in Table 2a. Although the panel noted that the RTC impact estimates disagreed across their two models (dummy and spline) for six of the seven crime categories, Wilson emphasized the similar finding of murder rate declines in the two models. The agreement in the murder estimates led him to heartily endorse the MGLC view. Indeed, after dismissing articles that had cast doubt on the MGLC hypothesis (such as Black and Nagin, 1998), on the grounds that they were "controversial," Wilson concluded: "I find the evidence presented by Lott and his supporters suggests that RTC laws do in fact help drive down the murder rate, though their effect on other crimes is ambiguous" (NRC, 2005, p. 271).

The committee penned a response to Wilson's dissent (separate from its overall evaluation of RTC legislation), which stressed that the only disagreement between the majority and Wilson (throughout the entire volume on gun issues) concerned the impact of RTC laws on murder. They noted that, while there were a number of negative estimates for murder using the Lott-Mustard approach, there were also several positive estimates that could not be overlooked. In addition, even the results for murder failed to support the MGLC contention when restricting the period of analysis to five years or less after law adoption.[5] The important task was to try to reconcile these contradictions—and the panel majority believed that was not possible using the existing data.

Committee member (and noted econometrician) Joel Horowitz was the ardent skeptic, and not without merit. Horowitz joined the refutation of Wilson but also authored his own appendix discussing at length the difficulties of measuring the impact of RTC laws on crime using observational rather than experimental data.[6] He began by addressing a number of flaws in the panel data approach. First, if factors other than the adoption of the RTC law change but are not controlled for in the model, then the resulting estimates would not effectively isolate the impact of the law (we demonstrate the

---

5. The importance of this restriction on the post-passage data was mentioned earlier: As states dropped out of the post-passage data, the estimated impact of RTC laws became badly biased (since one was no longer deriving the estimated effect from a uniform set of states).

6. While his chapter is directed at the analysis of RTC laws, Horowitz's comments applied to an array of empirical studies of policy that were discussed throughout the entire NRC volume.

likelihood of this possibility in Section 10). Second, if crime increases before the adoption of the law at the same rate it decreases after adoption, then a measured zero difference would be misleading. The same problem arises for multiyear averages. Third, the adoption of RTC laws may be a *response* to crime waves. If such an endogeneity issue exists, the difference in crime rates may merely reflect these crime waves rather than the effect of the laws. Lastly, as even Lott (2000) found in his data, RTC states differ noticeably from non-RTC states (e.g., RTC states are mainly Republican and had low but rising rates of crime). It would not be surprising if these distinctive attributes influence the measured effect of RTC laws. In this event, looking at the impact of RTC laws in current RTC states may not be useful for predicting impact if they are adopted in very different states.

Ideally, states would be randomly selected to adopt RTC laws, thereby eliminating the systematic differences between RTC states and non-RTC states. In the absence of such randomization, researchers introduce controls to try to account for these differences, which generates debate over which set of controls is appropriate. Lott (2000) defended his model by claiming that it included "the most comprehensive set of control variables yet used in a study of crime" (p. 153). We show here that this claim is gravely outdated. Moreover, Horowitz noted that not only are the data limited for these variables, it is also possible to control for too many variables—or too few. He pointed out that Donohue (2003) found a significant relationship between crime and *future* adoption of RTC legislation, suggesting the likelihood of omitted variable bias and/or the endogenous adoption of the laws. Horowitz concludes by noting that there is no test that can determine the right set of controls: "it is not possible to carry out an empirical test of whether a proposed set of X variables is the correct one . . . it is largely a matter of opinion which set [of controls] to use" (NRC, 2005, p. 307). Noting the likelihood of misspecification in the evaluation of RTC laws, and that estimates obtained from a misspecified model can be highly misleading, he concluded that there was little hope of reaching a scientifically supported conclusion based on the Lott-Mustard/NRC model.

3.1. The Serious Need for Reassessment

The story thus far has been discouraging for those hoping for illumination of the impact of legislation through econometric analysis. If the NRC majority is right, then years of observational work by numerous researchers,

topped off with a multiyear assessment of the data by a panel of top scholars, were not enough to pin down the actual impact of RTC laws. However, given that the panel only presented estimates based on the Lott-Mustard (1997) approach (except for a sparse model with no covariates, which we describe in Section 4), it is possible the committee overlooked quantitative models and potentially useful evidence that could have influenced their view on the topic. If Horowitz is right, then the entire effort to estimate the impact of state RTC policies from observational data is doomed. Indeed, there may be simply too much that researchers do not know about the proper structure of econometric models of crime. Notably, however, the majority did not join Horowitz in the broad condemnation of all observational microeconometrics for the study of this topic. Perhaps a model that better accounts for all relevant, exogenous, crime-influencing factors and secular crime trends could properly discern the effects of RTC laws. As we show below, a number of plausible explanations and factors were excluded from the committee's examination.

4. Attempts to Replicate the NRC Findings

Previous research on guns and crime has shown how data and methodological flaws can produce inaccurate conclusions. In a follow-up to their initial 2003 *Stanford Law Review* article, Ayres and Donohue (2003a) showed how coding errors can yield inaccurate estimates of the effect of RTC laws on crime. Commenting on a study in support of the MGLC premise by Plassman and Whitley (2003), Ayres and Donohue (2003a) described numerous coding flaws. After correcting these errors, the evidence supporting the MGLC hypothesis evaporated.

4.1. Panel Data Models with No Covariates

Since the NRC panel based their reported estimates on data provided by John Lott, we thought it prudent to carefully examine the NRC committee's own estimates. We first attempt to replicate the results of the report using the NRC 1977–2000 county data set, which the committee supplied to us. We begin with the committee's no-controls model, which, apart from the dummy and trend variables, only includes year and county fixed effects. The reported NRC estimates are presented in Table 1a, and the first two rows of Table 1b show our efforts at replicating them. While the estimates of the dummy variable model are reasonably close, the trend estimates are not at all

comparable: The sign on the estimates in the spline model switches when going from Table 1a to Table 1b for all crimes except auto theft. Table 1b also includes our own estimates from the more flexible version of these specifications—the hybrid model—which combines the dummy and trend approaches. In other words, taken at face value, Table 1b tells us that crime clearly worsened for six or seven crime categories after the passage of RTC laws, regardless of whether one used the dummy variable, spline, or hybrid models.

Table 1a. Estimated Impact of RTC Laws—Published NRC Estimates—No Controls, All Crimes, 1977–2000 (County Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−1.95	17.91***	12.34***	19.99***	23.33***	19.06***	22.58***
	1.48	1.39***	0.90***	1.21***	0.85***	0.61***	0.59***
2. Spline model	0.12	−2.17***	−0.65***	−0.88***	0.57***	−1.99***	−0.71***
	0.32	0.30***	0.20***	0.26***	0.19***	0.13***	0.13***

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 1b. Estimated Impact of RTC Laws—Using NRC County Data—No Controls, All Crimes, 1977—2000[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−2.58	18.40***	12.60***	19.70***	22.80***	19.00***	22.60***
	1.87	2.29***	1.40***	1.75***	1.69***	1.24***	1.08***
2. Spline model	−0.57*	2.36***	1.52***	2.43***	3.17***	2.23***	3.01***
	0.34*	0.39***	0.25***	0.31***	0.30***	0.24***	0.22***
3. Hybrid model							
Post-passage dummy	−0.06	16.20***	11.90***	17.40***	16.80***	17.70***	18.50***
	2.33	2.22***	1.69***	1.88***	1.86***	1.34***	1.20***
Trend effect	−0.56	0.58	0.22	0.51	1.32***	0.28	0.98***
	0.43	0.40	0.30	0.35	0.35***	0.27	0.25***

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 1c. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—No Controls, All Crimes, 1977–2000 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−2.20 1.87	27.80*** 3.53***	16.40*** 2.16***	19.50*** 2.06***	23.90*** 2.27***	22.80*** 2.06***	28.10*** 2.29***
2. Spline model	0.68** 0.28**	4.65*** 0.46***	4.31*** 0.26***	3.18*** 0.27***	4.72*** 0.28***	5.06*** 0.25***	6.02*** 0.27***
3. Hybrid model							
Post-passage dummy	−7.99*** 2.19***	12.00*** 3.08***	−3.50 2.72	8.91*** 2.32***	5.50** 2.70**	1.44 2.60	3.26 2.98
Trend effect	1.34*** 0.33***	3.66*** 0.37***	4.60*** 0.32***	2.44*** 0.30***	4.27*** 0.32***	4.94*** 0.31***	5.75*** 0.35***

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates.
*Significant at 10%; **significant at 5%; ***significant at 1%.

We contacted the committee to see if we might be able to understand why the efforts at replication were failing, but the files for reproducing their results and tables had not been retained.[7] Thus, we thought it wise to analyze county-level data by constructing our own data set, which we will refer to as the "updated 2009 data set." We create the same variables found in Lott's data—crime rates, demographic composition, arrest rates, income, population, and population density—and extend our new set as far forward as the data are available—2006 (the NRC data ended in 2000).[8] This data extension also gives us an opportunity to explore how the NRC's results are

---

7.   In an attempt to reconcile the divergence, we initially speculated that perhaps the NRC committee did not weight its panel data regressions by county population as we do throughout, but this turned out not to explain the difference. Our best guess is that the NRC did weight the regression by population since they essentially adopted the Lott and Mustard (1997) approach. We also determined that the NRC data set was missing all county identifiers for 1999 and 2000, so we speculated that this might explain the results (since data for any year with a missing country identifier would be omitted from the regression). Again, we could not replicate the NRC spline model results of Table 1a, whether we included all years of data or dropped 1999 and 2000.

8.   We also add 0.1 to all zero crime values before taking the natural log in our county-level data set, as the NRC did.

affected when using the most current data available. As we will see in Section 7, the additional years of data will also enable us to estimate the effect of six additional state adoptions of RTC laws, not present in the NRC analysis: Michigan (2001), Colorado (2003), Minnesota (2003), Missouri (2003), New Mexico (2003), and Ohio (2004).[9]

We obtained our crime data from the University of Michigan's Interuniversity Consortium for Political and Social Research, which maintains the most comprehensive collection of Uniform Crime Reports (UCR) data. Unfortunately, county-level crime data for 1993 are currently unavailable. The National Archive of Criminal Justice Data recently discovered an error in the crime data imputation procedure for 1993 and, for this reason, has made 1993 data inaccessible until the error has been corrected. Thus, for all of the following tables with estimates using our updated data, we are missing values for 1993.

Table 1c reproduces Table 1b using our own newly constructed data set (with 1993 omitted). In the case of every crime-model permutation, the use of this new data set further weakened the crime-reducing effects of RTC laws.[10] The bottom line is that (1) we cannot replicate the NRC no-controls estimates of Table 1a whether we use our own newly constructed county data or the data used by the NRC committee and (2) the best estimates in the no-controls model overwhelmingly show that all crime was *higher* after RTC laws adoptions.

4.2. Panel Data Models with Covariates

After failing to replicate the NRC "no-covariates" model, we next undertook the same replication exercise with the "covariates" model, which adds to the county and year fixed effects model the following Lott-Mustard explanatory variables: arrest rate, county population, population density, real per capita income variables, and thirty-six variables designed to capture the

---

9.   Kansas and Nebraska adopted RTC laws in 2006, which is too late to be captured in our analysis, since we assume a state to be an "RTC state" beginning in the first full year after a law's passage.

10.   Table 1c differs from Table 1b in two respects—it uses our new data set instead of the NRC, and it omits 1993 data. To see how important the 1993 omission is, we reproduced Table 1b (using the NRC data) dropping that year, which turned out to have little effect on the estimates.

county's demographic composition.[11] Although we have already noted Lott's claim that this is "the most comprehensive set of control variables yet used in a study of crime," in fact, this set of variables omits many important influences on crime, which we will reintroduce in Section 8.

To be clear about our approach, we use annual county-level crime data (and later, state-level data) for the United States from 1977 through either 2000 (to conform to the NRC report) or 2006 (the last year for which data are available). We explore the impact of RTC laws on seven Index I crime categories by estimating the reduced-form regression:

$$Y_{it} = \eta \text{RTC}_{jt} + \alpha_i + \theta_t + \beta_{jt} + \gamma X_{ijt} + \epsilon_{it}, \tag{1}$$

where the dependent variable Y_{it} denotes the natural log of the individual violent and property crime rates for county i and year t. Our explanatory variable of interest—the presence of an RTC law within state j in year t—is represented by RTC_{jt}. The exact form of this variable shifts according to the three variations of the model we employ (these include the Lott-Mustard dummy and spline models, as well as the Ayres and Donohue hybrid model).[12]

The variable α_i indicates county-level fixed effects (unobserved county traits) and θ_t indicates year effects. As we will discuss below, there is no consensus on the use of state-specific time trends in this analysis, and the NRC report did not address this issue. Nevertheless, we will explore this possibility, with β_{jt} indicating state-specific trends, which are introduced in selected models. Since neither Lott and Mustard (1997) nor the NRC (2005) examines state

---

11.   The NRC uses the Lott-Mustard method of calculating arrest rates, which is the number of arrests for crimes divided by the contemporaneous number of crimes. Econometrically, it is inappropriate to use this contemporaneous measure since it leaves the dependent variable on both sides of the regression equation (a better approach would lag this variable one year, as discussed in Ayres and Donohue, 2009). Another issue about the arrest rates is unclear: The NRC report does not indicate whether it uses the individual Index I crime categories to compute arrest rates, or alternatively, if they use the broad categories of violent and property crimes, as has been used in recent articles (Moody and Marvell, 2008). We adopt this latter approach for all tables in this article, although we also explored the possibility of arrest rates for individual crimes. Regardless of which arrest rate we used, our estimates still diverged considerably from the estimates presented by the NRC.

12.   As noted previously, in the dummy variable approach, the RTC variable is a dichotomous indicator that takes on a value of 1 in the first full year that a state j has an RTC law. In the spline model, the RTC variable indicates the number of post-passage years. The hybrid specification contains both dummy and trend variables.

trends, this term is dropped when we estimate their models. The term X_{ijt} represents a matrix of observable county and state characteristics thought by researchers to influence criminal behavior. The components of this term, however, vary substantially across the literature. For example, while Lott uses only "arrest rates" as a measure of criminal deterrence, we discuss the potential need for other measures of deterrence, such as incarceration levels or police presence, which are measured at the state level.

In Tables 2a–c, we follow the same pattern as that of Tables 1a–c: We begin by showing the NRC published estimates (Table 2a) and then show our effort at replication using the NRC data set (Table 2b). We then show the estimates obtained from our reconstruction of the county data set from 1977 through 2000 (Table 2c, which omits 1993 data).[13] The basic story that we saw above with respect to the no-covariates model holds again: We cannot replicate the NRC results using the NRC's own data set (compare Tables 2a and b), and omitting 1993 data does not make a substantive difference. Once again, our Table 2c estimates diverge wildly from the Table 2a estimates, which appeared in the NRC report. As we will see in a moment, the results that Professor Wilson found to be consistent evidence of RTC laws reducing murder (see Table 2a) were probably inaccurate (see Table 2c).

Table 2a. Estimated Impact of RTC Laws—Published NRC Estimates—Lott-Mustard Controls, All Crimes, 1977–2000[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−8.33*** 1.05***	−0.16 0.83	3.05*** 0.80***	3.59*** 0.90***	12.74*** 0.78***	6.19*** 0.57***	12.40*** 0.59***
2. Spline model	−2.03*** 0.26***	−2.81*** 0.20***	−1.92*** 0.20***	−2.58*** 0.22***	−0.49*** 0.13***	−2.13*** 0.19***	−0.73*** 0.13***

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

_____

13. Once again, we explored whether omitting 1993 data had an impact on the results, and again our Table 2 estimates looked quite similar when 1993 data were dropped.

Table 2b. Estimated Impact of RTC Laws—Using NRC Data—with Lott-Mustard Controls, All Crimes, 1977–2000[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−3.80*	10.50***	11.20***	11.20***	16.80***	11.00***	17.60***
	2.14*	2.18***	1.55***	1.81***	1.54***	0.98***	0.86***
2. Spline model	−0.61	1.38***	1.91***	1.63***	2.61***	1.62***	3.12***
	0.38	0.36***	0.25***	0.32***	0.29***	0.19***	0.17***
3. Hybrid model							
Post-passage dummy	−2.51	9.77***	7.01***	9.02***	12.20***	8.92***	9.72***
	2.63	2.28***	1.76***	1.92***	1.74***	1.06***	0.94***
Trend effect	−0.30	0.18	1.05***	0.53	1.11***	0.52**	1.92***
	0.47	0.36	0.27***	0.33	0.34***	0.22**	0.19***

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 2c. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—with Lott-Mustard Controls, All Crimes, 1977–2000 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−3.80**	9.82***	8.96***	5.44***	13.60***	4.36***	12.90***
	1.87**	2.74***	1.34***	1.45***	1.40***	0.95***	0.88***
2. Spline model	−0.26	0.48	1.10***	0.26	1.50***	0.30**	1.16***
	0.28	0.33	0.18***	0.21	0.19***	0.15**	0.14***
3. Hybrid model							
Post-passage dummy	−3.98*	11.40***	6.34***	6.39***	10.60***	4.53***	11.80***
	2.22*	2.62***	1.48***	1.66***	1.57***	1.05***	0.94***
Trend effect	0.04	−0.38	0.63***	−0.23	0.70***	−0.04	0.28*
	0.33	0.30	0.20***	0.25	0.22***	0.16	0.15*

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

4.3. Potential Problems with the NRC Models and Data

Before turning to the implications of the errors in the NRC estimates, we note a few small errors in the NRC data that we corrected in all our tables. First, we identified an extraneous demographic variable that caused a substantial number of observations to drop from the NRC data set (over 20,000).[14] We do not know if the committee dropped this variable before conducting its analysis, but we drop it in our own analysis.[15] Second, Philadelphia's year of adoption is coded incorrectly—as 1995 instead of 1996. Third, Idaho's year of adoption is coded incorrectly—as 1992 instead of 1991. Fourth, the area variable, which is used to compute county density, has missing data for years 1999 and 2000.[16]

The major differences in Table 2a (the NRC committee's estimates) and Table 2c (what we think is the best estimate of what the NRC intended to present) are profound enough that they might well have changed the nature of the report. Recall that Wilson had looked at the NRC's results (Table 2a) and decided that since the dummy and spline estimates were both consistent and statistically significant for only one crime—murder—these were the only estimates that should be accepted. But applying this same logic to the Table 2c estimates would lead to the drastically different conclusion that for four crimes—aggravated assault, auto theft, burglary, and larceny—Table 2c provides uniform evidence that

---

14.   The variable is called "ppnpermpc." We stumbled into using this variable as we tried to incorporate Lott and Mustard's thirty-six demographic variables, which denote the percentage of each county's population that falls into each of six age-groups based on three racial categories for men and for women. Twelve of these variables begin with the prefix "ppn," which will then be included in the analysis if one uses a STATA command that groups together all variables with this common "ppn" prefix. For example, "ppnm2029" indicates the percentage of a county population that is male and neither white nor black. We do not know how the ppnpermpc variable fits into this grouping (or even if it is meant be a part of this group of variables). The mean value of this variable is –3.206657, with the individual observations ranging from –12.05915 to 4.859623. While the other ppn variables reflect some sort of percentage, the mean negative value obviously indicates that this variable is not a percentage.

15.   We found that whether or not we include this variable, we cannot replicate the NRC's results (in Table 2a).

16.   Because the NRC area numbers are the same for a county across all years, we fill in this gap by simply using the 1998 values for these two years. (However, we note that area should not be constant across all years, as the Census updates these data every decade.) We include complete, updated area data in our new data set.

RTC laws *increase* crime (while the evidence for the other crimes is mixed). One might go further and say that all the Table 2c dummy and spline estimates show crime *increases*, except for murder.

Although we speculate that Table 2c reflects where the NRC panel should have ended up if it had wanted to repeat Lott and Mustard's county data analysis, there is actually far more that the committee could have done to go beyond Table 2c to test the validity of the MGLC premise. We emphasize, though, that this is not necessarily a strong criticism of the NRC majority since it concluded (in our view, correctly) that the evidence was already too fragile to draw strong conclusions, and further support for this assessment would merely have been cumulative. Nevertheless, we now turn to some avenues of inquiry that Wilson might have considered before adopting the Lott and Mustard (1997) conclusion vis-à-vis murder.

5. Debate over the Clustering of Standard Errors

5.1. Is Clustering Necessary?

To this point we have said little about the important question of estimating the standard errors in panel-data regressions. The estimates presented thus far follow the NRC in providing heteroskedasticity-robust standard errors. Research has found, though, that the issue of whether to "cluster" the standard errors has a profound impact on assessments of statistical significance. This issue gained prominence beginning primarily with a 1990 article by Brent Moulton. Moulton (1990) pointed to the possible need for the clustering of observations when treatments are assigned at a group level. In such cases, there is an additive source of variation that is the same for all observations in the group, and ignoring this unique variation leads to standard errors that are underestimated. Lott, however, suggests that clustered standard errors are not needed (Lott, 2004), claiming that county-level fixed effects implicitly control for state-level effects, and therefore, clustering the standard errors on state is unnecessary.

On this point, the NRC committee (2005) sided with Lott, stating that "there is no need for adjustments for state-level clustering" (p. 138). However, we *strongly* believe the committee was mistaken in this decision. One must account for the possibility that county-level disturbances may be correlated within a state during a particular year by clustering the standard errors by state. There is also a second reason for clustering that the NRC report

did not address. Specifically, serial correlation in panel data can lead to major underestimation of standard errors. Indeed, Bertrand, Duflo, and Mullainathan (2004) point out that even the Moulton correction alone may be insufficient for panel data estimators that utilize more than two periods of data due to autocorrelation in both the intervention variable and the outcome variable of interest. Wooldridge (2003, unpublished manuscript), as well as Angrist and Pischke (2009), suggest that clustering the standard errors by state (along with heteroskedasticity-robust standard errors) will help address this problem, and at least provide a lower bound on the standard errors.

5.2. Using Placebo Laws to Test the Impact of Clustering

Our reading of the influential literature on this issue suggests to us that clustering would make a major difference in the results generated by the Lott and Mustard models that the NRC report adopted in its analysis. But who is correct on the clustering issue—Lott, Mustard, and the NRC panel on the one hand, or Angrist, Pischke, and several other high-end applied econometricians on the other? To address this important question, we run a series of placebo tests. In essence, we randomly assign RTC laws to states, and reestimate our model iteratively (1,000 times), recording the number of times that the variable(s) of interest are "statistically significant." For this experiment, we use our most flexible model: the hybrid model (that incorporates both a dummy and a trend variable) with the controls employed by the NRC.

We run three versions of this test. First, we first generate a placebo law in a random year for all fifty states and the District of Columbia. Once the law is applied, it persists for the rest of our data period, which is how laws are coded in the original analysis. In our second test, we apply a placebo law in a random year to the thirty-two states that actually implemented RTC laws during the period we are analyzing. The remaining nineteen states assume no RTC law. Finally, we randomly select thirty-two states to receive a placebo law in a random year. The results of these three tests are presented in Table 3a.

Given the random assignment, one would expect to reject the null hypothesis of no effect of these randomized "laws" roughly 5% of the time if the standard errors in our regressions are estimated correctly. Instead, the table reveals that the null hypothesis is rejected 50–70% of the time for murder and robbery with the dummy variable and even more frequently with the trend variable (60–74%). Clearly, this exercise suggests that the standard errors used in the NRC report are far too small.

Table 3a. Hybrid Model—Percentage of Significant Estimates (at the 5% Level)—Using Updated 2009 County-Level Data—Lott-Mustard Controls, without Clustered Standard Errors, 1977–2006 (without 1993 Data)[a]

		Dummy Variable (%)	Trend Variable (%)
1. All 50 states + DC	Murder	50.2	67.4
	Robbery	56.7	65.6
2. Exact 32 states	Murder	64.2	71.9
	Robbery	59.8	67.2
3. Random 32 states	Murder	57.8	59.9
	Robbery	70.6	74.2

[a]Simulation based on NRC with-controls model, which, similar to above estimations, includes year and county fixed effects, and weighting by county population. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.

Table 3b. Hybrid Model—Percentage of Significant Estimates (at the 5% Level)—Using Updated 2009 County-Level Data—Lott-Mustard Controls, with Clustered Standard Errors, 1977–2006 (without 1993 Data)[a]

		Dummy Variable (%)	Trend Variable (%)
1. All 50 states + DC	Murder	8.9	11.5
	Robbery	8.1	8.1
2. Exact 32 states	Murder	10.0	11.0
	Robbery	9.2	7.1
3. Random 32 states	Murder	11.2	13.5
	Robbery	10.3	8.8

[a]Simulation based on NRC with-controls model, which, similar to above estimations, includes year and county fixed effects, and weighting by county population. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.

Table 3b replicates the exercise of Table 3a, but now uses the cluster correction for standard errors (on state). Table 3b suggests that clustering standard errors does not excessively reduce significance, as the NRC panel feared. In fact, the percentages of "significant" estimates produced in all three versions of the test still lie well beyond the 5% threshold. Similar results are found when we replicate Tables 3a and b while employing the dummy model instead of the hybrid model (we do not show those results here). All these tests show that if we do not cluster the standard errors, the likelihood of obtaining significant estimates is astonishingly (and unreasonably) high. The conclusion we draw from this exercise is that clustering is clearly needed to adjust the standard

errors in these panel data regressions. Accordingly, we will use this clustering adjustment for all remaining regressions in this article.

5.3. Does Clustering Influence the Results?

To get a sense of how clustering would have changed the NRC's estimates, we run the NRC model with standard errors clustered on state using our county-level data. Table 4 shows that clustering the standard errors in this model eliminates most of the statistical significance we saw in Table 2c (the same model but without clustering). Importantly, the significance of the negative coefficients for murder disappears. On this basis, one might suspect that had this set of results been used, the conclusions of the panel may have been quite different. These estimates—which we believe are now more accurate—provide no support for the claim that RTC laws reduce crime and, in fact, reveal evidence that aggravated assault, auto theft, and larceny all rise by between 9 and 14%. While this might suggest that RTC laws *increase* crime, the auto theft and larceny results do not readily comport with any plausible theory about the impact of RTC laws, and so we would proceed

Table 4. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—with Lott-Mustard Controls, with Clustered Standard Errors, All Crimes, 1977–2000 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−3.80	9.82	8.96*	5.44	13.60**	4.36	12.90***
	6.25	11.20	5.33*	5.53	5.83**	3.58	3.97***
2. Spline model	−0.26	0.48	1.10	0.26	1.50*	0.30	1.16
	0.80	1.22	0.81	0.85	0.83*	0.50	0.82
3. Hybrid model							
Post-passage dummy	−3.98	11.40	6.34	6.39	10.60*	4.53	11.80***
	7.08	10.20	4.43	5.69	6.18*	3.92	2.95***
Trend effect	0.04	−0.38	0.63	−0.23	0.70	−0.04	0.28
	0.89	0.86	0.76	0.81	0.77	0.49	0.65

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

with caution in interpreting those results (even if we had more confidence in the Lott-Mustard model than we do given the concern over omitted variables).[17]

6. Debate over the Inclusion of Linear Trends

An important issue that the NRC did not address was whether there was any need to control for state-specific linear trends. Inclusion of state trends could be important if, for example, a clear pattern in crime rates existed before a state adopted an RTC law that continued into the post-passage period. In contrast, there is also a potential danger in using state-specific trends if their inclusion inappropriately extrapolates a temporary swing in crime long into the future. Lott and Mustard (1997) never controlled for state-specific trends in analyzing handgun laws, while Moody and Marvell (2008) always controlled for these trends. Ayres and Donohue (2003b) presented evidence with and without such trends.

Table 5 replicates the NRC's full model (with the appropriate clustering adjustment) from Table 4 while adding linear state trends to this county-data model. Strikingly, Table 5 suggests that RTC laws increase aggravated assault by roughly 3% each year, but no other statistically significant effect is observed. Thus, the addition of state trends eliminates the potentially problematic result of RTC laws increasing property crimes, which actually increases our confidence in these results. Certainly, an increase in gun carrying and prevalence induced by an RTC law could well be thought to spur more aggravated assaults. Nonetheless, one must at least consider whether the solitary finding of statistical significance is merely the product of running seven different models, is a spurious effect flowing from a bad model, or reflects some other anomaly (such as changes in the police treatment of

---

17. Lott and Mustard offered a crime substitution theory based on a view that if RTC laws reduced robbery (because criminals feared encountering armed victims), the criminals might turn to property crimes that were less likely to result in armed resistance. Note, though, that Table 4 gives no support for a robbery reduction effect, so the premise of the crime substitution story is not supported.

Table 5. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—with Lott-Mustard Controls, with Clustered Standard Errors and State Trends, All Crimes, 1977–2000 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable	−6.17	−10.80	3.00	−5.31	0.21	−5.19	−0.40
model	5.31	8.27	3.60	5.66	5.85	3.55	3.04
2. Spline model	−1.21	−2.64	3.02**	−0.06	0.82	0.00	1.18
	1.46	3.48	1.23**	2.26	1.27	1.29	1.12
3. Hybrid model							
Post-passage	−5.14	−8.28	−0.64	−5.69	−0.83	−5.63	−1.95
dummy	5.07	5.65	3.79	6.28	5.99	3.95	3.25
Trend effect	−0.87	−2.09	3.06**	0.32	0.88	0.38	1.31
	1.43	3.28	1.29**	2.42	1.30	1.40	1.19

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

domestic violence cases, which could confound the aggravated assault results).[18]

7. Extending the Data through 2006

Thus far, we have presented panel data regression results for the period 1977–2000. Since more data are now available, we can further test the strength of the MGLC premise over time by estimating the NRC Lott-Mustard covariates specification on data extended through 2006. Table 6a presents our estimates (with clustering), which can be compared with Table 4 (which also clusters the standard errors in the main NRC model, but is estimated on the shorter time period).

---

18.   We tested this theory by creating a new right-hand side dummy variable that identified if a state passed legislation requiring law enforcement officials to submit official reports of all investigated domestic violence cases. Eight states have passed this legislation of which we are aware: Florida (1984), Illinois (1986), Louisiana (1985), New Jersey (1991), North Dakota (1989), Oklahoma (1986), Tennessee (1995), and Washington (1979). We included this dummy variable when running both the NRC specification (through 2000) and our preferred specification (through 2006), and found that this dummy indicator of domestic violence reporting statutes did not undermine the finding that RTC laws increase aggravated assaults.

Table 6a. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—with Lott-Mustard Controls, with Clustered Standard Errors, All Crimes, 1977–2006 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−5.44 5.91	10.40 13.20	11.40** 4.84**	3.10 4.47	14.40** 6.65**	7.48* 3.85*	12.90*** 3.96***
2. Spline model	−0.28 0.60	0.61 1.03	1.05 0.69	0.39 0.54	0.99 0.61	0.44 0.43	1.07** 0.51**
3. Hybrid model							
Post-passage dummy	−5.35 6.05	9.77 12.00	8.39** 3.48**	1.69 5.43	12.60** 5.91**	6.99* 3.99*	10.10*** 3.68***
Trend effect	−0.02 0.61	0.14 0.74	0.65 0.63	0.30 0.65	0.39 0.47	0.10 0.44	0.59 0.49

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 6b. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—with Lott-Mustard Controls, with Clustered Standard Errors and State Trends, All Crimes, 1977–2006 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−4.45 4.44	−13.00 8.14	3.44 3.13	−0.22 5.48	3.81 4.84	−0.77 3.53	1.51 3.10
2. Spline model	−0.96 0.96	−4.51 3.74	1.72* 0.94*	−0.95 1.60	−0.91 1.10	−0.82 1.04	−0.66 0.87
3. Hybrid model							
Post-passage dummy	−3.98 4.55	−10.70 7.01	2.53 3.09	0.31 5.55	4.36 4.67	−0.32 3.64	1.89 3.08
Trend effect	−0.86 0.98	−4.26 3.69	1.66* 0.93*	−0.96 1.62	−1.01 1.08	−0.82 1.07	−0.70 0.89

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

This comparison reveals that the additional six years of data somewhat strengthen the evidence that RTC laws *increase* aggravated assault, auto theft, burglary, and larceny. Table 6b simply adds state trends to the Table

6a models, which can then be compared to Table 5 (clustering, state trends, and 1977–2000 data). Collectively, these results suggest that the added six years of data do not appreciably change the results from the shorter period. The inclusion of state trends on the longer data set renders all estimates insignificant except for the evidence of marginally significant *increases* in aggravated assault.

8. Revising the Lott-Mustard Specification

We have already suggested that the Lott-Mustard specification that the NRC employed is not particularly appealing along a number of dimensions. The most obvious problem—omitted variable bias—has already been alluded to: the Lott and Mustard (1997) model had no control for incarceration, which Wilson considered to be one of the most important influences on crime in the last twenty years. In addition to a number of important omitted variables, the Lott-Mustard model adopted by the NRC includes a number of questionable variables, such as the highly dubious ratio of arrests to murders, and the thirty-six (highly collinear) demographic controls.[19]

To explore whether these specification problems are influencing the regression estimates, we revise the NRC models in a number of ways. First, we drop the flawed contemporaneous arrest rate variable and add in two preferable measures of state law enforcement/deterrence: the incarceration rate and the rate of police.[20] Second, we add two additional controls to capture economic conditions: the unemployment rate and the poverty rate, which are also state-level variables. Finally, mindful of Horowitz's admonition that the Lott-Mustard model might have *too many* variables (including demographic controls that are arguably irrelevant to the relationship between the guns and crime, and may have a spurious, misleading effect), we decided not to follow the NRC in using the thirty-six demographic controls employed by Lott-Mustard. Instead, we adhered to the more customary practice in the econometrics of crime and controlled only for the demographic groups considered to be most

---

19. For extended discussion on the abundant problems with this pseudo arrest rate, see Donohue and Wolfers (forthcoming).

20. We also estimated the model with the arrest rate (lagged by one year to avoid endogeneity concerns), and the results were qualitatively similar.

involved with criminality (as offenders and victims), namely the percentage of black and white males between ages ten and thirty years in each county.[21]

The results with this new specification are presented in Tables 7a and b (which correspond to Tables 6a and b estimated using the Lott-Mustard specification). In particular, one sees a strong adverse shift for murder. Note that had the NRC panel used our preferred specification while maintaining its view that neither clustering nor controls for state trends are needed, then we would have overwhelming evidence that RTC laws increase crime across every crime category. We do not show these regression results since we are convinced that clustering is needed, although of course when we cluster in Table 7a, the point estimates remain the same (while significance is drastically reduced).

It would indeed be a troubling state of the world if the NRC view on clustering (and linear trends) were correct, for in that event, RTC laws would increase every crime category other than murder by 20–40% (the dummy model) or increase it by 2–4% every year (the spline model)—all at the 0.01 level.[22] In fact, the version of Table 7a in which the standard errors are not adjusted by clustering generates a finding that RTC laws increase murder at the 0.10 level in the spline model and at the 0.05 level in the trend term of the hybrid model. When we do cluster, however, as shown in Table 7a, we are left with large positive point estimates but far fewer significant results: Nonetheless, this more reasonable specification suggests that RTC laws increase aggravated assault, robbery, and larceny. Interestingly, adding state trends in Table 7b wipes out all statistical significance.

This discussion again highlights how critical the choices of clustering and state trends are to an assessment of RTC laws. Using neither, the data suggest these laws are harmful. With only clustering, RTC laws show (marginally significant) signs of increases for two violent crime categories as well as for larceny. In our preferred specification (without state trends), the effect of RTC laws on murder seems to basically be zero. With both clustering

---

21. To test the robustness of this specification to alternations in the demographic controls used, we also estimated the following models: Only black men between ages ten and forty years; black, white, and Hispanic men between ages ten and forty years; only black men between ages ten and thirty years; black and white men between ages ten and thirty years; and black, white, and Hispanic men between ages ten and forty years. The results were again qualitatively similar across our tests.

22. These results are not presented here since standard errors clustered on state are clearly needed. The authors can provide these results upon request.

Table 7a. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—with Preferred Controls, with Clustered Standard Errors, All Crimes, 1977–2006 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−0.44	21.30	21.60	19.30	24.80	26.60	29.50
	7.13	19.40	19.00	14.50	21.10	22.40	26.00
2. Spline model	0.31	2.34	3.16	2.64*	3.12	3.59	4.20
	0.79	1.83	1.89	1.46*	2.11	2.27	2.61
3. Hybrid model							
Post-passage dummy	−2.72	12.60	7.40	7.92	12.00	11.10	10.90
	6.96	15.40	15.80	12.10	16.80	18.20	20.50
Trend effect	0.45	1.70	2.78*	2.24*	2.51	3.03	3.64*
	0.81	1.39	1.62*	1.27*	1.74	1.94	2.15*

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this "preferred" specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 7b. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—with Preferred Controls, with Clustered Standard Errors and State Trends, All Crimes, 1977–2006 (without 1993 Data)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−3.11	−15.50	0.02	1.15	1.89	−3.98	−3.22
	4.81	10.80	9.70	7.25	9.89	10.90	12.50
2. Spline model	−0.41	−6.69	0.61	−0.82	−0.97	−1.92	−2.25
	1.31	4.77	2.44	2.28	2.66	2.83	3.15
3. Hybrid model							
Post-passage dummy	−2.97	−13.00	−0.22	1.48	2.29	−3.25	−2.35
	5.08	9.98	10.30	7.64	10.40	11.50	13.10
Trend effect	−0.35	−6.46	0.61	−0.85	−1.01	−1.87	−2.21
	1.35	4.76	2.54	2.35	2.76	2.96	3.29

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this "preferred" specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

and state trends, all statistically significant effects are wiped out. The only conclusion from both the NRC/Lott-Mustard model and our preferred specification (on county data) is that there is no robust evidence that RTC laws

provide any net benefits, and there is a greater likelihood that RTC laws may cause either some or a great deal of harm.

9. State versus County Crime Data

In their initial study, Lott and Mustard (1997) tested the MGLC hypothesis by relying primarily on county-level data from the FBI's UCR.[23] These FBI reports present yearly estimates of crime based on monthly crime data from local and state law enforcement agencies across the country. The NRC report followed Lott and Mustard in this choice and presented regression estimates using only county data. Unfortunately, according to criminal justice researcher Michael Maltz, the FBI's county-level data are highly problematic.

The major problem with county data stems from the fact that law enforcement agencies voluntarily submit crime data to the FBI. As a result, the FBI has little control over the accuracy, consistency, timeliness, and completeness of the data it uses to compile the UCR reports. In a study published in the *Journal of Quantitative Criminology*, Maltz and Targonski (2002) carefully analyzed the shortcomings in the UCR data set and concluded that UCR county-level data are unacceptable for evaluating the impact of RTC laws. For example, in Connecticut, Indiana, and Mississippi, over 50% of the county-level data points are missing crime data for more than 30% of their populations (Maltz and Targonski, 2002). In another thirteen states, more than 20% of the data points have gaps of similar magnitude. Based on their analysis, Maltz and Targonski (2002) concluded that:

County-level crime data cannot be used with any degree of confidence . . .. The crime rates of a great many counties have been underestimated, due to the exclusion of large fractions of their populations from contributing to the crime counts. Moreover, counties in those states with the most coverage gaps have laws permitting the carrying of concealed weapons. How these shortcomings can be compensated for is still an open question . . . it is clear, however, that in their current condition, county-level UCR crime statistics cannot be used for evaluating the effects of changes in policy. (p. 316–17)

Because of the concerns raised about county-level crime data, it is prudent to test our models on state-level data. According to Maltz and Targonski (2003), state-level crime data are less problematic than county-level data because the

---

23.   Lott and Mustard present results based on state-level data, but they strongly endorse their county-level over their state-level analysis: "the very different results between state- and county-level data should make us very cautious in aggregating crime data and would imply that the data should remain as disaggregated as possible" (Lott and Mustard, 1997, p. 39).

FBI's state-level crime files take into account missing data by imputing all missing agency data. County-level files provided by National Archive of Criminal Justice Data, however, impute missing data only if an agency provides at least six months of data; otherwise, the agency is dropped completely (Maltz, 2006). As with our estimations using county-level data, we compiled our state-level data from scratch, and will refer to it as "Updated 2009 State-Level Data."

Unsurprisingly, the regression results reproduced using state-level data are again different from the NRC committee's estimates using county-level data. This is shown in Table 8a, which presents the results from the NRC's specification (the Lott-Mustard model) on state data, with the cluster adjustment.[24] Table 8b simply adds state trends. When we compare these state-level estimates to the county-level estimates (using the updated 2009 county-level data set), we see that there are marked differences. Considering the preceding discussion on the reliability—or lack thereof—of county data, this result is unsurprising. Importantly, state-level data through 2006 show not a hint of statistically significant evidence that RTC laws reduce murder.[25] None of the state results is robust to the addition or exclusion of state linear trends.

Tables 9a and b below repeat Tables 8a and b, but use the model with our preferred set of explanatory variables instead of the Lott and Mustard (1997) model. The main question raised by these estimations is whether state trends are needed in the regression models. If not, there is evidence that RTC laws increase assault and larceny. If state trends are needed, some muddiness returns but RTC laws appear to increase aggravated assault, while declines in rape are marginally significant.

10. Additional Concerns in the Evaluation of Legislation Using Observational Data

We now turn to three critical issues that must be considered when using panel data to evaluate the impact of legislation and public policy (and gun

---

24. Our placebo test on county data showed that standard errors needed to be adjusted by clustering. In Appendix A, we again find that clustering is needed for state data. Thus, all our state-level estimates include clustering.

25. We also estimate the model on data through 2000 (the last year in the NRC report), though those results are not shown here. The results similarly do show not any statistically significant evidence that RTC laws reduce murder. Moreover, we also estimate the NRC's no-controls model on the state-level data. See Appendix B for these results.

Table 8a. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Lott-Mustard Controls, with Clustered Standard Errors, All Crimes, 1977—2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable	−4.94	−5.04**	1.44	−6.96**	0.31	−4.97**	2.32
model	3.61	2.29**	4.11	2.90**	3.98	2.22**	1.58
2. Spline model	−0.03	−0.49	0.80	−0.16	−0.87**	−0.44	0.40
	0.54	0.33	0.66	0.60	0.42**	0.45	0.29
3. Hybrid model							
Post-passage	−5.62	−3.77	−1.69	−7.41**	4.00	−3.92*	1.03
dummy	4.25	2.36	3.26	3.59**	4.88	2.03*	1.80
Trend effect	0.19	−0.35	0.86	0.12	−1.02**	−0.29	0.36
	0.58	0.36	0.64	0.64	0.50**	0.46	0.32

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 8b. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Lott-Mustard Controls, with Clustered Standard Errors and State Trends, All Crimes, 1977–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable	−3.32	−3.33	−1.12	−3.36	2.64	−1.93	1.21
model	3.47	2.20	2.78	3.04	2.71	1.37	1.07
2. Spline model	0.42	0.34	2.49***	0.46	−1.95***	0.35	0.39
	0.82	0.88	0.61***	1.00	0.72***	0.79	0.60
3. Hybrid model							
Post-passage	−3.83	−3.78	−3.33	−3.90	4.51	−2.33	0.92
dummy	3.58	2.42	2.84	3.10	2.85	1.62	1.28
Trend effect	0.61	0.54	2.67***	0.66	−2.19***	0.47	0.35
	0.81	0.92	0.63***	1.00	0.77***	0.83	0.64

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables (adopted from the Lott-Mustard model) include arrest rate, county population, population density, per capita income measures, and thirty-six demographic composition measures indicating the percentage of the population belonging to a race-age-gender group.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 9a. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with Clustered Standard Errors, All Crimes, 1977–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−2.93	−0.62	5.05	5.36	7.03	2.24	6.72**
	3.94	3.76	3.71	4.28	6.05	3.00	2.98**
2. Spline model	−0.16	−0.44	1.09*	0.64	0.45	0.00	0.57
	0.61	0.54	0.60*	0.75	0.62	0.39	0.46
3. Hybrid model							
Post-passage dummy	−2.75	1.71	0.15	3.09	6.29	2.82	5.22*
	3.75	3.52	3.56	4.74	5.49	3.21	3.05*
Trend effect	−0.04	−0.52	1.09*	0.50	0.17	−0.13	0.34
	0.63	0.56	0.63*	0.83	0.56	0.43	0.50

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this ''preferred'' specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 9b. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with Clustered Standard Errors and State Trends, All Crimes, 1977–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	0.54	−3.61*	−2.03	2.40	8.17*	1.51	1.89
	2.72	1.83*	3.05	3.67	4.16*	2.18	1.83
2. Spline model	0.83	0.08	3.10**	0.51	−1.84**	−0.22	−0.15
	0.87	0.79	0.81**	1.29	0.82**	0.88	0.74
3. Hybrid model							
Post-passage dummy	0.11	−3.70*	−3.68	2.17	9.26**	1.65	1.99
	2.86	1.96*	3.15	3.96	4.24**	2.41	1.97
Trend effect	0.83	0.19	3.21***	0.44	−2.11**	−0.27	−0.20
	0.89	0.79	0.82***	1.35	0.84**	0.91	0.77

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this ''preferred'' specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

laws in particular). First, we discuss the possibility of difficult-to-measure omitted variables, and how such variables can shape estimates of policy impact. We are particularly concerned with how the crack epidemic of the 1980s and 1990s may bias results in the direction of finding a beneficial effect. Second, we explore pre-adoption crime trends in an attempt to examine the plausibly endogenous adoption of RTC legislation. Finally, given that the intent of right-to-carry legislation is to increase gun-carrying in law-adopting states, we explore whether these laws may have had a particular effect on gun-related assaults (which is the one crime category that has generated somewhat consistent results thus far).

10.1. Further Thoughts on Omitted Variable Bias

As discussed above, we believe it is likely that the NRC's estimates of the effects of RTC legislation are marred by omitted variable bias. In our attempt to improve (at least to a degree) on the original Lott-Mustard model, we included additional explanatory factors, such as the incarceration and police rates, and removed extraneous variables (such as unnecessary and collinear demographic measures). We recognize, however, that there are additional criminogenic influences for which we cannot fully control. In particular, we suspect that a major shortcoming of all the models presented is the inability to account for the possible influence of the crack cocaine epidemic on crime.[26]

---

26. Although Lott and Mustard (1997) do make a modest attempt to control for the potential influence of crack cocaine through the use of cocaine price data based on the U.S. Drug Enforcement Administration's STRIDE data, we find their approach wanting for both theoretical and empirical reasons. First, a control for crack should capture the criminogenic influence of the crack trade on crime. We know that prior to 1985, there was no such influence in any state and that after some point in the early to mid-1990s this criminogenic influence declined strongly. Since there is little reason to believe that cocaine prices would be informative on the criminogenic influence of crack in particular geographic areas, it is hard to see how the cocaine price data could be a useful control. Second, the data that Lott and Mustard use are themselves questionable. Horowitz (2001) argues forcefully that STRIDE data are not a reliable source of data for policy analyses of cocaine. The data are mainly records of acquisitions made to support criminal investigations in particular cities, and are not a random sample of an identifiable population. Moreover, since the STRIDE data are at the city level, we are not sure how this would be used in a county-level analysis. The data were collected for twenty-one cities, while there are over three thousand counties in the United States. In addition, the data are missing for 1988 and 1989, which are crucial years in the rise of the crack epidemic in poor urban areas. Lott and Mustard drop those years of analysis when including cocaine prices as a control.

Many scholars now suggest that rapid growth in the market for crack co-caine in the late 1980s and the early 1990s was likely one of the major influ-ences on increasing crime rates (and violent crimes in particular) during this period (Levitt, 2004). Moreover, the harmful criminogenic effect of crack was likely more acute in urban areas of states slow to adopt RTC laws. Meanwhile, many rural states adopted such laws during this era. If this was indeed the case, this divergence between states could account for much of the purported "crime-reducing" effects attributed by Lott and Mustard to gun laws (which were then supported by scholars such as James Q. Wilson). The regression analysis would then identify a relationship between rising crime and the failure to adopt RTC legislation, when the actual reason for this trend was the influence of crack (rather than the passage of the RTC law).

We now explore how results from our main models vary when we restrict the analysis to the time periods before and after the peak of the American crack epidemic. According to Fryer et al. (2005), the crack problem through-out most of the country peaked at some point in the early 1990s. Coinciden-tally, the original Lott-Mustard period of analysis (1977–1992) contains years that likely represent the height of crack-induced crime problem. With this in mind, we run our main regressions after breaking up our data set into two periods: the original Lott-Mustard period of analysis (1977–1992) and the post–Lott-Mustard period (1993–2006). We first present the results for the era that includes the crack epidemic (1977–1992) on our preferred model. We run these regressions (with clustered standard errors) on state-level data, with and without state trends. These results are presented in Tables 10a and b. We then estimate the same models on the post-crack period (see Tables 11a and b).

Note that the regression results in Table 10 from the initial Lott-Mustard sixteen-year time period (1977–1992) do suggest that rape, robbery, and aggra-vated assault are dampened by RTC laws if state trends are not needed and that murder may have declined if state trends are needed. If we look at the following fourteen-year period from 1993 to 2006 in Table 11, however, the conclusion flips around: Now, there is evidence that all four violent crimes *rose* when states adopted RTC laws. This evidence supports the theory that the Lott-Mustard find-ing was likely the result of the crime-raising impact of crack in non-RTC states.

Figure 8 depicts a measure of crack prevalence for the period 1980–2000 in the five states with the greatest crack problem as well as the five states with the

Table 10a. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with Clustered Standard Errors, All Crimes, 1977–1992[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−3.69	−12.10***	−6.55	−4.85	7.28	−3.73	0.12
	3.81	3.41***	4.66	4.07	4.73	2.45	1.52
2. Spline model	−0.88	−2.87***	0.52	−2.28***	0.51	−0.34	−0.10
	1.44	0.80***	1.70	0.72***	1.13	0.83	0.33
3. Hybrid model							
Post-passage dummy	−2.32	−7.59**	−11.80**	1.08	9.07*	−4.37	0.54
	4.70	3.01**	5.64**	5.32	4.61*	3.87	1.82
Trend effect	−0.56	−1.83***	2.13	−2.42**	−0.73	0.26	−0.17
	1.67	0.59***	1.47	1.08**	0.85	0.97	0.42

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this ''preferred'' specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
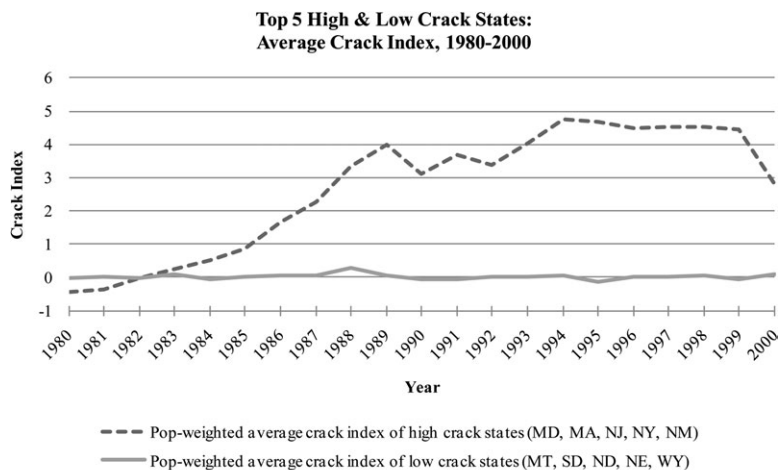*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 10b. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with State Trends and Clustered Standard Errors, All Crimes, 1977–1992[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−5.61	−4.14	−2.02	−3.78	−0.04	−3.05	1.28
	3.57	3.61	3.70	4.25	3.84	2.23	1.96
2. Spline model	−5.41**	0.27	−0.05	−4.35*	−1.62	−2.36	0.37
	2.45**	1.11	1.17	2.48*	2.20	1.43	1.15
3. Hybrid model							
Post-passage dummy	2.47	−6.67*	−2.89	3.08	3.17	0.18	1.16
	4.31	3.52*	5.10	6.91	4.98	4.26	2.02
Trend effect	−6.01**	1.88	0.65	−5.10	−2.38	−2.41	0.09
	2.51**	1.18	1.84	3.30	2.64	2.11	1.26

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this ''preferred'' specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 11a. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with Clustered Standard Errors, All Crimes, 1993–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	3.12	−3.47	1.36	3.64	2.46	3.58	0.27
	3.61	2.47	3.54	4.89	4.50	2.57	2.74
2. Spline model	1.11*	−0.21	1.91**	1.78**	−0.30	0.35	0.08
	0.63*	0.68	0.74**	0.87**	0.80	0.71	0.55
3. Hybrid model							
Post-passage dummy	2.36	−3.35	0.03	2.42	2.70	3.37	0.22
	3.82	2.46	4.05	4.73	4.33	2.57	2.76
Trend effect	1.09*	−0.17	1.91**	1.75**	−0.34	0.31	0.08
	0.64*	0.67	0.76**	0.87**	0.77	0.70	0.55

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this ''preferred'' specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table 11b. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with State Trends and Clustered Standard Errors, All Crimes, 1993–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	3.12	0.27	2.38	3.81	2.83	0.89	0.33
	3.62	2.66	2.59	3.33	3.39	2.19	1.83
2. Spline model	−1.99	2.61**	4.34***	−0.17	−5.53*	−0.71	−1.49
	2.00	1.16**	1.53***	1.89	2.77*	1.74	1.31
3. Hybrid model							
Post-passage dummy	4.04	−0.75	0.79	4.04	5.12	1.20	0.93
	3.87	2.46	2.40	3.48	3.43	2.29	1.98
Trend effect	−2.44	2.69**	4.25**	−0.62	−6.10**	−0.84	−1.59
	2.10	1.14**	1.61**	1.95	2.99**	1.80	1.42

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this ''preferred'' specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Top 5 High & Low Crack States:
**Average Crack Index, 1980-2000**

Source: Authors' calculations based on the crack index of Fryer et al. (2005).

Figure 8. Prevalence of Crack in the Five Most and the Five Least Crack-Affected States.

least crack, according to Fryer et al. (2005). Figure 9 shows the murder rates over time for these two sets of states. We see that crime rose in the high-crack states when the crack index rises in the mid-to-late 1980s, but that the crack index does not turn down in those states at the time crime started to fall.

Top 5 High & Low Crack States:
Average Murder Rate, 1980-2003
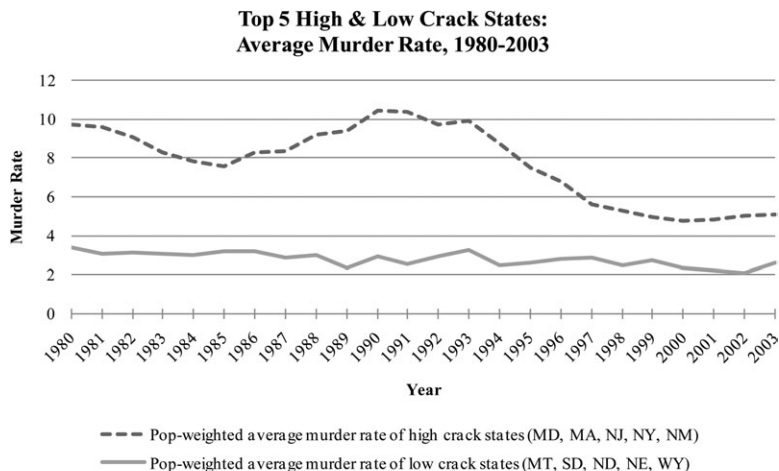
Source: Bureau of Justice Statistics (2009).

Figure 9. Murder Rates in the Five Most and the Five Least Crack-Affected States.

Apparently, the rise of the crack market triggered a great deal of violence but once the market stabilized, the same level of crack consumption could be maintained while the violence ebbed.

Of course, omitting an appropriate control for the criminogenic influence of crack is problematic if the high-crack states tend not to adopt RTC laws and the low-crack states tend to adopt. This is in fact the case: All the five "high-crack" states are non-RTC states during this period, whereas four of the five "low-crack" states are RTC states (all four adopted an RTC law by 1994).[27] The only exception is Nebraska, a state that did not adopt an RTC law until 2007, which is outside the scope of our current analyses.[28]

Table 12. Population-Weighted Statistics of RTC-Adopting States between 1977 and 1990[a]

State	Year of RTC Law Adoption	Murder Rate	Crack Index
Indiana	1980	6.53	0.17
Maine	1985	2.53	−0.04
North Dakota	1985	1.29	0.01
South Dakota	1986	2.10	−0.03
Florida	1987	11.73	0.67
Virginia	1988	7.90	0.65
Georgia	1989	12.28	0.92
Pennsylvania	1989	5.73	0.65
West Virginia	1989	5.65	0.32
Idaho	1990	3.56	0.30
Mississippi	1990	11.65	0.25
Oregon	1990	4.85	0.76

Notes: Source—Fryer et al. (2005) and Bureau of Justice Statistics (2009).
[a]The crack index data come from the Fryer et al. (2005) study, which constructs the index based on several indirect proxies for crack use, including cocaine arrests, cocaine-related emergency room visits, cocaine-induced drug deaths, crack mentions in newspapers, and Drug Enforcement Administration drug busts. The article does suggest that these values can be negative. The state with the lowest mean value of the crack index over our data period is Maine (−0.04) and the state with the highest mean value is New York (1.15). (The article does suggest that the crack index values can be negative.)

---

27. New Mexico, one of the five highest crack states, adopted its RTC law in 2003. Wyoming and Montana adopted RTC laws in 1994 and 1991, respectively. North Dakota and South Dakota adopted their laws prior to the start of our data set (pre-1977), although the dates are contested (Lott and Mustard, 1997; Moody and Marvell, 2008).

28. In fact, out of the ten states with the lowest crack cocaine index, seven adopted an RTC law by 1994. The exceptions are Nebraska, Minnesota (2003), and Iowa (no RTC law).

Moreover, as Table 12 reveals, the twelve states that adopted RTC laws during the initial Lott-Mustard period (1977–1992) had crack levels substantially below the level of the five high-crack states shown in Figures 8 and 9. None of the RTC adopters shown in Table 12 has an average crack index value that even reaches 1, while Figure 9 reveals that the high-crack states had a crack level in the neighborhood of 4 or 5.

In other words, over the initial Lott-Mustard period of analysis (ending in 1992), the criminogenic influence of crack made RTC laws look beneficial since crack was raising crime in non-RTC states. In the later period, crime fell sharply in the high-crack states, making RTC states look bad in comparison. Therefore, the effects estimated over this entire period will necessarily water down the initial Lott-Mustard results. The hope is that estimating the effect over the entire period will wash out the impact of the omitted variable bias generated by the lack of an adequate control for the effect of crack.

10.2. Endogeneity and Misspecification Concerns

To this point, our analysis has remained within the estimation framework common to the NRC/Lott-Mustard analyses, which implicitly assumes that passage of RTC legislation in a given state is an exogenous factor influencing crime levels. Under this assumption, one can interpret the estimated coefficient as an unbiased measure of RTC laws' collective impact.

We probe the validity of this strong claim by estimating a more flexible year-by-year specification, adding pre- and post-passage dummy variables to the analysis.[29] Pre-passage dummies can allow us to assess whether crime trends shift in unexpected ways prior to the passage of a state's RTC law. Autor, Donohue, and Schwab (2006) point out that when analyzing the impact of state-level policies using panel data, one would ideally see lead dummies that are near zero. The graphs that we present below, though, suggest the possible presence of systematic differences between RTC law adopters that can complicate or thwart the endeavor of obtaining clean estimates of the impact of RTC laws.

Figures 10–13 present the results from this exercise in graphical form. Using our preferred model as the base specification, we introduce dummies for the eight years preceding and the first eight years following adoption. We

---

29.   In Appendix C, we further analyze the issue of misspecification and model fit by analyzing residuals from the regression analysis.
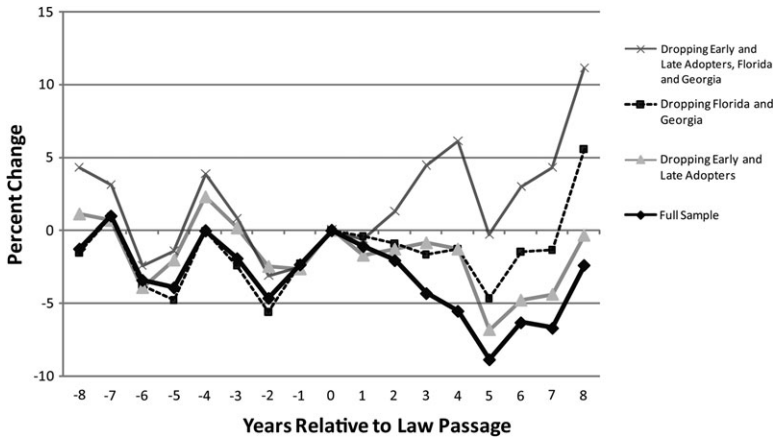
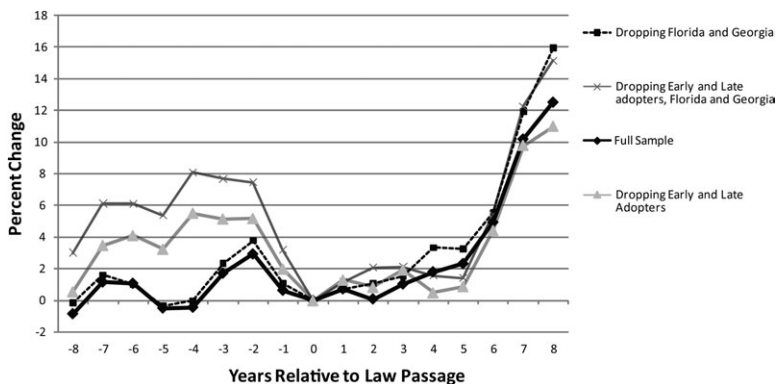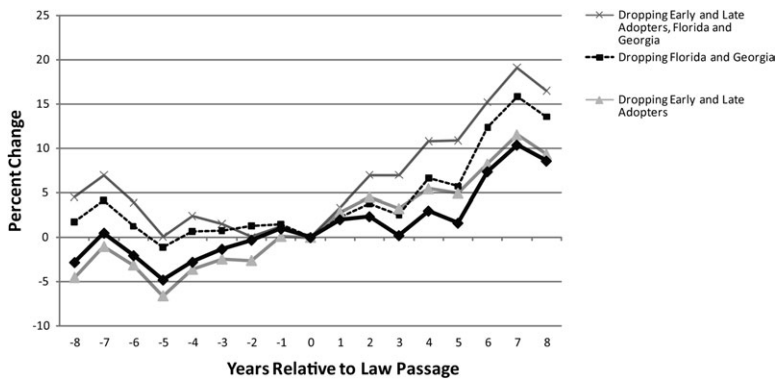**Figure 10.** Normalized Year-by-Year Estimates of the Impact of RTC Laws on Murder.
Notes: Estimations include year and county fixed effects, state trends, and are weighted by county population. The control variables include incarceration and police rates, unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.

first estimate this regression for each violent crime category over the full sample of RTC states. However, because of the presence of one state that adopted its RTC law just three years after our data set begins, and eight states that adopted laws within the five years before our data set ends, we have nine
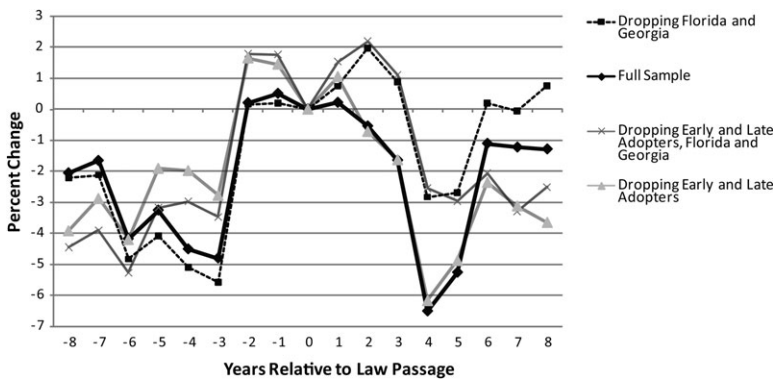
Figure 11. Normalized Year-by-Year Estimates of the Impact of RTC Laws on Rape.
Notes: Estimations include year and county fixed effects, state trends, and are weighted by county population. The control variables include incarceration and police rates, unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.

Figure 12. Normalized Year-by-Year Estimates of the Impact of RTC Laws on Assault.
Notes: Estimations include year and county fixed effects, state trends, and are weighted by county population. The control variables include incarceration and police rates, unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.

states that cannot enter into the full set of pre- and post-adoption dummy variables. Because Ayres and Donohue (2003a) showed that the year-by-year estimates can jump wildly when states drop in or out of the individual year estimates, we also estimate the year-by-year model after dropping out

Figure 13. Normalized Year-by-Year Estimates of the Percent Change in Robbery.
Notes: Estimations include year and county fixed effects, state trends, and are weighted by county population. The control variables include incarceration and police rates, unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.

the earliest (1980) and latest (post-2000) law-adopting states. In this separate series of regressions, our estimates of the full set of lead and lag variables are based on the set of all twenty-five adopters between 1985 and 1996.[30]

Unfortunately, the graphs raise concerns about the presence of endogenous adoption that complicate our thinking about the influence of RTC laws on violent crime. If one looks at the four lines in Figure 10, one sees four different sets of year-by-year estimates of the impact of RTC laws on murder. The lines have been normalized to show a zero value in the year of adoption of an RTC law. Let us begin with the bottom line (looking at the right-hand side of the figure) and the line just above it. The lower line represents the naive year-by-year estimates from the preferred model estimated on the 1977–2006 period, while the line just above it drops out the early and late adopters, so that the estimated year-by-year estimates are based on the "clean" sample of twenty-five adopters for which complete data are available from eight years prior to adoption through eight years after adoption. One immediately sees that the trimmed estimates are different and less favorable to the MGLC hypothesis, as evidenced by the higher values in the post-passage period. They also look superior in the pre-passage period in that on average the pre-passage dummies are closer to zero for the trimmed set of estimates (the mean of the pre-passage dummies is x for the trimmed estimate and Y for the naive estimate).[31]

How should we interpret these trimmed sample estimates? One possibility is to conclude that on average the pre-passage estimates are reasonably close to zero and then take the post-passage figures as reasonable estimates of the true effect. If we do this, none of the estimates would be statistically significant, so one could not reject the null hypothesis of no effect. But note that the pre-passage year-to-year dummies show an oscillating pattern that is not altogether different from what we see for the post-passage values. Without

30.   The states that drop out (with dates of RTC law passage in parentheses) include Indiana (1980), Michigan (2001), Colorado (2003), Minnesota (2003), Missouri (2003), New Mexico (2003), Ohio (2004), Kansas (2006), and Nebraska (2006).

31.   Note that this bias in favor of a deterrent effect for murder would also be operating in the aggregate estimates, further suggesting that the true aggregate estimates would be commensurately less favorable for the deterrence hypothesis than the ones we presented earlier in this article—and in all other articles providing unadjusted aggregate estimates.

the odd drop when moving from Year 4 to Year 5 and subsequent rise in values through Year 8, the zero effect story would seem more compelling, but perhaps the drop merely reflects a continuation of the pre-passage oscillations, which are clearly not the product of the passage of RTC laws.[32]

Perhaps what is most important is not the oscillations but rather the trend just prior to passage. This might suggest that rising crime in fact increases the likelihood that a state would adopt an RTC law. In particular, since murder is typically the crime most salient in the media, we suspect it has the greatest effect on implementation of purported crime control measures such as RTC legislation. Of course, this would suggest an endogeneity problem that would also likely lead to a bias in favor of finding a deterrent effect. The mechanism driving this bias would presumably be that rising crime strengthens the National Rifle Association's push for the law, and the mean reversion in crime would then falsely be attributed to the law by the naive panel data analysis (incorrectly premised on exogenous RTC law adoption). Post-adoption murder rates again decline—often to within the neighborhood of pre-law levels. We do, however, uncover some interesting findings when estimating (more cleanly) the year-by-year effects on the twenty-five states for which we have observations across the full set of dummy variables.

Another striking feature we note is the strong influence of Florida and Georgia on our estimates of the impact of RTC laws on murder and rape. When we remove these two states, the post-adoption trend lines for murder and rape shift upward substantially. Moreover, when dropping them from the set of RTC states that already excludes the early and late adopters—still leaving us with twenty-three RTC states to analyze—we see that murder increases in each post-adoption year except one. As previous articles have noted, Florida experienced enormous drops in murder during the 1990s that may have been completely unrelated to the passage of its RTC policy. Donohue (2003) points out that the 1980 Mariel boatlift temporarily added many individuals prone to committing crimes to Florida's population, causing a massive increase in crime in Florida during the 1980s. Thus, it is plausible that the massive 1990s crime reductions in Florida were not driven by the

---

32.   The ostensible pronounced drop in murder five years after adoption (exists for the full data set, as well, but it is part of a continuing downward trend in murder that simply reaches a trough five years after passage).

adoption of the state's RTC law but rather a return to traditional population dynamics that were less prone to violent crime (again, a reversion to the mean). This is important to consider given the strong downward pull of Florida on aggregate murder rates.

The line based on dropping Florida and Georgia from the trimmed sample would suggest that for the twenty-three other states, the impact of RTC laws on murder was highly pernicious—and increasingly so as the sharp upward trend in the last three years would suggest. Again a number of interpretations are possible: (1) Florida and Georgia are unusual and the best estimate of the impact of RTC laws comes from the trimmed sample that excludes them (and the early and late adopters); (2) there is heterogeneity in the impact of RTC laws, so we should conclude that the laws help in Florida and Georgia, and tend to be harmful in the other twenty-three states; and (3) omitted variables mar the state-by-state estimates but the aggregate estimates that include Florida and Georgia may be reasonable if the state-by-state biases on average cancel out.

Note that Figure 11, which presents the comparable year-by-year estimates of the impact of RTC laws on rape, shows a similar yet even more extreme pattern of apparent spikes in crime leading to adoption of RTC laws followed by a substantial amount of mean reversion. The somewhat unsettling conclusion from Figures 10 and 11 is that RTC laws might look beneficial if one only had data for four or five years, but this conclusion might be substantially reversed if a few additional years of data were analyzed. Taken as a whole, these two figures show the sensitivity of the estimates to both the time period and sample of states that are analyzed.

Further casting doubt on the possibility that drops in murder and rape could be attributed to the passage of RTC laws, a dramatically different picture emerges from our year-by-year analysis of these laws' impact on assault and robbery rates. The general story here seems to be that assault increases markedly over the time period after law passage, which squares with our results discussed in the previous sections. One observes positive coefficient changes that are initially modest, but these increase dramatically and uniformly over the second half of the post-passage period. Moreover, in contrast to the year-by-year murder and rape estimates, assault trends are not demonstrably different when we alter the sample to exclude early and late adopters, as well as Florida and Georgia. The pattern is generally unaffected by sample, giving us some confidence that RTC laws may be having an adverse

impact on the rate of assault. Robbery rates similarly increase over time after the passage of RTC laws, although not as dramatically.

Something to consider, however, is how one should interpret the assault trends in light of the murder trends just discussed. If, for example, the decline in murder to pre-law levels after RTC laws' passage is nothing more than a "mean reversion" effect, it is conceivable that the apparent increase in assault simply represents mean reversion in reverse (from relatively low to high). It is important to note, however, that while assault does return to its pre-law levels a few years after passage, the coefficients continue to rise dramatically, with no hint of any subsequent mean regression. Thus, a more plausible way to interpret the near uniform increases in assault coefficients is that aggravated assault did actually increase over time with the passage of RTC legislation, which strongly undercuts the "MGLC" thesis. Interestingly, the robbery data (Figure 13) suggest either a pernicious effect similar to that on aggravated assault (particularly for the trimmed estimates dropping only early and late adopters) or a strong upward trend in crime, starting well before passage, that might be taken as a sign of the absence of any impact of RTC laws on robbery.

10.3 Effects of RTC Laws on Gun-related Assaults

Thus far in our analysis, we have yet to consider whether RTC laws affect aggravated assaults committed with a firearm differently than aggravated assaults overall. This is important to consider given that the 1990s witnessed huge movements in reported assaults due to cultural shifts around the issue of domestic violence. Many of these crimes would not have involved guns, making it possible that our results above suggesting increased rates of assault in RTC states are actually a statistical artifact of changing crime-reporting norms. For this reason, gun-related aggravated assaults may be an arguably more reliable statistic for measuring RTC laws' impact than overall aggravated assaults.

To test this possibility, we estimate our preferred regression using gun-related aggravated assaults as the dependent variable (both with and without state-specific trends) in Table 13 below. Comparing these new results with the assault estimates in Tables 9a and 9b above, our bottom-line story of how RTC laws increases rates of aggravated assault does not change much when limiting our analysis to assaults involving a gun. Without state trends, we see large positive estimates, some of which are significant at the 10% level. With

Table 13. Estimated Impact of RTC Laws on Gun-related Aggravated Assaults—Using Updated 2009 State-Level Data—With Preferred Controls, With Clustered Standard Errors, All Crimes, 1977–2006[a]

	Without State Trends (%)	With State Trends (%)
1. Dummy variable model:	15.50*	0.67
	8.11*	7.48
2. Spline model:	2.23*	5.64*
	1.27*	3.12*
3. Hybrid model:		
Postpassage dummy	7.76	−2.19
	7.76	7.13
Trend effect	1.90	5.71*
	1.28	3.08*

[a]Estimations include year and county fixed effects and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this ''preferred'' specification include: incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

state trends, we again see some significant evidence that gun-related aggravated assault rates are increased by RTC legislation. These results solidify our overall confidence in the array of estimates we present above that suggests that RTC laws raise rates of aggravated assault.

11. Conclusions

In this article, we have explored the NRC panel's 2005 report detailing the impact of RTC gun laws on crime. Using the committee's models as a starting point for our analysis, we highlight the importance of thoroughly considering all the possible data and modeling choices. We also highlight some issues that should be considered when evaluating the NRC report.

Data reliability is one concern in the NRC study. We corrected several coding errors in the data that were provided to us by the NRC (which had originally been obtained from John Lott). Accurate data are essential to making precise causal inferences about the effects of policy and legislation—and this issue becomes particularly important when we are considering topics as controversial as firearms and crime control. We attempted to mitigate any uncertainty over data reliability by re-collecting the data. However, when attempting to replicate the NRC specifications—on both the NRC's and

our own newly constructed data sets—we consistently obtained point estimates that differed substantially from those published by the committee.

Thus, an important lesson for both producers and consumers of econometric evaluations of law and policy is to understand how easy it is to get things wrong. In this case, it appears that Lott's data set had errors in it, which then were transmitted to the NRC committee for use in evaluating Lott and Mustard's hypothesis. The committee then published tables that could not be replicated (on its data set or a new corrected data set), but which made at least Professor James Q. Wilson think (incorrectly it turns out—see our Tables 2a–c) that running Lott-Mustard regressions on both data periods (through 1992 and through 2000) would generate consistently significant evidence that RTC laws reduce murder. This episode suggests to us the value of making publicly available data and replication files that can re-produce published econometric results. This exercise can both help to uncover errors prior to publication and then assist researchers in the process of replication, thereby aiding the process of ensuring accurate econometric estimates that later inform policy debates.

A second lesson is that the "best practices" in econometrics are evolving. Researchers and policy makers should keep an open mind about controversial policy topics in light of new and better empirical evidence or methodologies. Case in point: The NRC report suggested that clustering standard errors on the state level in order to account for serial correlation in panel data was not necessary to ascertain the impact of RTC laws on crime. However, most applied econometricians nowadays consider clustering to be advisable in the wake of a few important articles, including one in particular by Bertrand, Duflo, and Mullainathan (2004) on difference-in-differences estimation. The evidence we present corroborates the need for this standard error adjustment. Our placebo tests showed that standard errors are greatly understated without clustering, and we believe strongly that this adjustment is vital for both county-level and state-level analyses of gun laws and crime. Otherwise, statistical significance is severely exaggerated and significant results are detected where none in fact exists.

A third lesson relates to the potential flaws in the Lott-Mustard (and by extension, the NRC) approach and specification. Issues—such as the inclusion of state-specific linear trends, the danger of omitted variable bias, and the choice of county-level over state-level data, all of which the NRC neglected to discuss—clearly have enough impact on the panel data

estimates to influence one's perception of the MGLC theory and thus warrant closer examination. These issues were not all arcane (although many were, such as the need to control for state trends). By now, empirical researchers should be well acquainted with omitted variable bias, and the increases in the prison and police populations were known major factors influencing the pattern of U.S. crime in recent decades (Wilson, 2008). Yet, the Lott-Mustard model—adopted by the NRC—had no control for incarceration or police![33] On that basis alone, Wilson might well have hesitated before accepting the MGLC hypothesis on the basis of the Lott-Mustard or NRC results. Yet, Lott, with at best questionable support for his view that RTC laws reduce murder, now claims that Wilson, one of the most eminent criminologists of our time, supports his position (Lott, 2008). Clearly, the consequences of embracing fragile empirical evidence can be severe.

Granted, much of the work of applied econometricians is of the sort that was set forth by the NRC as evidence on the impact of RTC laws. The committee, though, found this evidence inadequate to reach a conclusion, doubtless because the results seemed too dependent on different modeling choices. But Horowitz is even more nihilistic, essentially rejecting all applied econometric work on RTC legislation, as indicated by his following independent statement in an appendix to the NRC's (2005) report:

It is unlikely that there can be an empirically based resolution of the question of whether Lott has reached the correct conclusions about the effects of right-to-carry laws on crime. (p. 304)

Of course, if there can be no empirically based resolution of this question, it means that short of doing an experiment in which laws are randomly assigned to states, there will be no way to assess the impact of these laws. The econometrics community needs to think deeply about what the NRC report and the Horowitz appendix imply for the study of legislation using panel data econometrics and observational data.

Finally, despite our belief that the NRC's analysis was imperfect in certain ways, we agree with the committee's cautious final judgment on the

---

33. The Lott-Mustard model omitted a control for the incarceration rate (which is indicated implicitly—though not explicitly—in the notes to each table of the NRC report, which listed the controls included in each specification).

effects of RTC laws: "with the current evidence it is not possible to de-termine that there is a causal link between the passage of right-to-carry laws and crime rates." Our results here further underscore the sensitivity of guns crime estimates to modeling decisions.[34] If one had to make judg-ments based on panel data models of the type used in the NRC report, one would have to conclude that RTC laws likely increase the rate of aggra-vated assault. Further research will be needed to see if this conclusion sur-vives as more data and better methodologies are employed to estimate the impact of RTC laws on crime.

Appendix A
Using Placebo Laws to Test the Impact of Clustering in the State Data

Using state-level data, we again conduct our experiment with placebo laws to examine the effects of clustering the standard errors. As seen in Tables A1–4, we find results similar to those generated with our county data: Without clustering, the Type 1 error rates are often an order of magnitude too high or worse for our murder and robbery regressions (see Tables A1 and A3). In fact, even *with* clustered stan-dard errors (Tables A2 and A4), the rejection of the null hypothesis

Table A1. Hybrid Model—Percentage of Significant Estimates (at the 5% Level)—Using Updated 2009 State-Level Data—Lott-Mustard Controls, without Clustered Standard Errors, 1977–2006 (without 1993 Data)

		Dummy Variable (%)	Trend Variable (%)
1. All 50 states + DC	Murder	47.1	67.2
	Robbery	46.0	61.7
2. Exact 32 states	Murder	48.5	57.3
	Robbery	51.2	71.1
3. Random 32 states	Murder	49.3	64.2
	Robbery	50.0	66.0

---

34.   For a quick and clear sense of how sensitive estimates of the impact of RTC laws are, see Appendix D, where we visually demonstrate the range of point estimates we obtain throughout our analysis.

Table A2. Hybrid Model—Percentage of Significant Estimates (at the 5% Level)—Using Updated 2009 State-Level Data—Lott-Mustard Controls, with Clustered Standard Errors, 1977–2006 (without 1993 Data)

		Dummy Variable (%)	Trend Variable (%)
1. All 50 states + DC	Murder	18.5	22.6
	Robbery	12.5	15.4
2. Exact 32 states	Murder	17.1	19.4
	Robbery	15.2	20.3
3. Random 32 states	Murder	22.0	22.7
	Robbery	16.3	18.2

Table A3. Dummy Variable Model—Percentage of Significant Estimates (at the 5% Level)—Using Updated 2009 State-Level Data—Lott-Mustard Controls, without Clustered Standard Errors, 1977–2006 (without 1993 Data)

		Dummy Variable (%)
1. All 50 states + DC	Murder	44.3
	Robbery	46.7
2. Exact 32 states	Murder	50.3
	Robbery	49.4
3. Random 32 states	Murder	51.9
	Robbery	50.8

Table A4. Dummy Variable Model—Percentage of Significant Estimates (at the 5% Level)—Using Updated 2009 State-Level Data—Lott-Mustard Controls, with Clustered Standard Errors, 1977–2006 (without 1993 Data)

		Dummy Variable (%)
1. All 50 states + DC	Murder	18.0
	Robbery	14.1
2. Exact 32 states	Murder	16.0
	Robbery	16.4
3. Random 32 states	Murder	22.7
	Robbery	14.3

(that RTC laws have no significant impact on crime) occurs at a relatively high rate. This finding suggests that, at the very least, we should include clustered standard errors to avoid unreasonably high numbers of significant estimates.

Appendix B
Panel Data Models Over the Full Period with no Covariates

The NRC panel sought to underscore the importance of finding the correct set of covariates by presenting panel data estimates of the impact of RTC without covariates but including county and year fixed effects. For completeness, this appendix presents these same estimates for the preferred models (with and without state trends) on both county and state data for the period from 1977 to 2006. If one compares the results from these four tables with no controls with the analogous tables using the preferred model for the same time period, one sees some interesting patterns. For example, if we compare the county results without state trends from both our preferred specification (Table 7a) and the no-controls specification (Table B1), we see that the results are quite similar in terms of magnitude and direction, although adding in our suggested covariates seems to both dampen the coefficients and reduce their significance. The basic story from our analysis is again strengthened: There seems to be virtually no effect of RTC laws on murder, while if there is *any* RTC effect on other crimes generally, it is a crime-*increasing* effect. The results are slightly less similar when we compare those from the models that include state trends (Tables 7b and B2). While we see that estimates are similar for murder, rape, robbery, and auto theft, the estimates for assault, burglary, and larceny change in either magnitude or direction (or both) when adding controlling factors to the model. In general, though, we only see decreases when adding state trends to either specification, and even then, the results are much too imprecise to make causal inferences. When we shift to a comparison of the state-level results, we again see similarities between the preferred and no-controls specifications. When looking at the results without state trends, we see that the estimates are very similar in terms of direction, although the no-controls estimates are often larger in magnitude and more

Table B1. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—No Controls, with Clustered Standard Errors, All Crimes, 1977–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−0.55	33.10	27.30	25.50*	33.50	35.90	38.00
	8.30	22.60	18.90	14.60*	21.50	22.00	25.50
2. Spline model	0.35	3.35*	3.20*	2.86**	3.42*	3.85*	4.27*
	0.76	1.94*	1.66*	1.36**	2.01*	2.00*	2.29*
3. Hybrid model							
Post-passage dummy	−3.48	21.40	14.30	14.30	21.40	21.50	21.30
	8.07	18.70	16.90	12.70	17.60	18.90	21.60
Trend effect	0.54	2.17*	2.41*	2.07*	2.24	2.66*	3.09*
	0.72	1.25*	1.27*	1.08*	1.48	1.54*	1.69*

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table B2. Estimated Impact of RTC Laws—Using Updated 2009 County-Level Data—No Controls, with State Trends and Clustered Standard Errors, All Crimes, 1977–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−2.80	−13.10	5.02	3.10	5.58	1.50	2.98
	5.03	10.60	9.31	7.71	9.47	10.50	11.70
2. Spline model	−0.54	−4.74	1.95	−0.37	−0.14	−0.78	−0.80
	1.23	4.06	2.30	2.33	2.52	2.45	2.61
3. Hybrid model							
Post-passage dummy	−2.52	−10.50	3.94	3.35	5.73	1.97	3.48
	5.22	10.10	10.20	8.27	10.20	11.40	12.80
Trend effect	−0.48	−4.52	1.87	−0.44	−0.26	−0.82	−0.87
	1.27	4.07	2.42	2.42	2.63	2.61	2.80

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates.
*Significant at 10%; **significant at 5%; ***significant at 1%.

statistically significant. When doing a similar comparison of the specifications that now adds in state trends, we also see similar results for nearly all crimes. The exception is aggravated assault, for which we see that our preferred specification produces more

Table B3. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—No Controls, with Clustered Standard Errors, All Crimes, 1977–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable	−1.79	8.33	11.70**	20.00**	24.70**	18.30***	16.60***
model	7.54	8.22	4.62**	7.90**	11.60**	6.69***	4.04***
2. Spline model	0.08	0.78	1.47**	1.98**	2.03*	1.73**	1.63***
	0.88	0.90	0.64**	0.96**	1.17*	0.72**	0.46***
3. Hybrid model							
Post-passage	−3.22	5.90	5.36	13.30*	19.60**	12.70**	11.00***
dummy	6.96	5.81	3.82	7.36*	9.00**	4.96**	3.69***
Trend effect	0.26	0.45	1.17*	1.24	0.90	0.99*	1.00**
	0.89	0.71	0.63*	0.96	0.86	0.56*	0.42**

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table B4. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—No Controls, with State Trends and Clustered Standard Errors, All Crimes, 1977–2006[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable	−0.31	−4.66**	0.62	3.43	8.38	1.10	0.92
model	3.73	2.00**	3.36	4.92	5.28	2.93	2.37
2. Spline model	0.78	−0.54	2.46***	0.29	−0.16	−0.20	−0.46
	0.93	0.92	0.91***	1.39	1.71	0.80	0.63
3. Hybrid model							
Post-passage	−0.80	−4.39**	−0.90	3.30	8.63	1.25	1.24
dummy	3.67	2.03**	3.37	5.30	5.17	3.23	2.55
Trend effect	0.80	−0.44	2.48***	0.21	−0.39	−0.23	−0.49
	0.93	0.91	0.92***	1.43	1.70	0.84	0.67

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates.
*Significant at 10%; **significant at 5%; ***significant at 1%.

negative estimates for the dummy model (although this result is not particularly precise). Again, when the comparison is taken as a whole, support is lacking for the view that RTC laws lead to reductions in crime.

Appendix C
Trimming the Sample to Address Questions of Model Fit

Given our concerns about how well the guns crime econometric models fit all 50 U.S. states (plus DC), we decided to examine the residuals from various regressions models. For example, one potentially important issue is whether one should include linear state trends in our models. To further explore this issue, we examined the variance of the residuals for the aggravated assault regression estimates using our preferred models on state data for the period through 2006—both with and without state trends.[35] In particular, we found that the residual variance was high for smaller states, even when we do not weight our regressions by population.[36] We explored how these "high–residual variance" states (defined from the aggravated assault regressions on our preferred model through 2006) might be influencing the results. We estimated our preferred model (both with and without state trends) after removing the 10% of states with the highest residual variance. This step is also repeated after removing the highest 20% of states in terms of residual variance. Our full-sample results for our preferred specification (which includes clustered standard errors, and is run over the entire time period) are shown in Tables 11a and b (without and with state trends, respectively). The results from our two trimmed set of states are presented below. Tables C1 and C2 should be compared to Table 11a (no state trends) and Tables C3 and C4 should be compared to Table 11b (adding in state trends). Removing high–residual variance states (based on the aggravated assault regressions) has little impact on the story told in Table 11a (no state trends): There was no hint that RTC laws reduce crime in Table 11a and this message comes through again in Tables C1 and C2. All three of these tables show at least some evidence that RTC laws increase aggravated

---

35. Since our most robust results across the specifications in this article were for aggravated assault, we focused specifically on the residuals obtained using assault rate as the dependent variable.

36. We removed the population weight for this exercise because it is likely that when regressions are weighted by population, the regression model will naturally make high-population states fit the data better. As a result, we expect that residuals for smaller states will be higher. We find, however, that the results are qualitatively similar even when we obtain the residuals from regressions that include the population-weighting scheme.

Table C1. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with Clustered Standard Errors, All Crimes, 1977–2006, Dropping States with Highest Residual Variance (Top 10%: MT, ME, WV, NH, TN)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−3.53	−0.98	4.33	5.04	6.80	1.38	5.75*
	4.02	3.95	3.15	4.41	6.27	3.05	2.96*
2. Spline model	−0.13	−0.50	1.16**	0.66	0.57	0.01	0.57
	0.62	0.56	0.57**	0.77	0.63	0.39	0.47
3. Hybrid model							
Post-passage dummy	−3.69	1.65	−1.21	2.53	5.26	1.69	3.94
	3.80	3.69	3.22	4.98	5.80	3.30	2.98
Trend effect	0.04	−0.58	1.21*	0.55	0.34	−0.07	0.40
	0.64	0.58	0.60*	0.86	0.58	0.43	0.50

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this "preferred" specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table C2. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with Clustered Standard Errors, All Crimes, 1977–2006, Dropping States with Highest Residual Variance (Top 20%: MT, ME, WV, NH, TN, NE, VT, HI, OH, KY)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable model	−4.99	−0.28	3.94	5.80	8.13	2.86	6.75**
	4.23	4.28	2.40	4.97	6.60	3.20	3.23**
2. Spline model	−0.16	−0.50	0.84*	0.90	0.71	0.29	0.71
	0.66	0.59	0.47*	0.83	0.70	0.37	0.50
3. Hybrid model							
Post-passage dummy	−5.38	2.53	0.15	2.09	6.16	1.91	4.39
	3.93	3.95	3.05	5.54	6.13	3.64	3.37
Trend effect	0.09	−0.61	0.83	0.81	0.43	0.21	0.52
	0.68	0.61	0.54	0.92	0.66	0.43	0.55

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this "preferred" specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures.
*Significant at 10%; **significant at 5%; ***significant at 1%.

Table C3. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with State Trends and Clustered Standard Errors, All Crimes, 1977–2006, Dropping States with Highest Residual Variance (Top 10%: MT, NH, VT, WV, KY)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable	1.17	−3.56	−0.13	2.28	7.82**	1.31	1.77
model	2.95	2.16	2.82	3.75	3.26**	2.03	1.66
2. Spline model	0.80	0.15	2.83***	0.32	−2.01**	−0.31	−0.21
	0.91	0.81	0.82***	1.37	0.83**	0.91	0.79
3. Hybrid model							
Post-passage	0.73	−3.71	−1.77	2.14	9.13***	1.51	1.93
dummy	3.12	2.32	2.80	4.04	3.23***	2.31	1.84
Trend effect	0.77	0.27	2.89***	0.25	−2.29***	−0.35	−0.27
	0.95	0.83	0.84***	1.42	0.83***	0.95	0.83

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this "preferred" specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures. *Significant at 10%; **significant at 5%; ***significant at 1%.

Table C4. Estimated Impact of RTC Laws—Using Updated 2009 State-Level Data—with Preferred Controls, with State Trends and Clustered Standard Errors, All Crimes, 1977–2006, Dropping States with Highest Residual Variance (Top 20%: MT, NH, VT, WV, KY, NE, NV, SD, ND, DE, IN)[a]

	Murder (%)	Rape (%)	Aggravated Assault (%)	Robbery (%)	Auto Theft (%)	Burglary (%)	Larceny (%)
1. Dummy variable	2.09	−2.88	−1.35	4.63	8.94***	1.42	2.41
model	2.97	2.29	2.78	3.44	3.18***	2.14	1.68
2. Spline model	0.92	0.25	2.42***	0.63	−2.11**	−0.43	−0.12
	0.97	0.83	0.80***	1.44	0.88**	0.99	0.83
3. Hybrid model							
Post-passage	1.69	−3.03	−2.50	4.39	10.00***	1.63	2.50
dummy	3.09	2.40	2.83	3.71	3.18***	2.40	1.87
Trend effect	0.88	0.32	2.48***	0.53	−2.35**	−0.47	−0.18
	1.01	0.84	0.81***	1.50	0.87**	1.02	0.87

[a]Estimations include year and county fixed effects, and are weighted by county population. Robust standard errors are provided beneath point estimates. The control variables for this "preferred" specification include incarceration and police rates (lagged one year to avoid potential endogeneity issues), unemployment rate, poverty rate, county population, population density, per capita income, and six demographic composition measures. *Significant at 10%; **significant at 5%; ***significant at 1%.

assault. Removing the high–residual variance states from the models with state trends does nothing to shake the Table 11b finding that RTC laws increase aggravated assault. The somewhat mixed results for auto theft seen in Table 11b also remain in Tables C3 and C4. Of the states dropped from Table C1, the following four states adopted RTC laws during the 1977–2006 period (with date of adoption in parentheses): Montana (1991), Maine (1985), West Virginia (1989), and Tennessee (1994). Of the additional states dropped from Table C2, the following four states adopted RTC laws during the 1977–2006 period (with date of adoption in parentheses): Ohio (2004), Kentucky (1996), Indiana (1980), and Oklahoma (1995).[37] Results from Table C3 come from dropping similar RTC states to Table C1, although Kentucky (1996) is dropped rather than Tennessee, and New Hampshire (1959) is dropped rather than Maine.[38] Finally, in addition to the five RTC states that were dropped in Table C3, Table C4 dropped the following four RTC states: Nevada (1995), South Dakota (1986), North Dakota (1985), and Indiana (1980).

Appendix D
Summarizing Estimated Effects of RTC Laws Using Different Models, State Versus County Data, and Different Time Periods

This appendix provides graphical depictions of sixteen different estimates of the impact of RTC laws for the dummy and spline models for specific crimes using different data sets (state and county), time periods (through 2000 or through 2006), and models (Lott-Mustard versus our preferred model and with and without state trends). For example, Figure D1 shows estimates of the impact on murder using

---

37. In implementing our protocol of dropping high–residual variance states, we examined the residuals of the dummy and spline models separately to identify the high-variance states. While they match across models for three of the four tables, in the case of Table C4, the ordinal rank of the states in terms of residual variance were slightly different for the dummy versus the spline model. For this table, Indiana had the 9th highest residual variance when looking at the dummy model results, while North Dakota had the 11th highest variance. For the spline results, the residual variance ranks of these two states were reversed. Thus, for this table, we dropped both states to estimate our regressions.

38. The dropped states are slightly different between Tables C1 and C3, as well as between Tables C2 and C4, because the state ranks based on residual variances differed when the models were run with and without state trends.
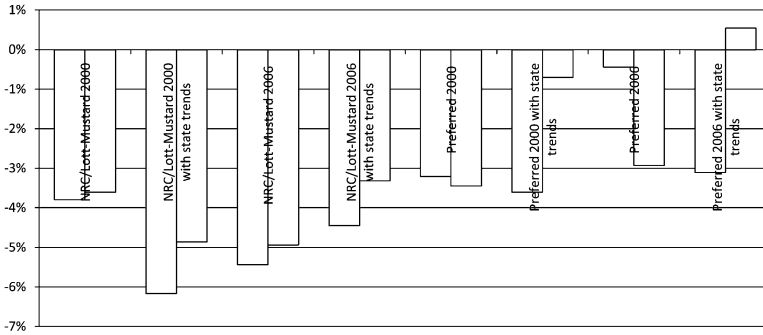
Figure D1. Various Murder Estimates (Dummy Model).
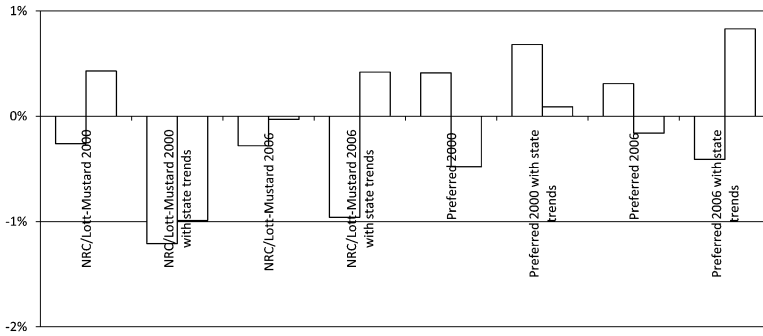
Figure D2. Various Murder Estimates (Spline Model).
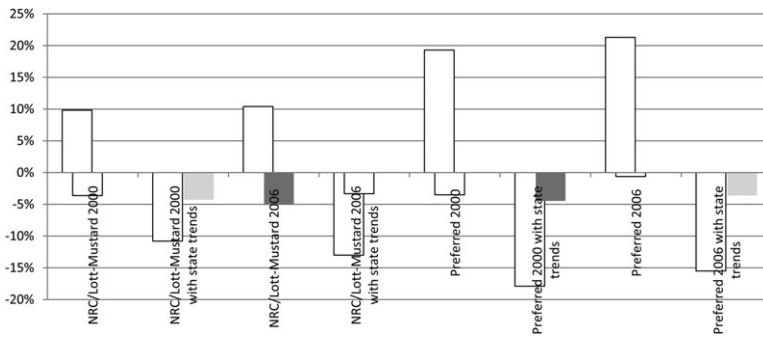
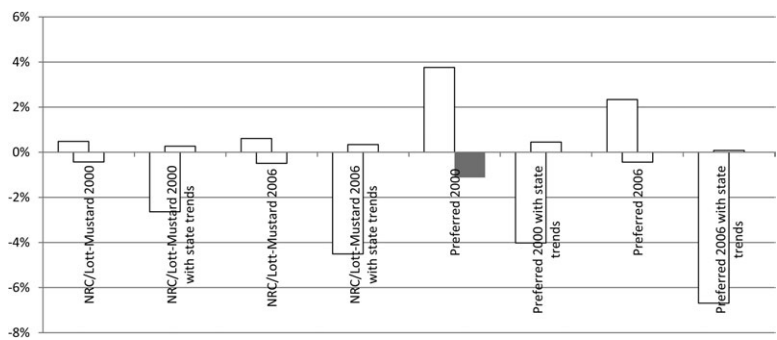**Figure D3.** Various Rape Estimates (Dummy Model).
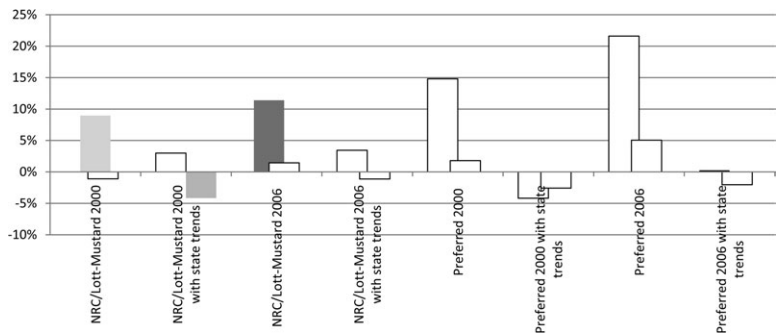
Figure D4. Various Rape Estimates (Spline Model).

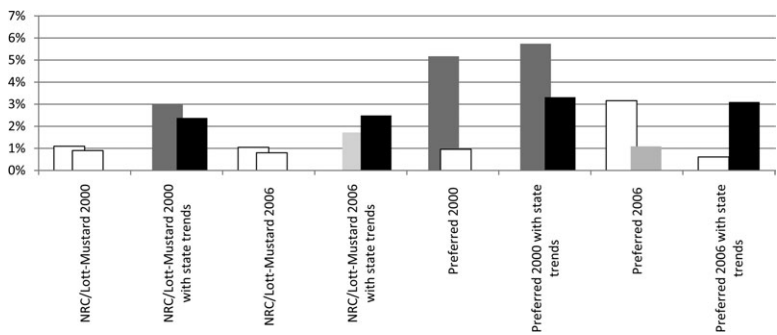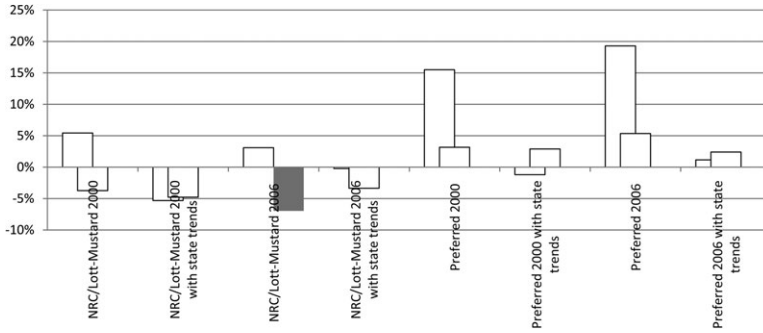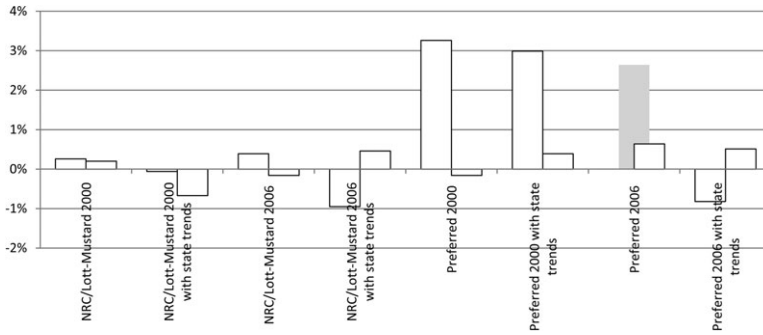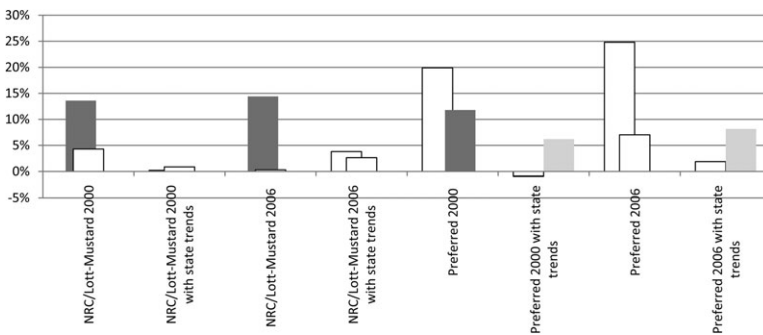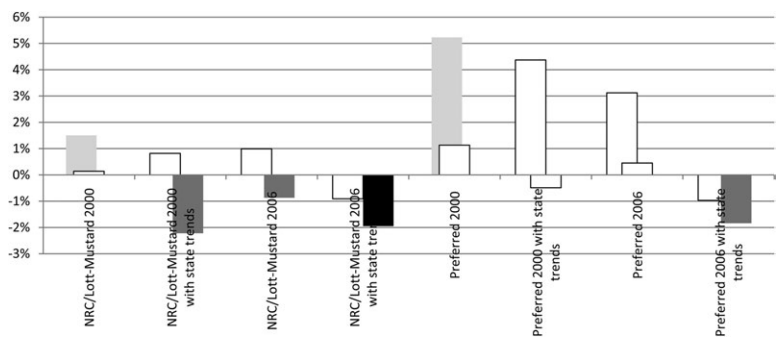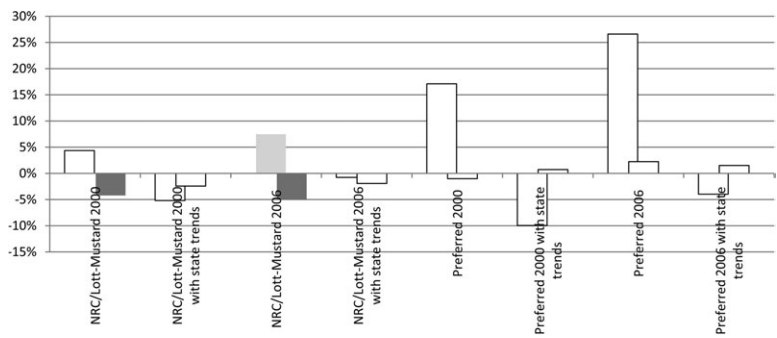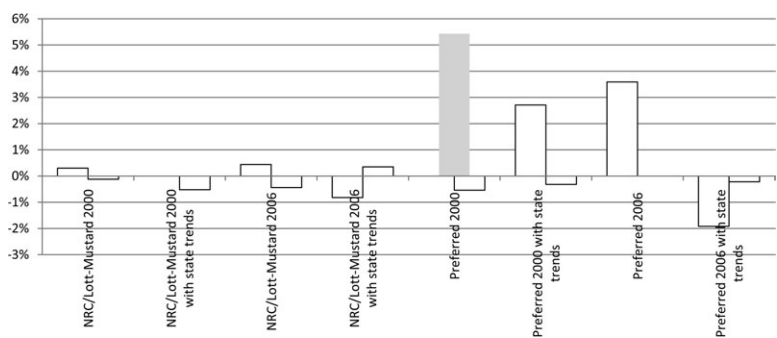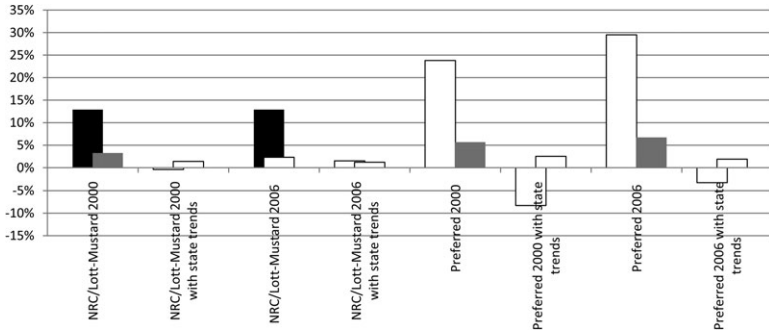**Figure D5.** Various Assault Estimates (Dummy Model).

Figure D6. Various Assault Estimates (Spline Model).

Figure D7. Various Robbery Estimates (Dummy Model).

Figure D8. Various Robbery Estimates (Spline Model).

Figure D9. Various Auto Theft Estimates (Dummy Model).

Figure D10. Various Auto Theft Estimates (Spline Model).

Figure D11. Various Burglary Estimates (Dummy Model).

Figure D12. Various Burglary Estimates (Spline Model).

Figure D13. Various Larceny Estimates (Dummy Model).
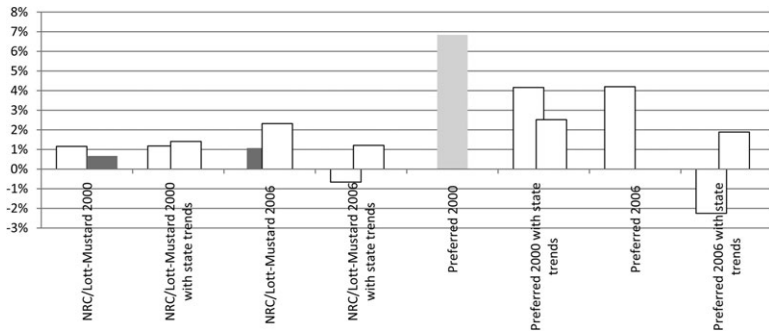
Figure D14. Various Larceny Estimates (Spline Model).

the dummy model, designed to capture the average effect of RTC laws during the post-passage period. The first bar in each of the eight groupings corresponds to county-level estimates; the second bar corresponds to state-level estimates, for a total of sixteen estimates per figure. The value of the figures is that they permit quick visual observation of the size and statistical significance of an array of estimates. Note, for example, that none of the estimates of RTC laws on murder in either Figure D1 or D2 is significant at even the 0.10 threshold. This sharp contrast to the conclusion drawn by James Q. Wilson on the NRC panel is in part driven by the fact that all the estimates in this appendix come from regressions in which we adjusted the standard errors by clustering. In contrast to the wholly insignificant estimates for murder, the estimates of the impact of RTC laws on aggravated assault in Figure D6 are generally significant as

indicated by the shading of the columns, where again no shading indicates insignificance, and the shading darkens as significance increases (from a light gray indicating significance at the 0.10 level, slightly darker indicating significance at the 0.05 level, and black indicating significance at the 0.01 level). Note that the overall impression from Figure D6 is that RTC laws increase aggravated assault. Even in Figure D6, though, one can see that some of the estimates differ between county- and state-level data and tend to be strongest in state data controlling for state trends.

Figure D5, which provides estimates of the effect of RTC laws on aggravated assault using the dummy model (rather than the spline model of Figure D6), reveals that the conclusion that RTC laws increase aggravated assault is model dependent: If the dummy model is superior, and if we confine our attention to the complete 1977–2006 data set, the conclusion that RTC laws increase aggravated assault only holds in the Lott-Mustard county data model. In Figure D14, the state-level estimates of the preferred specifications (without state trends) through 2000 and 2006 are essentially zero (no impact), so only the county-level estimates show up in the graph.

References

Angrist, Joshua, and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton, NJ: Princeton University Press.

Autor, David, John J. Donohue, and Stewart Schwab. 2006. "The Costs of Wrongful-Discharge Laws," 88 *Review of Economics and Statistics* 211–31.

Ayres, Ian, and John J. Donohue. 2003a. "The Latest Misfires in Support of the More Guns, Less Crime Hypothesis," 55 *Stanford Law Review* 1371–98.

———. 2003b. "Shooting Down the More Guns, Less Crime Hypothesis," 55 *Stanford Law Review* 1193–312.

———. 2009. "More Guns Less Crime Fails Again: The Latest Evidence from 1977-2006," 6 *Econ Journal Watch* 218–38. Available at: http://www.aier.org/aier/publications/ejw_com_may09_ayresdonohue.pdf (accessed September 1, 2009).

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" 119 *Quarterly Journal of Economics* 249–75.

Black, Dan A., and Daniel S. Nagin. 1998. "Do Right-to-Carry Laws Deter Violent Crime?" 27 *Journal of Legal Studies* 209–19.

Collins, Gail. 2009. "Have Gun, Will Travel," *The New York Times*. July 31.

Donohue, John J. 2003. "The Impact of Concealed-carry Laws" in J. Ludwig and P. J. Cook, eds., *Evaluating Gun Policy,* 287–324. Washington, DC: Brookings Institution Press.

Donohue, John J., and Justin Wolfers. 2009. "Estimating the Impact of the Death Penalty on Murder." 11 *American Law and Economics Review* 249–309.

Fryer, Roland, Paul Heaton, Steven Levitt, and Kevin Murphy. 2005. "Measuring the Impact of Crack Cocaine," NBER Working Paper Series No. W11318. National Bureau of Economic Research, Cambridge, MA.

Horowitz, Joel L. 2001. "Should the DEA's STRIDE Data Be Used for Economic Analyses of Markets for Illegal Drugs?" 96 *Journal of the American Statistical Association* 1254–71.

Levitt, Steven D. 2004. "Understanding Why Crime Fell in the 1990s: Four Factors That Explain the Decline and Six That Do Not," 17 *Journal of Economic Perspectives* 163–90.

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

———. 2004. "Right-to-Carry Laws and Violent Crime Revisited: Clustering, Measurement Error, and State-by-State Breakdowns." Available at: http://ssrn.com/abstract=523002 (accessed September 1, 2009).

———. 2008. "Do Guns Reduce Crime?" *Intelligence Squared Debate Series.* Available at: http://intelligencesquaredus.org/wp-content/uploads/Guns-Reduce-Crime-102808.pdf (accessed September 1, 2009).

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

Ludwig, J. 1998. "Concealed Gun-Carrying Laws and Violent Crime: Evidence from State Panel Data," 18 *International Review of Law and Economics* 239–54.

Maltz, Michael D. 2006. Analysis of Missingness in UCR Crime Data. NCJ 215343. Washington, DC: U.S. Department of Justice.

Maltz, Michael D., and J. Targonski. 2002. "A Note on the Use of County-Level Crime Data," 18 *Journal of Quantitative Criminology* 297–318.

———. 2003. "Measurement and Other Errors in County-Level UCR Data: A Reply to Lott and Whitley," 19 *Journal of Quantitative Criminology* 199–206.

Moody, Carlisle E., and Thomas B. Marvell. 2008. "The Debate on Shall-Issue Laws," 5 *Econ Journal Watch* 269–93. http://www.aier.org/ejw/archive/doc_view/3610-ejw-200809?tmpl=component&format=raw.

Moulton, Brent. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," 72 *Review of Economics and Statistics* 334–38.

National Research Council. 2005. *Firearms and Violence: A Critical Review*. Washington, DC: National Academies Press.

Plassman, Florenz, and John Whitley. 2003. "Confirming More Guns, Less Crime," 55 *Stanford Law Review* 1313–69.

U.S. Department of Justice–Federal Bureau of Investigation. 2009. "Crime—National or State Level: Data with One Variable." Uniform Crime Reporting Statistics. Available at: http://www.ucrdatatool.gov/Search/Crime/State/TrendsInOneVar.cfm. (accessed September 1, 2009).

Wilson, James Q. 2008. "What Do We Get from Prison?" *The Volokh Conspiracy.* Available at: http://volokh.com/posts/chain_1213046814.shtml (accessed November 20, 2009).

Wooldridge, Jeffrey M. 2003. "Cluster-Sample Methods in Applied Econometrics," 93 *American Economic Review* 133–38.

Zimring, Franklin, and Gordon Hawkins. 1997. "Concealed Handguns: The Counterfeit Deterrent," 7 *The Responsive Community* 46–60.