The Right to Carry Has Not Increased Crime: Improving an Old Debate Through Better Data on Permit Growth Over Time

By William English, PhD*

Draft: July 18, 2021

*Over the last 30 years, a majority of US states adopted Right-to-Carry (RTC) laws at the same time that crime rates dramatically decreased. A large literature has examined whether RTC laws contributed to or slowed this decline in crime, with most studies concluding that they have no significant effect on crime. However, this research has been plagued by methodological challenges, many of which are exacerbated by the common approach of modeling the effect of RTC laws using a binary dummy variable to indicate a one-time change in policy. Recently, Donohue, Aneja and Weber (2019*a*) have employed a novel synthetic control approach which they suggest indicates that RTC laws significantly increase violent crime. However, we show that this analysis is highly sensitive to modeling choices, and Donohue et al. chose a specification that has been criticized by Kaul et al. (2017) as mistaken because it prevents covariates from exercising any influence on the development of predicted crime rates. Correcting this to properly incorporate covariates dramatically changes the estimated effect in many states; and comprehensive synthetic control analysis reveals no significant effect on crime. Given the methodological challenges inherent in binary approaches to modeling the effects of RTC laws, we gather data on the growth of carry permits in states over time, which allows us to investigate the phenomenon of interest - the actual ability to carry - in a manner that is theoretically more valid and econometrically more powerful. Employing two different methods for estimating missing data - modeling the growth of permits as a logistical growth process and imputing missing data using the Amelia II package - we find that the growth in carry permits has no effect on violent crime rates, homicide rates, firearm homicide rates, or non-firearm homicide rates. This study provides further, strong evidence that the dramatic growth in the ability to carry firearms for self-defense in recent decades has not exacerbated crime rates.*

I. Introduction

Laws that permit ordinary citizens to carry handguns for self-defense have expanded significantly over the last 30 years in the United States, as have the number of concealed carry permits issued,

* William English: McDonough School of Business, Georgetown University, william.english@georgetown.edu.

at the same time that violent crime and homicide rates have fallen by roughly 50%. Whether right to carry (RTC) laws have accelerated these declines in crime or slowed them has been the subject of extensive scholarly debate that has yielded contradictory findings and little consensus that these laws have had any effect at all.

A recent, comprehensive review of this literature published by RAND concludes that there is "inconclusive evidence" for the effect of shall-issue laws on robberies, assaults, rapes, firearm homicides, and total homicides (Gresenz, 2018/2020).[1] The review suggests that "limited evidence" exists that shall-issue laws may increase violent crime overall, based on two recent studies by Donohue, Aneja and Weber (2019*a*) and Durlauf, Navarro and Rivers (2016), but also notes that this literature has been plagued by methodological challenges. Indeed, this is a theme of Durlauf, Navarro and Rivers (2016)'s paper, which, after critically examining a number of common modeling choices, concludes that "the evidence that shall issue right-to-carry laws generate either an increase or decrease in crime on average seems weak."

This recent judgment echoes the conclusions reached by a 2005 National Research Council report on the subject of firearms and violence which likewise emphasized methodological difficulties with this research, particularly: "(a) the sensitivity of the empirical results to seemingly minor changes in model specification, (b) a lack of robustness of the results to the inclusion of more recent years of data (during which there were many more law changes than in the earlier period), and (c) the statistical imprecision of the results (Council et al., 2005)."

Many of these methodological challenges are exacerbated by the common approach in this literature of using a binary coded "dummy variable" to assess a law's impact. The years before a law is passed in a state are coded as 0's and years after as 1's, providing little variation from which to draw inferences, while also making models vulnerable to confounding from other secular trends and poorly equipped to capture variation in the impact of a law over time. Indeed, in addition to the methodological issues flagged by the NRC report, the RAND essay on "Methodological Challenges to Identifying the Effects of Gun Policies" specifically identifies coding of the time effects as one of the greatest challenges in this literature while noting limitations with spline and hybrid model approaches (Schell, 2018).

This is a challenge that has been increasingly well understood within econometric literature focused on difference-in-differences methods. As Wolfers (2006) noted in a seminal article examining the effect of divorce laws, modeling time trends is critical in conventional difference-in-differences analysis with binary treatments because the effects of a legal change may take time to develop. More recently, Goodman-Bacon (2018) has also shown that when there is variation in the timing of

[1] This review was updated in April of 2020, with the most recent version available at: `https://www.rand.org/research/gun-policy/analysis/concealed-carry/violent-crime.html` and is part of a larger research project on "The Science of Gun Policy" ed by Morrall (2018).

treatments (e.g. laws passed in different states in different years), standard difference-in-difference estimators will be biased if the effects of these laws change over time.

These are particularly serious concerns in the case of right-to-carry laws, as there are strong theoretical reasons to expect that the impact of these laws will take time to develop. The number of concealed carry permits issued the year an RTC law goes into effect is generally small, comprising only fraction of a percent of a state's population. However, in many states, the number of permit holders has grown to around 10% of the adult population over time.[2] If the mechanism through which RTC laws affect crime involves the actual ability to carry a handgun, then the mere passage of a law is a poor proxy for this. Rather, what needs to be evaluated is how the growth of the number of people permitted to carry over time affects crime.

Using data from 12 states that report the number of carry permits issued every year that their RTC has been in effect, along with partial data obtainable for remaining states, we model the growth of permitted carry in general and develop estimates for permitted carry rates in states that have missing data. This allows us to assess the effect of permitted carry rates on crime rates in a manner that is theoretically more valid and econometrically more powerful than prior approaches using binary coding. Examining model specifications with a variety of controls, we find that all model specifications indicate that carry permit rates have no significant effect on homicide rates, firearm homicide rates, non-firearm homicide rates or violent crime rates. While there is weak evidence under some specifications that carry permit rates may be associated with higher property crime rates, consistent with the original thesis of Lott and Mustard (1997) that criminals will substitute away from crimes that might involve an armed defendant, contrary to the claims of Donohue, Aneja and Weber (2019*a*), we find no evidence that the ability to carry is associated with increased violent crime rates or homicide rates.

This paper proceeds as follows: We begin with a brief literature review in which we examine common modeling choices in this literature with a focus on choices that produce the most prominent outlier results reported by Donohue, Aneja and Weber (2019*a*) suggesting a positive association between carry and violent crime. We show that simple model corrections that are more theoretically defensible undo their results, including their recent synthetic control results. In the second section, we summarize state level data on carry rates over time and introduce two approaches for estimating the growth of rates of carry in states with missing data. Corresponding subsections report analysis of the effects of carry permit rates on crime accompanied by a variety of robustness checks. Overall, these results suggest that higher rates of permitted carry have no significant effect on homicide rates or violent crime rates. We conclude with a summary of the state of the literature and suggestions

[2] Note that the denominator in this statistic is adult population. For purposes of continuity with Donohue et al.'s data, we use percentage of a state's total population for the analysis introduced later in this paper, which is the reason that those rates are lower.

for future research.

II. Literature Review and Modeling Controversies

The review of literature on concealed carry and crime conducted by RAND in 2018 and updated in 2020 provides a thorough introduction to the history of scholarly debates on this subject.[3] In summarizing the current state of the literature, the review focuses on 18 studies conducted in recent years that the authors believe did not have serious methodological concerns, having observed that this literature is rife with methodological challenges. Based on these 18 studies, the authors judge that there is "inconclusive evidence" that shall-issue laws have any effect on robberies, assaults, rapes, firearm homicides, and total homicides.[4] However, the review suggests that there is "limited evidence" that shall-issue laws may increase violent crime. In coming to this conclusion, the authors rely heavily on recent work published by Donohue, Aneja and Weber (2019*a*), while discounting studies by Hamill et al. (2019), Helland and Tabarrok (2004), Plassmann and Whitley (2003) that found no effect, in part because two of these studies relied on older data (as did Durlauf, Navarro and Rivers (2016)).

The RAND review rightly notes the value of using data that includes more recent years and it highlights a number of specific methodological considerations that need to be taken into account in this research.[5] We review the most serious of these challenges below because they help explain the fragility of the outlier results of Donohue, Aneja and Weber (2019*a*) while also demonstrating why analysis that incorporates the growth of permits over time is such a significant methodological advance.

A. *Challenge 1: Coding of State Laws- transition dates and may-issue states with high permit rates*

First, there have been disagreements in the literature regarding the coding of state laws, small changes in which can have a large impact on model conclusions.

One challenge concerns the coding of states that have "may-issue" policies, which allow authorities to issue carry permits with discretion. The criteria for discretion vary between these state, with some states issuing permits at levels that rival or exceed "shall-issue" states, while others seldom issue permits at all, making them look more like no-carry states. Connecticut is a may-issue

[3]It should be noted that Lott has argued that the RAND review neglects a number of studies that found that RTC laws reduce violent crime. For a summary of these, see Lott (2011).

[4]Note that empirical research evaluating the effect of RTC laws has generally not focused on examining their effects on gun accidents and suicide, presumably because such laws, which deal with the ability of individuals to carry in public, do not affect people's ability to posses guns at home. In recent reviews, RAND notes that literature examining the impact of RTC laws on suicides or accidents is sparse and judges that "Evidence for the effect of shall-issue concealed-carry laws on unintentional firearm injuries and deaths is inconclusive," and "Evidence for the effect of shall-issue concealed-carry laws on total suicides, firearm suicides, and firearm self-injuries is inconclusive." See `https://www.rand.org/research/gun-policy/analysis/concealed-carry/unintentional-injuries.html`, `https://www.rand.org/research/gun-policy/analysis/concealed-carry/suicide.html`

[5]Key methodological concerns are summarized more succinctly in related reports that are part of this larger RAND project: `https://www.rand.org/research/gun-policy/analysis/essays/methodological-challenges-to-identifying-the-effects-of-gun-policies.html`

state that has often been classified as a shall-issue carry state because more than 9% of the adult population has been licensed to carry. However, Massachusetts is a may-issue state that has often been classified as a no-carry state, including by Donohue, Aneja and Weber (2019*a*), despite the fact that more than 7% of the adult population has been licensed to carry - a rate that is on par with Texas and above roughly a dozen shall-issue states.

Related to this, there have been disputes about when to begin counting some states as shall-issue states. For example, the major legislative change that turned Virginia into shall-issue concealed carry state occurred in 1988, when the applicable law was changed to the following:

> "The court, after consulting the law-enforcement authorities of the county or city and receiving a report from the Central Criminal Records Exchange, shall issue such permit if the applicant is of good character, has demonstrated a need to carry such concealed weapon, which need may include but is not limited to lawful defense and security, is physically and mentally competent to carry such weapon and is not prohibited by law from receiving, possessing, or transporting such weapon (Cramer and Kopel, 1994)."

However, lack of clarity about the criteria for issuance, particularly the language concerning a demonstrated "need to carry," which courts in some jurisdictions used to deny permits, resulted in another legislative change in 1992, requiring judges to renew permits "unless there is good cause shown for refusing to reissue a permit." Yet another legislative change was made in 1995, which removed the language of "need" altogether; and it was not until 1995 that concealed carry permit applications started rising dramatically. Manski and Pepper (2018) follow Lott and Mustard (1997) and code 1989 as the transitional date for Virginia in their detailed study of the effects of concealed carry in Virginia, Maryland, and Illinois, which found heterogeneous effects that vary by type of crime and over time. However, the RAND review discards Manski and Pepper (2018) because of their use of this earlier date.

Donohue, Aneja and Weber (2019*a*) use the 1995 date and point out that the 2005 National Research Council report endorsed using 1995 as well. Their reasoning is straightforward. Data show that it was not until 1995 that the carry permit applications in Virginia started to dramatically increase. In the appendix to Donohue, Aneja and Weber (2019*a*), they illustrate this with a simple graph reproduced below in figure 1. However, while the point is well taken, it begs the question as to why analysis of other states should not be treated in the same manner.

For example, Figure 1 shows permit applications in Colorado in the years following the legislative change that introduced shall-issue. As in Virginia, for the first few years after the change, only a few thousand individuals applied, and it is not until after 5 years that application numbers begin to break 20,000, which is the level at which Donohue, Aneja and Weber (2019*a*) begin to consider Virginia a carry state. Vertical lines in both figures indicate this discrepancy in coding.

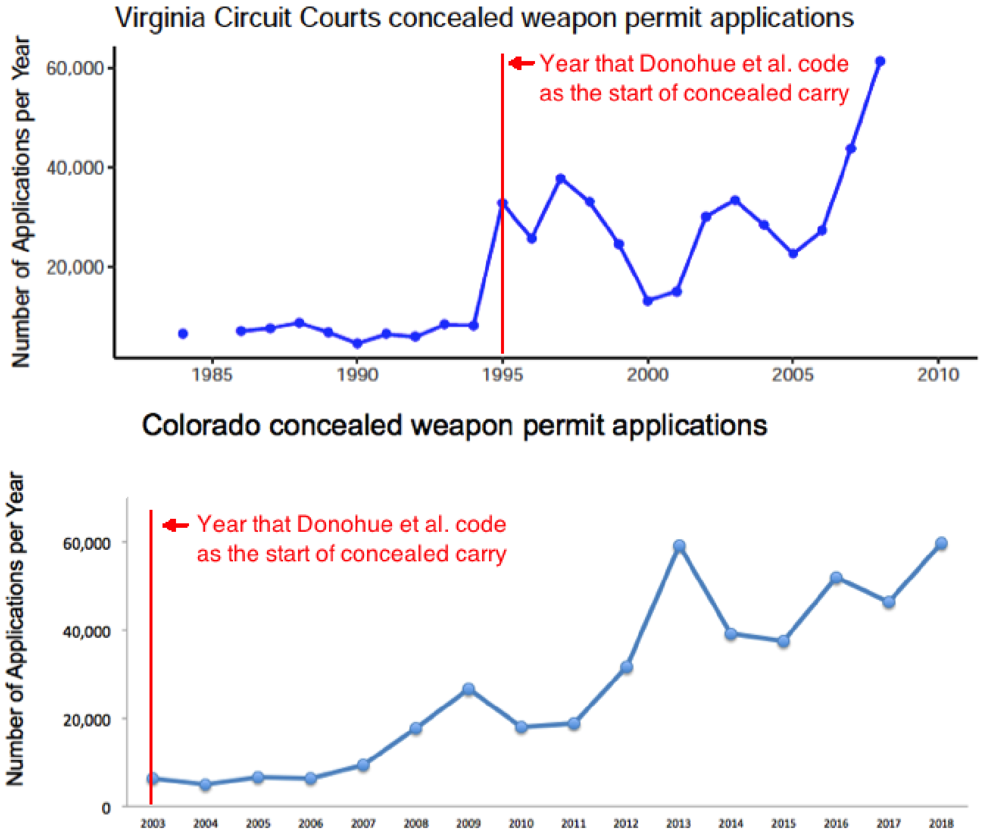


Figure 1. : Coding of the Start of Shall-Issue in Virginia vs Colorado in Donohue et al. (2019)

As we will see, these patterns are not unusual, with few permits issued the years immediately following a shall-issue legal change but rates accelerating rapidly in later years. For example, although only about 1% of the adult population in Virginia was licensed by 1996, this grew to over 9% by 2018. In Florida, less than half a percent of the adult population was licensed by 1990, a little over two years after its RTC law went into effect at the end of 1987, while this grew to over 10% by 2018. In Michigan, less than a half of a percent of the adult population was licensed in the year after its RTC law passed in 2001, but this grew to about 9% by 2018.

As noted by many who have reviewed this literature, results are highly sensitive to the cut points at which one codes the binary change from a no-carry to a carry regime. One significant advantage of using the actual number of permits issued to analyze the effects of carry on crime is that this make the analysis less sensitive to disputed cut points in an all-or-nothing binary coding approach while also modeling the phenomenon of interest more accurately.[6]

[6]Note that the problem of binary cut points also affects other research that has claimed a positive relationship between RTC laws and crime, particularly Siegel et al. (2017).

B. Challenge 2: Error Modeling

A second set of methodological issues flagged by both the Rand review and the 2005 NRC report relate to error modeling, statistical power, and poorly calibrated standard errors.

As Helland and Tabarrok (2004) note in one of the earliest methodological examinations of this literature, "Failing to take into account serial correlation and grouped data can dramatically reduce standard errors suggesting greater certainty in effects than is actually the case." The treatment of error modeling was explicitly discussed in the NRC report. While the report sided with Lott in not considering state-level clustered error essential, Aneja, Donohue III and Zhang (2011) argued emphatically in favor of using state-level clustering.

More recent analysis by Moody and Marvell (2020) further supports this approach in the context of investigating the impact of policy changes on crime, demonstrating that standard error biases are substantial when standard errors are not clustered in the presence of autocorrelation. Moreover, they show that, without state level clustering, extreme standard error bias can occur when a small number of policy changes are coded using a binary variable, although there are also cases for which clustering alone is not sufficient to correct standard error bias. Their analysis illustrates why these problems are exacerbated by the use of a binary dummy variable to code a policy change, as the series of 0's and 1's is highly autocorrelated, echoing Bertrand, Duflo and Mullainathan (2004)'s observation that differences-in-differences models are liable to produce incorrect standard error estimates. Another strength of using actual permit data as indicative of an RTC law's treatment effect over time, rather than binary coding of a one-time policy change, is that it mitigates the problem of autocorrelation. However, because of the fixed effects design, we follow Donohue, Aneja and Weber (2019*a*) in using state clustered errors in all panel-based analysis.

A second, related virtue of using the growth in permits over time to analyze the impact of carry on crime is that it provides a more compelling framework to deal with the challenge of trends. Binary dummy variable approaches test for a difference in average crime rates before and after an RTC law goes into effect, while spline model approaches test for whether trends in crime are altered. Donohue and Ayres (2003) develop a "hybrid model" meant to combine both approaches to measure the immediate and long-run impact of these policy changes. However, this generic approach to modeling policy effects over time raises a question of whether and how to control for state trends, as these are significant in many model specifications and could confound inferences drawn from spline and hybrid models. In versions of their 2017 working paper (Donohue, Aneja and Weber, 2017*a*,*b*) Donohue et al. further experiment with spline models, which effectively model RTC laws as having a greater impact over time. In nearly all specifications, including the most defensible specifications, their spline models find no significant effects on murder or violent crime rates.

How to appropriately model trends is one focus of the critical exchange between Moody and Marvell (2018) and Donohue (2018), in which the former point out that excluding state specific trends could introduce serious omitted variable bias into models. Donohue acknowledges that the issue of state-specific trends is a "challenging one." On the one hand, "adding state trends could be helpful if it corrects for an important omitted variable, but it could also be harmful because the state trends will not just pick up a pre-existing trend but will also pick up any effects of the RTC law that unfold over time in a similar fashion to the pre-existing trend." Moreover, from 1977 to 2014, movements in crime rates are clearly not linear, showing a large rise then a large fall, which is a problem for the use of state-specific linear trends over long time periods. The importance of trends to difference-in-differences analysis is also the reason that Donohue, Aneja and Weber (2019*a*) invest considerable effort examining parallel trends conditions and in recommending a synthetic controls approach.

Concerns about trends confounding the results are considerably mitigated by the use of the number of carry permits issued as a continuous treatment variable, which moves away from a difference-in-differences framework. This data varies by year and reflects changes in the phenomenon of interest rather than a crude before/after comparison or generic trends that are assumed to persist at the same rate far into the future.

One of the more serious methodological concerns with Donohue, Aneja and Weber (2019*a*)'s analysis, which has been noted by a number of critics, is their use of "analytic weights" based on state population. In their most recent paper, they do not discuss or explain their use of analytic weights except for the suggestion made in a table note that this was done so that OLS estimations would be "weighted by population." However, the variables being examined are already in terms of population rates (e.g. number of violent crimes per 100,000 residents), which are, by definition, adjusted for population.

Moody and Marvell (2019) recently published an extended critique of Donohue, Aneja and Weber (2019*a*)'s use of analytic weights. They note that Donohue, Aneja and Weber (2019*a*) never explain their rationale for employ these weights, but the standard rationale for doing so is to address heteroskedasticity. However, there's little reason to believe that analytic weights are appropriate for redressing heteroskedasticity in this context. Fist, there's no theoretical basis for expecting that data from larger states is measured with greater precision than the exact same data from smaller states - that, for example, the murder rate in California is measured more accurately than the murder rate in Delaware. These data are based on the same reporting standards and do not represent averages derived from sub-samples of different sizes. Second, as Moody and Marvell (2019) show, testing for heteroskedasticity and weighting to correct for it using Feasible Generalized Least Squares yield results that are dramatically different from the analytic weights approach but almost identical to results without analytic weights. Thus, the most accurate method of dealing with heteroskedasticity in this context rejects the use of analytic weights.

As Durlauf, Navarro and Rivers (2016) note in their criticism of this practice in the context of county level analysis, "the use of population weights will overweight observations from more populous counties, leading to invalid confidence intervals, and potentially misleading point estimates." For this and other reasons, they conclude that the "use of population weights to control for heteroskedasticity in crime rates, has so little evidentiary support that it is reasonably excluded from the analysis." [7] Moody and Marvell (2020) likewise judge that the use of analytic weights is unwarranted based on Monte Carlo simulations, noting that "standard error bias increases with greater regression weight," and the net effect is that large states end up having an outsized influence on the dummy coefficient.

In their response to Moody and Marvell (2019), Donohue, Aneja and Weber (2019*b*) say that addressing the problem of heteroskedasticity is not their "primary rationale" for using analytic weights. Rather, they claim that they overweight large states because laws in such states affect more people. They argue that this is a "conceptually superior" approach because it better captures the net impact on Americans. However, this misconstrues the aim of this analysis, which is to assess the comparative effect of RTC laws. States are the relevant unit of analysis. At a conceptual level, we want to know how RTC laws affect crime rates in states that adopt them versus states that don't. Recall that models include state fixed effects, and we can also include population as a control variable, which we do in some later specifications, if there is reason to believe that the size of states matters. However, the use of analytic weights effectively requires that small states with RTC laws must show much larger reductions in crime for them to be judged equivalent to large non-RTC states like California and New York when these states have only a small decrease in crime. The net effect of using analytic weights in this context is to greatly inflate the impact of large states in Donohue et al.'s analysis. If the aim is to accurately asses the effect of RTC laws on state crime rates, analytic weights are clearly inappropriate.

As Table 1 shows, removing analytic weights from Donohue, Aneja and Weber (2019*a*)'s panel model substantially changes the results. The large and significant association claimed between carry laws and violent crime rates disappears, while the coefficient of the non-significant association between carry laws and firearms murder rates changes from positive to negative.

Although Donohue, Aneja and Weber (2019*a*) do not mention it in their paper, they also use analytic weights in a manner that is arguably misleading to construct the first figure in their paper, which summarizes the change in violent crime rates between two dates - 1977 and 2014 - in states

[7] As Durlauf, Navarro and Rivers (2016) note, Solon, Haider, and Wooldridge (2013) provide an informative examination of justifications for using weights and ways in which they can be misapplied in empirical work.

Table 1—: Donohue et al. Panel Data Estimates with State- & Year-Fixed Effects, Donohue Regressors, 1977–2014, Yield Different Results with Model Corrections

	(1) Murder Rate	(2) Firearm Murder Rate	(3) Nonfirearm Murder Rate	(4) Violent Crime Rate	(5) Property Crime Rate
Initial Donohue Model	2.27	2.90	1.53	9.02**	6.49*
	(5.05)	(6.74)	(3.32)	(2.90)	(2.74)
No Analytic Weights	0.32	-1.22	2.23	0.65	4.68*
	(3.98)	(5.37)	(3.68)	(3.68)	(2.26)
No Analytic Weights &	-13.14	-13.96	6.80	2317.71	15808.03†
No Ln Transformation	(36.44)	(32.33)	(9.65)	(1990.37)	(8763.34)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Robust clustered standard errors in parentheses. All models use Donohue et al. (2019)'s data and regressors, include year- and state-fixed effects, and employ clustered errors by state.

that adopted RTC laws versus states that didn't. They intend this to show that "The decline in violent crime rates has been far greater in states with no RTC." However, the information conveyed in this figure provides little evidence in support of this larger thesis. For example, simply shifting these dates one year, to 1978 and 2013, reveals a substantively different pattern of rate changes. Moreover, the patterns are remarkably different when one examines the actual data without analytic weights.

Figure 2 reproduces Donohue et al.'s first bar graph, which used analytic weights to calculate the change in average violent crime rates across nine states that never adopted RTC laws between 1977 and 2014, and compares this with what the actual rate data shows. Rather than a 42.3% decline between these two years, we witness a 24.4% decline. Again, this simple comparison says little with regard to the larger inferential question about the net impact of carry on crime, but it demonstrates how the reporting of basic rates, which are already denominated on a per-capita basis, can be distorted through the use of analytic weights.

Another modeling choice that affects Donohue, Aneja and Weber (2019*a*)'s analysis in more subtle ways is the use of natural log transformations of crime rates. As Manski and Pepper (2018) note, this is commonly done in the literature, although Manski and Pepper prefer to focus on studying state-year crime rates (without log transformations) as the phenomenon of primary concern.

There can be both technical and theoretical reasons to log-transform rate data. From a technical perspective, log-transformations can change highly skewed data into normally distributed data, with properties that are more appropriate to standard regression assumptions and more likely to yield a normal distribution of regression residuals. However, as illustrated in Appendix C-E,

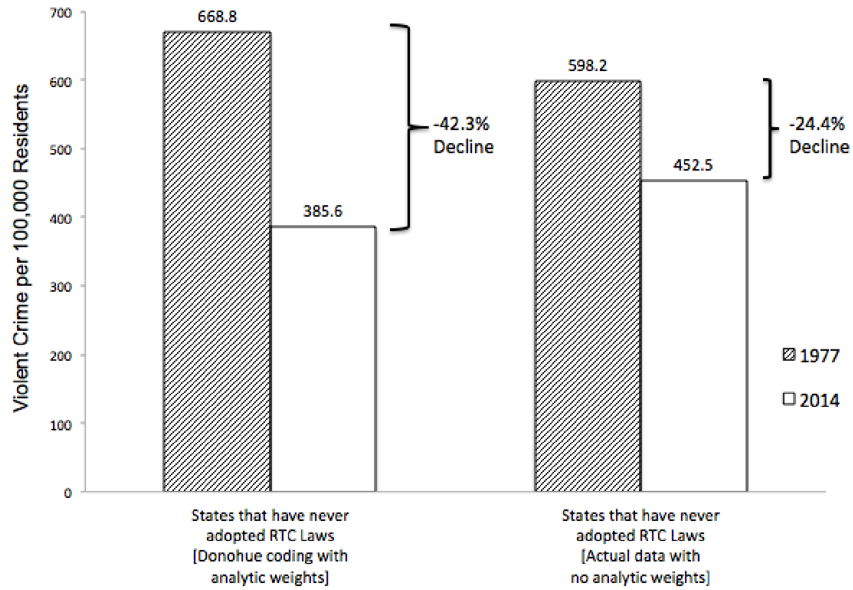


Figure 2. : Comparison of Violent Crime Rate Changes in No Carry States with and without Analytic Weights

which plot histograms for each type of crime rate across all states, crime rates are not highly skewed. Conducting a formal test of skewedness, 74.5% of states do not have skewed homicide data, 70.6% do not have skewed violent crime data, and 98.0% do not have skewed property crime data. Moreover, natural log transformations of this data produce almost as many states with non-normal distributions as were observed with the original rate data, as calculated using a joint test of skewness and kurtosis (27.5% vs 29.4% for homicide rate data, 17.6% vs 35.3% for violent crime rate data, and 29.4% vs 41.2% for property crime rate data). Finally, regressions employing rate data produce a normal distribution of residuals. In sum, there is no compelling technical need to employ natural log transformations.

From a theoretical perspective, there are reasons to analyze both log-transformed and standard rate data, as they provide different substantive insights. A log-linear model estimates coefficients that indicate approximately how a unit change in an independent variable relates to a percent change in the dependent variable. A log-log model estimates coefficients that indicate approximately how a percent change in an independent variable relates to a percent change in the dependent variable. And a linear-log model estimates coefficients that indicate approximately how a percent change in an independent variable relates to a unit change in the dependent variable. The relevance of both forms of analysis - examining percent changes in crime rates as well as unit changes - becomes clear when one examines the large disparity between crime rates in no-carry and carry states.

Figure 3 plots average violent crime rates by year in states with and without RTC laws. The number of sates that have RTC laws, and are thus included in the RTC state average, increases

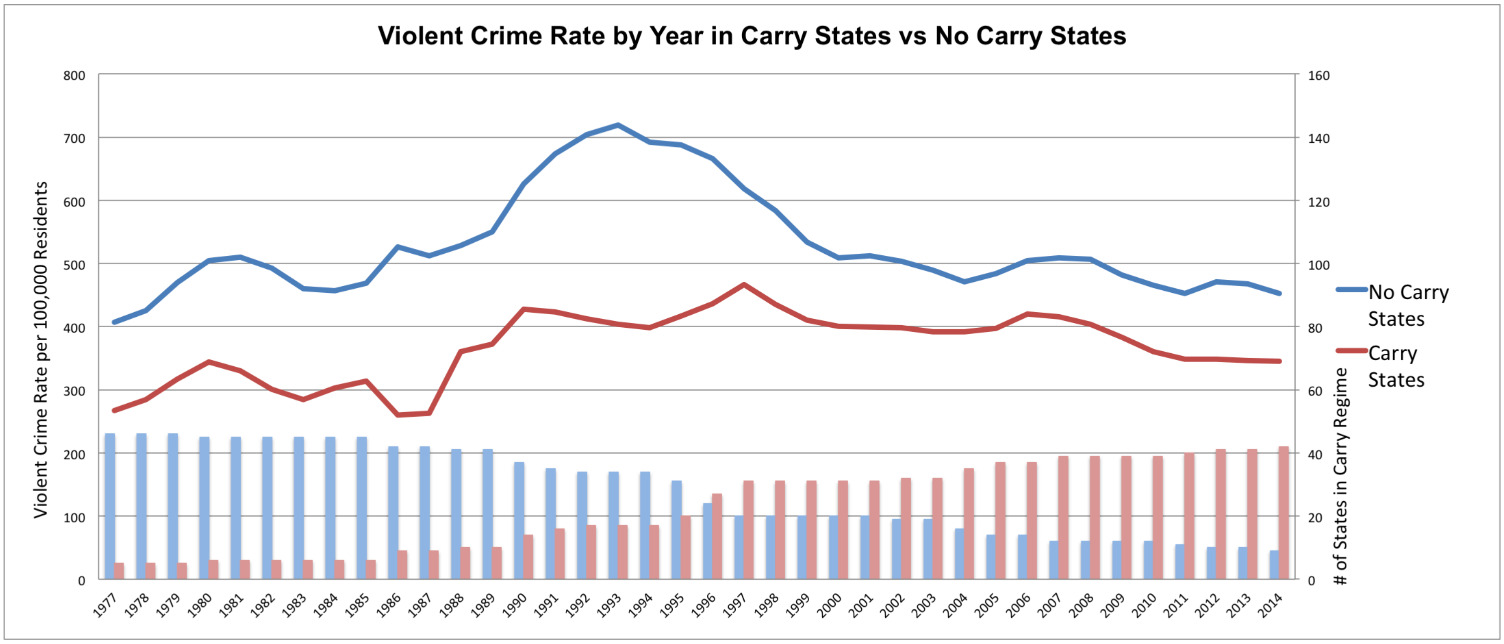


Figure 3. : Crime Rate Changes by Year in States with Different Carry Regimes

over time, as indicated in the lower bar graph indexed on the right y axis.

States without concealed carry have considerably higher violent crime rates in every year and, in some years, nearly double the rate of states with concealed carry. Note that this graph uses Donohue et al.'s coding of Massachusetts as a state with no RTC law, despite the high rate at which its residents hold carry permits, and coding it as a carry state would only exacerbate the disparity in violent crime rates between carry and no-carry states.

Because of the considerably higher rates of violent crime in states without RTC laws, the effect of unit changes in violent crimes rates will differ from percent changes, making both potentially worthwhile objects of inquiry. For example, while the populations of Washington DC and Vermont were nearly identical in size in 1994 (589,239 vs 583,836), Washington DC's violent crime rate of 2,662.6 per 100,000 residents was more than 27 times higher Vermont's rate of 96.9 per 100,000 residents. This means that if violent crime decreased by 50 incidents per 100,000 residents in both locations, this would constitute a 51.6% percent reduction in Vermont's rate but only a 1.9% reduction in DC's rate. Clearly, both unit rate changes and percent changes are worth investigating in this context and they may yield different rates of variance with implications for error modeling. Thus, we report results for both state-year crime rates and natural log transformed crime rates. As Table 1 and Table 2 (introduced below) indicate, estimated coefficients may change sign across these alternative specifications.

C. Challenge 3: Model overfitting and covariate selection

Overfitting is a concern in most empirical research. Functional forms and covariates that maximize the goodness of fit on a given data set often perform poorly on out of sample data. Moreover, as co-variates are added, the parameter space of a model increases exponentially as does the amount of data needed to make inferences of the same quality (De Marchi, 2005). Thus, there is a tradeoff between introducing additional covariates, at the risk of overfitting, and omitting variables that do influence a process, at the risk of generating omitted variable bias. Theoretical priors can assist with reasonable co-variate selection, but there is no simple test to determine which covariates to include.

Understandably, the question of which covariates to include has been a matter of ongoing debate in this literature, with different scholars settling on their "preferred" controls. In Table 2 of their paper, Donohue, Aneja and Weber (2019*a*) summarize their preferred controls versus Lott and Mustard. One key difference is that Donohue et al. use 6 age-sex-race demographic variables while Lott and Mustard use a more fine-grained assortment of 36 age-sex-race demographic variables. We agree with Donohue et al. that this approach of using a large number of highly co-linear demographic variables is not well supported theoretically and is problematic from a statistical perspective, as it is liable to introduce noise and generate spurious correlations. Donohue et al. also use per capita beer consumption, lagged per capita incarceration rate, and lagged police staffing rate, while Lott and Mustard use violent or property arrest rate[8] and state population. The two sets of authors also use slightly different measures of economic factors in addition to real per capita personal income (poverty rate and unemployment rate versus real per capita income maintenance, real per capita retirement payments, and real per capita unemployment insurance payments) and different measures of urbanization (percentage of state population living in metropolitan statistical areas vs population density).

In table 4 of their paper, Donohue et al. present coefficients meant to illustrate that their top line panel data estimates of the impact of RTC laws reveal a significant association with violent crime and property crime when their own preferred demographic controls are substituted in place of Lott and Mustard's (we follow the convention of Donohue, Aneja and Weber (2019*a*) and use the abbreviation DAW for their regressors and LM for Lott and Mustard (1997)'s when useful

[8]Donohue et al. raise concerns about the way in which arrest rates are calculated, while also noting that usable arrest data is absent in some state-years, and they make the decision to lag the arrest rates used in all of their LM regression models. Thus, it should be noted that they do not test the LM model as it was originally formulated. Also, for DC, arrest rate data is missing from 1997-2001, and there are concerns about the integrity of this data in other years. From 1978-1997, the average value in DC is 31.8 and in 1996 it is reported to be 29.16. However, the 2002 value is reported to be 0.67, with a 2001-2014 average of 1.27. No explanation is offered for this dramatic change, but it appears to correspond to a period in which DC's criminal justice system was overhauled and its prison population integrated into the federal Bureau of Prisons system. A number of criminal justice tracking statistics are not continuous through this period, and DC's incarceration data is missing from 2002-2014. In sum, there are strong reasons to doubt the integrity of the Violent/Property Crime Arrest Rate data in the Lott and Mustard model for DC, which suggests a drop of well over an order of magnitude between five years of missing data.

Table 2—: Alternative Panel Data Estimates with State- & Year-Fixed Effects, 1979–2014, Yield Different Results with Model Corrections

Panel A: LM Regressors Including 36 Demographic Variables					
	Murder Rate	Firearm Murder Rate	Nonfirearm Murder Rate	Violent Crime Rate	Property Crime Rate
	(1)	(2)	(3)	(4)	(5)
Initial Donohue Model	-5.17	-3.91	-5.70*	-1.38	-0.34
	(3.33)	(4.82)	(2.45)	(3.16)	(1.71)
No Analytic Weights	-2.69	-0.06	-5.22†	-5.39	0.25
	(2.85)	(3.97)	(2.72)	(4.22)	(1.96)
No Analytic Weights & No Ln Transformation	-0.77	1.69	-4.49	-77.55	1454.73
	(24.35)	(22.85)	(7.03)	(1655.37)	(7874.79)

Panel B: LM Regressors with 6 Donohue Demographic Variables					
	Murder Rate	Firearm Murder Rate	Nonfirearm Murder Rate	Violent Crime Rate	Property Crime Rate
	(1)	(2)	(3)	(4)	(5)
Initial Donohue Model	3.75	4.34	2.64	10.03*	7.59*
	(5.92)	(7.85)	(4.02)	(4.81)	(3.72)
No Analytic Weights	1.61	-1.11	5.64	1.96	6.56*
	(3.93)	(4.98)	(4.10)	(4.61)	(2.54)
No Analytic Weights & No Ln Transformation	-2.77	-15.16	13.04	2264.39	25191.33*
	(30.55)	(25.94)	(9.39)	(2289.93)	(9820.84)

Panel C: LM Regressors with 6 Donohue Demographic Variables, Adding Incarceration & Police Controls					
	Murder Rate	Firearm Murder Rate	Nonfirearm Murder Rate	Violent Crime Rate	Property Crime Rate
	(1)	(2)	(3)	(4)	(5)
Initial Donohue Model	4.99	5.96	3.76	10.05*	8.10*
	(5.50)	(7.20)	(4.29)	(4.54)	(3.62)
No Analytic Weights	1.44	-0.79	5.08	1.03	6.25*
	(4.14)	(5.21)	(4.27)	(4.70)	(2.67)
No Analytic Weights & No Ln Transformation	-41.58	-47.49	7.64	1108.32	20587.54*
	(36.03)	(31.13)	(10.22)	(2486.26)	(10106.59)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Robust clustered standard errors in parentheses. All models use Donohue et al. (2019)'s data, include year- and state-fixed effects, and employ clustered errors by state.

for brevity here and in our code). However, as we show in Table 2, the association with violent crime disappears once the models are corrected to remove analytic weights. Thus, both in their own model (see Table 1 from earlier), the Lott and Mustard model, and their modified Lott and Mustard models, panel data estimates of the effect of RTC laws from 1977-2014 show no significant association with homicide rates, firearm homicide rates, non-firearm homicide rates, or violent crime rates (except for the Lott and Mustard model, which suggests that RTC laws **lowered** the non-firearm homicide rate). While the modified Lott and Mustard models do show a positive association between RTC laws and property crime, that is consistent with the original thesis of Lott and Mustard (1997), namely that criminals should be expected to substitute towards forms of crime in which they are less likely to encounter an armed defendant. In sum, while scholars may have reasonable disagreements about what covariates ought to be included, and results that are robust to different sets of covariates should generally be given more credence, ***all panel models suggest that, when corrected to remove analytic weights, there is no significant relationship between RTC laws and murder rates or violent crime rates.***

These results provide strong evidence against concluding that RTC laws increase violent crime or homicides. However, in recent years Donohue has argued in favor of moving away from panel models in favor of synthetic control analysis, which he has suggested is "less sensitive to modeling choices (Donohue, 2018)." Indeed the second half of Donohue, Aneja and Weber (2019*a*) is devoted to presenting state-level synthetic control analysis, which the authors claim shows that "RTC laws are associated with 13–15 percent higher aggregate violent crime rates 10 years after adoption."

However, it is misleading to say that synthetic control analysis is less sensitive to modelling choices. As Moody and Marvell (2019) explain, "The fixed-effects regression model, properly specified, controls for all relevant factors, pre- and post-treatment, including trends and state and year fixed effects. This cannot be said for the synthetic control model." In particular, there is a real danger that the synthetic control approach excludes information needed to make a valid inference:

> Obviously, in the SC model there is the possibility that the control variables could vary significantly during the post-treatment period, altering the gap between the treated and control states; also states could have different trends with respect to the outcome; and the method fails to control for unobserved heterogeneity. The problem is that there is nothing held constant in the treatment period, so the gap is a function of the trends, the changing control variables, and the state fixed effects.

Moody and Marvell (2019) observe that this is a significant concern in the context of Donohue et al.'s analysis given that states vary widely in both their historical, path-dependent characteristics and in policies that they implement over the time periods being examined. In sum, an approach that fails to incorporate important factors that influence crime is unlikely to produce valid inferences.

This is a concern both with the inherent limits of synthetic control analysis but also with modelling choices that Donohue et al. make in their execution of this analysis.

Synthetic control analysis does, in fact, rely on modeling choices,[9] and Donohue et al. make modeling choices that overfit the correspondence between states and their synthetic controls in the pre-RTC period. The result is that states and their synthetic controls are more likely to diverge post-RTC because co-variates that are powerful predictors of crime rates are not allowed to exert predictive influence on the crime rate estimates generated for the synthetic control. Properly including covariates significantly changes the synthetic control analysis for many states, and a comprehensive synthetic control analysis with covariates that includes all 32 states in which RTC laws were introduced between 1986-2007 shows that RTC laws have had no significant effect on violent crime or homicides.

Donohue, Aneja and Weber (2019*a*) acknowledge in Appendix K that there are different approaches in the literature for incorporating lags of the dependent variable as predictors when constructing a synthetic control. They also acknowledge that they are aware of Kaul et al. (2017)'s recent paper, entitled, "Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates." Yet, Donohue et al proceed to use all pre-intervention outcomes together with covariates for most of their synthetic control analysis.

The key arguments of Kaul et al. (2017) are worth exploring in detail and at length in order to understand the shortcomings of Donohue, Aneja and Weber (2019*a*)'s approach. As Kaul et al. summarize in their abstract:

> It is becoming increasingly popular in applications of synthetic control methods to include the entire pre-treatment path of the outcome variable as economic predictors. We demonstrate both theoretically and empirically that using all outcome lags as separate predictors renders all other covariates irrelevant. This finding holds irrespective of how important these covariates are for accurately predicting post-treatment values of the outcome, potentially threatening the estimator's unbiasedness. We show that estimation results and corresponding policy conclusions can change considerably when the usage of outcome lags as predictors is restricted, resulting in other covariates obtaining positive weights.

Kaul et al. (2017) suggest that much of the nascent synthetic control literature has mistakenly done precisely what Donohue et al. have and used both lags of the dependent variable in all

[9]Donohue et al. do an admirable job of describing the synthetic control approach, which was first developed and applied by Abadie, Diamond and Hainmueller (2010) and Abadie, Diamond and Hainmueller (2015). Given the exposition provided in their paper, we do not reproduce an extensive introduction to the method here. However, for a comprehensive overview that emphasizes "feasibility, data requirements, contextual requirements, and methodological issues related to the empirical application of synthetic controls" and characterizes "the practical settings where synthetic controls may be useful and those where they may fail" see Abadie (2021).

pre-treatment periods as well as covariates, with the assumption that the inclusion of covariates will allow them to exert predictive influence on the construction of synthetic control outcomes. Inclusion of covariates that are believed to influence an outcome is indeed essential to the synthetic control methodology. As Kaul et al. explain:

> ...the prime objective of SCM is to build a synthetic control that properly reflects how the treated unit would have evolved after the intervention if the latter had not taken place. To achieve this goal, covariates with predictive power for the variable of interest should be matched, too. As stated by Abadie et al. (2015), "it is of crucial importance that synthetic controls closely reproduce the values that variables with a large predictive power on the outcome of interest take for the unit affected by the intervention." Thus, it is important that the explicitly chosen covariates are allowed to influence the estimated synthetic control.

However, including lags of the dependent variable in all pre-treatment periods actually leads to the covariates being entirely ignored. Kaul et al. prove both theoretically and empirically that

> ...using all pre-treatment values of the outcome variable as separate predictors inevitably leads to every single covariate being ignored...This finding holds no matter what the covariates actually are and how important and helpful these might be in order to predict post-treatment values of the outcome variable...When ignoring relevant covariates, the statistician's principle of using all available data is violated and synthetic controls are not applied as they are intended to be (Gardeazabal and Vega-Bayo, 2017): "the synthetic control is primarily designed to use any covariates that help explain the outcome variable as predictors, and not only pre-treatment values of the outcome variable".

Why were researchers initially drawn to the inclusion of lags for all pre-treatment periods? This was attractive because it enables synthetic controls to closely match the pre-treatment path of the treated unit. In essence, however, it gives rise to an extreme case of overfitting while ensuring that covariates will exert no predictive influence on the construction of the synthetic control. As Kaul et al. summarize: "using all pre-treatment outcomes as separate predictors leads to optimizing the pre-treatment fit of the outcome only, rendering all covariates irrelevant. The upside of this, i.e., the achievement of an optimal pre-treatment fit of the outcome, comes at the cost of ignoring the entire set of covariates, leading to a potentially biased estimator."

Donohue et al.'s rationale for using lags in all pre-treatment periods is that it generates the best fit in the pre-treatment period. Kaul et al. show why this reasoning is mistaken. Moreover, this strategy is likely to be particularly troublesome in the context of the trends witnessed near the passage of RTC laws, many of which were passed in the mid 1990's just as long trends of rising crime rates were reversed. In addition to overfitting the data on secular trends that are about to change, Donohue et al.'s approach also eliminates the ability of covariates that are known to affect crime to actually influence the construction of the synthetic control. In sum, the danger in this approach is that, as Kaul et al. explain, "solely optimizing the pre-treatment fit of the dependent variable and ignoring the covariates can be harmful: the more the covariates are truly influential for future values of the outcome, the larger a potential bias of the estimated treatment effect can become, possibly leading to wrong policy conclusions."

While Donohue et al. engage in extensive comparisons of different approaches to synthetic control modeling in their Appendix, this analysis appears to not adequately investigate the approach recommended by Kaul et al, namely using the last pre-treatment value of the dependent variable (a single lag of the most recent pre-treatment year) in addition to covariates. Note that the rationale for restricting a model to a single lag is that, as more lag years are included, covariates can exercise less influence on the outcome.

Donohue et al. compare their covariates to Lott and Mustard's covariates in the presence of lags for all years and lags in three years, the net effect of which is to entirely eliminate or partially eliminate the influence that covariates can actually exercise. The notes for tables K1 and K2 also suggest that Donohue et al. used the non-nested option for the synth program for their analysis of Lott and Mustard covariates, while using the more accurate nested option for analysis of their own covariates. However, in many cases, results differ between these two specifications when run on otherwise identical commands. The best procedure for comparing Lott and Mustard's covariates with Donohue et al.'s would be to run both with only one lag for the most recent pre-treatment year, and to use the most sensitive form analysis on each, e.g. nested with the "allopt" option.

Similar concerns can be raised with regard to table K9, in which Donohue et al. compare the influence of different lag specifications on their economic predictors (i.e. the influence of covariates). Their calculations for yearly lags appear at odds with Kaul et al.'s findings. However, this may be an artifact of computational limitations of the synth package. As Donohue et al. note, even when using the nested option, the synth package that they use to conduct synthetic control analysis can produce different estimates based on seemingly inconsequential details such as which version of software is running, specifications of the computer running the command, and the order in which predictors are listed. Moreover, as Klößner et al. (2018) observe in a recent article noting limits with current synthetic control approaches, some "results are not being reproduced when alternative software packages are used or when the variables' ordering within the dataset is changed."

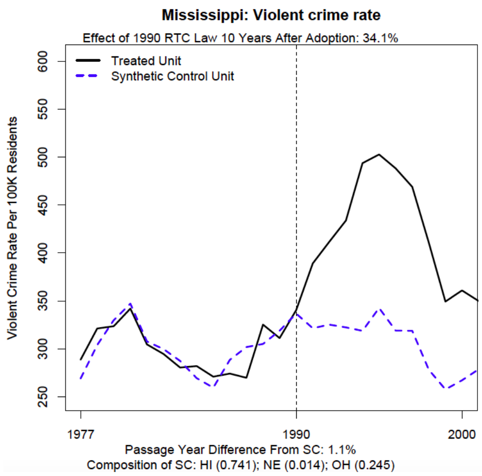
Indeed, we find that results sometimes differ between using the nested option alone and the more accurate and computationally demanding inclusion of the "allopt" option with the nested option. Also, in some cases, in order for the program to recognize and converge on zero coefficients as the most accurate estimates for covariates in the presence of lags in all years, user supplied custom V-weights must be entered (we demonstrate this in our publicly available code for the analysis of Mississippi). We suspect that this may account for the reason that in table K10 there is such little difference reported between the yearly lag specification and the 1 lag in the final year specification.

As best we can tell, Donohue et al. have not made the code for their synthetic control analysis publicly available, which makes it difficult to scrutinize their analysis in detail. However, we are able to reproduce their basic findings and graphs using lags in all pre-treatment periods.

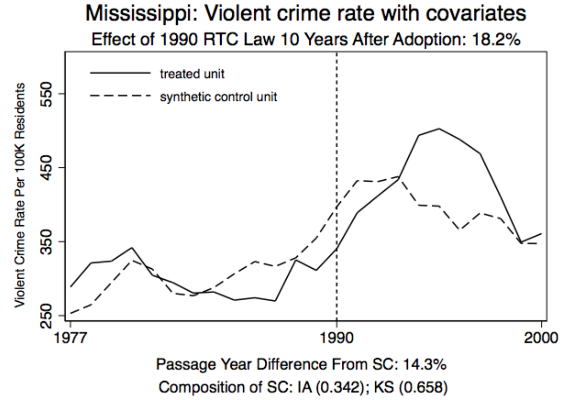
Critically, however, the approach recommended by Kaul et al. with one lag from the most recent pre-treatment period combined with covariates known to be influential in determining crime, yield results for many states that are dramatically different than Donohue et al.'s main estimates. For example, Figure 4 illustrates disparities in the synthetic control estimates for Mississippi, South Carolina, and North Carolina using Kaul et al.'s recommended approach. The estimate of the effect of the RTC regime on violent crime in Mississippi goes from an increase of 34.1% to an increase of 18%, in South Carolina it goes from an increase of 22.5% to an increase of only 1.1%, and in North Carolina it goes from an increase of 8.6% to a decrease of 2.6%. These are not small changes.

In deciding which covariates to include for this analysis we select those covariates that panel analysis reported in earlier tables suggests are most influential in determining crime. These are (with associated p-values from earlier panel analysis reported in parentheses): State Population (0.024), Real per Capita Personal Income (0.113), Real per Capita Income Maintenance (0.033), Population Density (0.00), Beer Consumption per Capita (0.00), Percentage of state population living in MSA (0.001), and 1-Year Lag of Number Incarcerated per 100,000 (0.019). For age-race demographics, which show up with different levels of predictive power based on how they are divided, we adopt a compromise between DAW and LM, which favors DAW's preference for parsimony, using 10-19 and 20-39 age groups for black males, white males, and other-race males. Also, although it is a powerful predictor in the Lott and Mustard model, we do not include the calculated 1-Year Lag of the Violent/Property Crime Arrest Rate assembled by Donohue et al. because of concerns expressed by Donohue et al. about its construction, the fact that is it missing in a number of state-years, and doubts about the accuracy of the data reported for DC. We refer to these 13 covariates as the "synthetic control" covariates.

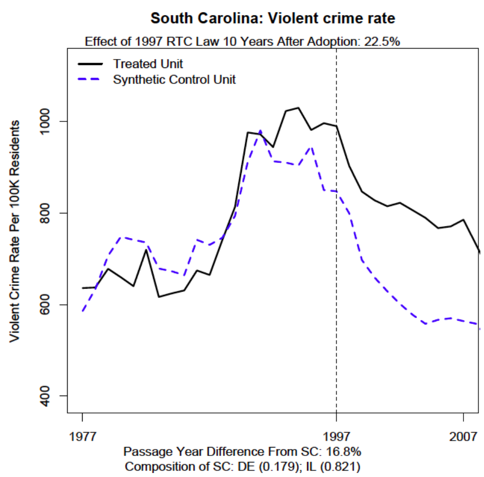
Finally, Donohue et al. mention in a footnote of their paper that they discard Washington, DC for their synthetic control analysis, despite including it in all previous panel analysis. Their rationale for doing so, explained briefly in the appendix, is that they argue DC is an "outlier." We find this argument unpersuasive, particularly given its inclusion in earlier analysis and the fact that DC has a larger population than Vermont and Wyoming, similar in size to Alaska, North Dakota, South Dakota, and Delaware. Although DC does have high levels of population density and violence,



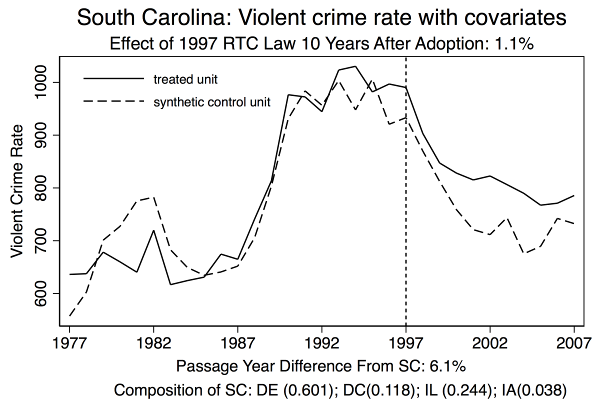
(a) MS: Donohue et al without DC or Covariates



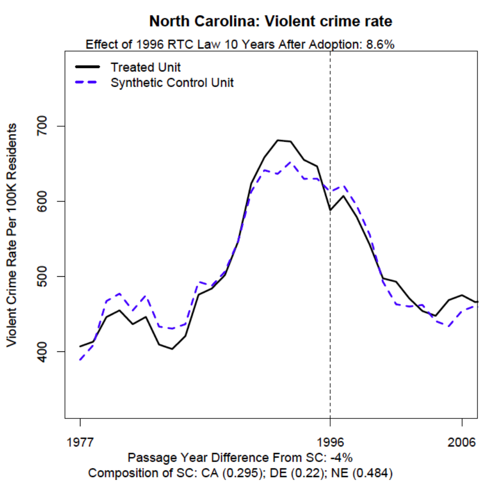
(b) MS: Corrected with DC and Covariates



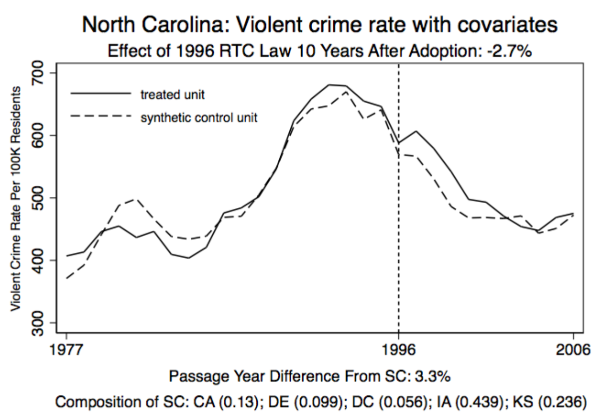
(c) SC: Donohue et al. without DC or Covariates



(d) SC: Corrected with DC and Covariates



(e) NC: Donohue et al without DC or Covariates



(f) NC: Corrected with DC and Covariates

Figure 4. : Inclusion of DC and covariates produces large changes in synthetic control estimates.

there are state-year units in which DC's violent crime rate is less than other states, such as Florida. Moreover, as stated above, we exclude the arrest rate variable that contains questionable DC data, so it is not a source of concern that would justify excluding DC. Donohue et al. assert that excluding DC from the analysis has little impact on their estimates of the effect of RTCs on violent crime. If true, they should not object to its inclusion, although we find many cases for which the inclusion of DC does impact synthetic control estimates. We include DC in our synthetic control analysis.

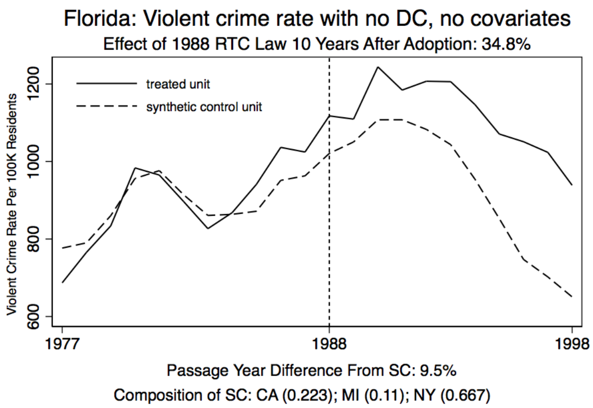
Florida is a useful state to examine to illustrate the impact of including covariates and DC in synthetic control analysis, as well as the effects of focusing analysis on different time periods. Figure 5 (a) displays Donohue et al.'s analysis for Florida, which suggests that 10 years after RTC passage, Florida's violent crime rate was 34.8% greater than the counterfactual synthetic control. Figure 5 (b) shows that this estimate decreases to 22.9% if DC is included in the same synthetic control analysis. Figure 5 (c) shows that this estimate jumps back up to 38.5% if covariates are included along with DC. However, Figure 5 (d) shows that if the time period for analysis is extended from 10 years to the full 26 years that this dataset allows us to investigate, the net effect is that Florida's violent crime rate is 8.5% *lower* than it would have been without an RTC law.

These examples demonstrate the sensitivity of synthetic control analysis of this data to a number of modeling choices, including the choice of whether to include covariates known to be highly correlated with crime, the choice of whether to use multiple lags, which decrease or eliminate the influence of covariates, the choice of whether to exclude some data, and the choice of whether to restrict analysis to particular time periods. However, we need to ask not only about the impact of these choices on the analysis of individual states, but also on the overall analysis of all states taken together.

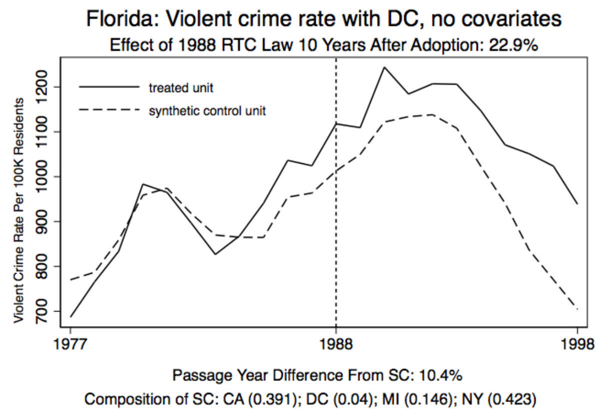
A recently released Stata package called "synth runner" automates the process of running multiple synthetic control estimations while allowing several units to receive treatments at different time periods (Galiani and Quistorff, 2017). The package also "conducts placebo estimates in space (estimations for the same treatment period but on all the control units)" and generates p-values by comparing the estimated main effect with the distribution of placebo effects. It is ideally suited to conduct comprehensive placebo control analysis of the effect of RTC laws on crime.

We use the synth runner package to analyze the 32 states in which RTC laws were introduced between 1986-2007.[10] This provides nine years for the pre-treatment period before the first RTC laws in this sample are adopted in 1986, and it provides seven years for the post-treatment period after the last RTC laws in this sample are adopted in 2007 (8 years of treatment data if one counts the first treated year). Restricting the analysis further one year in either direction (i.e. to 1987 or 2006) results in the loss of two states in each instance (Maine and North Dakota in 1986 and Kansas and Nebraska in 2007). As above, we include DC and covariates.

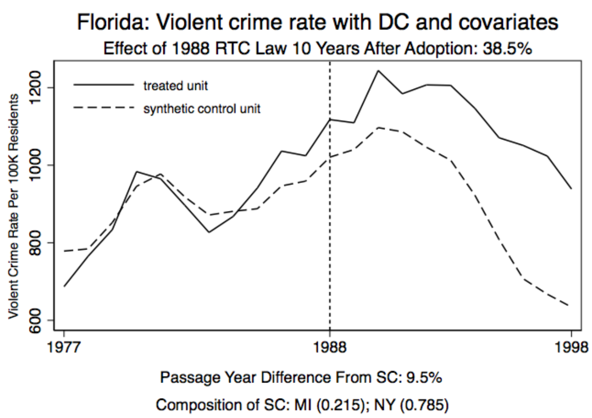
[10]Note that South Dakota, which is one of the 33 states that Donohue et al. examine individually in their synthetic control analysis, falls just before the cut off for this period and thus is not included because it would require dropping a year of pre-treatment data.



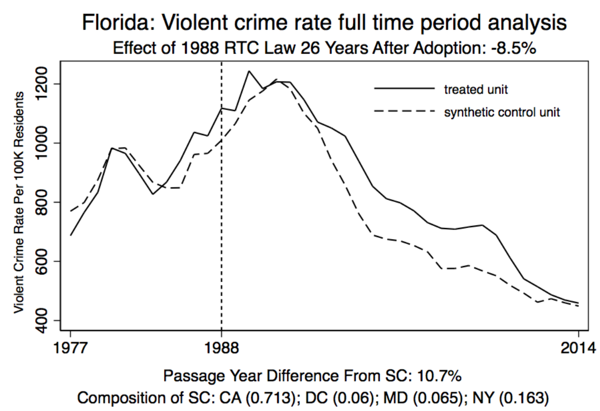
(a) Florida- Donohue et al., without DC or Covariates



(b) Florida with DC but no Covariates



(c) Florida with DC and Covariates



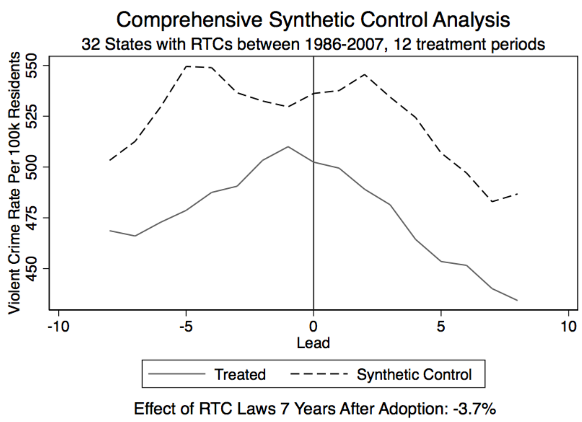
(d) Florida with DC, Covariates, and Full Time Period

Figure 5. : The inclusion of DC, the inclusion of covariates, and examination of additional time periods produce large changes in synthetic control estimates for Florida.

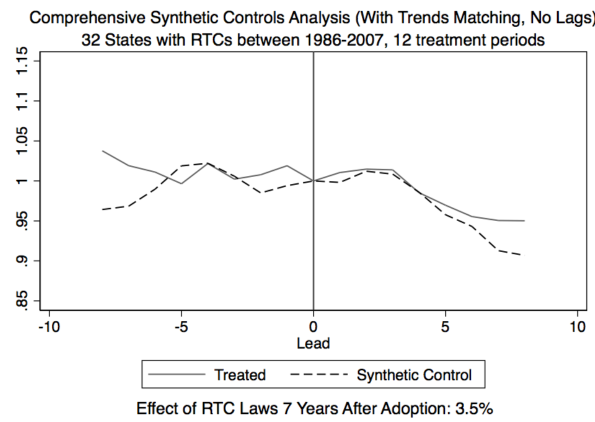
Figure 6 shows the results of this comprehensive synthetic control analysis by synth runner examining the effect of RTC laws on violent crime rates under two different specifications.[11] Version (a) shows the outcome with a single violent crime lag for the year before each treatment year, and version (b) shows the outcome without any violent crime lags but with trend matching (scaling each unit's outcome variable so that it is 1 in the last pre-treatment period). The first approach finds that RTC laws lowered violent crime rates by 3.7% by the 8th year of treatment compared to the synthetic control. Although standardized p-values suggest that this is a statistically significant difference, it should be noted that these are inflated because of the poor pre-treatment match quality, and the regular p-values are all greater than .72. In the trend matched version, violent crime rates are estimated to be 3.5% higher in RTC states in the 8th year of treatment compared

[11]Note that we use the regression based approach, as convergence is not achieved with the nested allopt option.

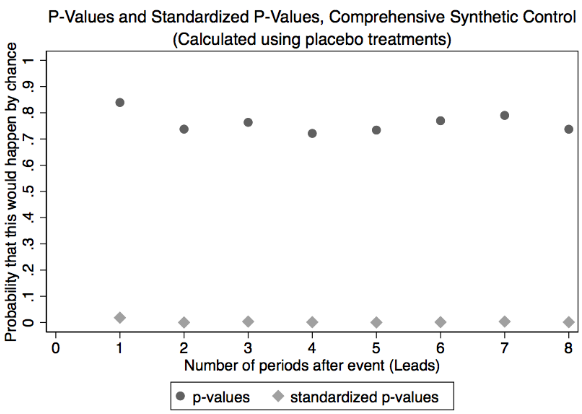
to the synthetic control but both p-values are above .2. In sum, comprehensive synthetic control analysis suggests that the differences in violent crime rates between RTC states and no-carry states are small and not statistically meaningful.



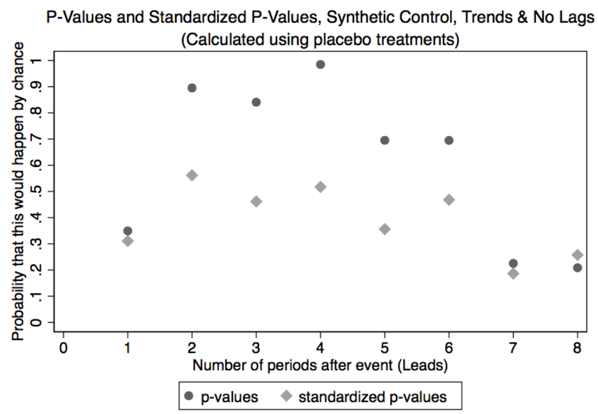
(a) Combined Synthetic Control Analysis



(b) Combined Synthetic Control Analysis, Trend Matching



(c) Placebo derived p-values



(d) Placebo derived p-values, trend matching

Figure 6. : Comprehensive synthetic control analysis of 32 states adopting RTC laws from 1986-2007 suggests small and insignificant effects on violent crime rates.

Given the relatively recent emergence of synthetic control methods, there will no doubt be further technical innovations with this approach in years to come along with additional parameters that can be tweaked and tests that can be performed. For example, the synth runner package provides tools for splitting pre-treatment periods into "training" and "validation" sections when pre-treatment periods are long and it employs sampling techniques that can be modified, as does the underlying synth package, as computational complexity increases. In any case, researchers will have to defend the modeling choices they make, particularly with regard to concerns about overfitting. Moreover, as Kaul et al. (2017) make clear, early adopters of this method may still have much to learn about its vulnerabilities and technical requirements.

We have shown that the synthetic control analysis by Donohue, Aneja and Weber (2019*a*) is not dispositive and that alternative, more theoretically and empirically defensible modeling choices yield very different results. Some states experience large shifts in the magnitude and direction of estimated effects; and comprehensive analysis suggests that RTC laws have no substantial or significant effect on violent crime rates. Ultimately, however, because synthetic control analysis focuses on a binary policy change, it remains vulnerable to concerns raised earlier about the proper way to code changes that track the underlying phenomenon of interest and the modeling of effects over time as the "treatment" intensity varies. This is yet another reason that examining the direct effect of the number of carry permits on crime over time constitutes a substantial contribution to this literature. The challenge is estimating data that is not reported in all states, which we address in the next section.

III. Data and Analysis of Carry Permit Rates Over Time

For the period that Donohue, Aneja and Weber (2019*a*) examine between 1977-2014, 5 states began this period with RTC laws in place, 37 states adopted RTC laws during this period, and 8 states plus DC were coded as not having RTC laws (note that this includes Massachusetts, which we discuss in more detail below). Of the 42 RTC states, 12 publicly report the number of permits issued for all years that their RTC law has been in effect: Florida, Virginia, Texas, Michigan, Colorado, Minnesota, Ohio, Kansas, Nebraska, Iowa, Wisconsin, and Illinois. Another 10 states have near perfect reporting of permit records for the last decade: Utah, Tennessee, Pennsylvania, Oregon, Oklahoma, North Carolina, Nevada, Kentucky, Indiana, and Arizona.

This leaves 20 states (plus Massachusetts, which makes 21) for which permit data is publicly reported more sporadically or not at all. In most cases, this data is tracked at the state or local level, but only released in response to inquiries by reporters or researchers. In 2012, in response to a request by Congress, the United States Government Accountability Office (GAO) issued a report that described the status of concealed carry permitting across U.S. states. Appendix V of that report indicates the number of valid permits by state, including numbers from many states that otherwise did not publicly report, as of December 31, 2011 (Office, 2012). More recently, John Lott has published a series of yearly reports under the auspices of the Crime Prevention Research Center that document the number of concealed carry permits issued by states, which includes numbers provided through direct inquiry to state and local officials.

Between 1977 and 2018, there are 1,067 state-years in which RTC laws were in effect (including Massachusetts, explained in more detail below). Using official state records, press reports citing state sources, the GAO report, and Crime Prevention Research Center reports, we are able to obtain concealed carry permit data for 408 state-years. The vast majority of this data comes directly from state sources, and more than 83% is not reliant on Crime Prevention Research Center reports. The entire data set, along with documentation of sources, is available as online material.

For each state-year with carry permit data, we divide the number of carry permits by the state population for that year to yield the "carry permit rate." Note that it is common in this literature to report carry rates using only the eligible adult population in a state as the denominator. These rates are obviously higher, in comparison, since the denominator is lower. While using the eligible adult population as the denominator does a better job of communicating the demand for carry permits among those who can possibly hold them, we calculate the carry permit rate based on overall state population in order to maintain uncontroversial continuity with the population data used by Donohue, Aneja and Weber (2019*a*).

Numbers for a few states warrant detailed explanation. Utah and Florida are widely known to issue a large number permits to non-residents, which many people seek because of reciprocity benefits that these permits can confer in other states. For both Utah and Florida, we use data for resident permits only, given that the number of non-resident permit holders is large and these individuals are less likely to be carrying on a regular basis within the state. Utah distinguishes between permits issued to residents and non-residents in reports from 2001-2004 and in 2013 and beyond. For years in which total permit numbers are not broken down between residents and non-residents, we multiply this total by the percentage breakdown from the nearest reported year in order to estimate resident permits for that given year (this applies to data for only three years). Florida reports total valid permits for every year dating back to 1988. More detailed data from the most recent four years shows that an average of 7.4% of licensees are from out of state. Thus, we deflate Florida's historical total valid permit data for each year by 7.4%.

As mentioned earlier, Donohue, Aneja and Weber (2019*a*) do not code Massachusetts as a RTC state, despite having a higher carry permit rate in recent years than about half of RTC states, and rates that are vastly higher than any other states coded as no-carry states. Moreover, during the period in question, the carry permit rate in Massachusetts was nearly identical to Connecticut's, which is a similar "may-issue" state that Donohue et al. do code as a RTC state. If one is interested in understanding the impact of the ability to carry on crime, it makes sense to take carry permit rates in Massachusetts into account.

The high carry permit rate in Massachusetts is an artifact of a major legislative change that took place in 1998, which essentially requires Massachusetts residents to obtain a "License to Carry" (LTC) in order to own a handgun. However, this license does allow recipients to carry a handgun in public for defense unless a specific restriction is placed on the license, which is a practice that is only common in a small number of local, urban police departments. Although an unconventional regime, Massachusetts has, in fact, licensed a large percentage of its population to carry since 1998.

We thus date the change to a carry regime in Massachusetts to 1998; and state reports provide concrete LTC data from 2010 to 2016.

Related to this question of how to code "may-issue" regimes, Iowa issued large numbers of permits as a may-issue state for many years before it transitioned to a shall-issue RTC regime in 2011. Iowa's RTC legislation was signed into law in April of 2010, and by the end of the year about 5.3% of the Iowa population was licensed to carry, even though the shall-issue criteria did not officially take effect until January 2011. Permit rates continued to rise to 7.7% by the end of 2011. Given the ability of individuals to submit applications in 2010 under the may-issue regime in anticipation of the changing standards, we code 2010 as the start of RTC in Iowa. This is a more accurate representation of the phenomena of interest and also helps ensure that statistical models are not biased by the dramatic shift that would be implied if rates went from 0 in 2010 directly to 7.7% in 2011.

Finally, it should be noted that in recent years a number of states have adopted so-called "constitutional carry" laws, which remove the need to obtain a permit in order to carry a handgun. While this may prove a challenge for research in future years that aims to quantify carry based on permit data as additional states drop permit requirements, it is not a substantial impediment to our current research. During the period we analyze between 1977 and 2014, only three states adopted constitutional carry: Alaska in 2003, Arizona in 2010, and Wyoming in 2011. Moreover, these states continued to grant permits to residents, which are valuable because they provide carry rights in other states through reciprocity agreements. This explains why, even after the passage of constitutional carry, carry permit rates continued to rise in these states. In Vermont, however, which many consider the original constitutional carry state, no permit data exists because the state has never issued permits. Given that Vermont's neighbor, New Hampshire, is similar in size and demographics and likewise had a longstanding RTC regime (dating to 1959), we use Lott's estimate for NH permits in 2014 in order to provide a single point estimate for the carry permit rate in Vermont in 2014.

Note that robustness tests contained in the appendix confirm that dropping Vermont, Massachusetts, Iowa, or DC has no substantive impact on the results.

In sum, our data on carry permit rates includes 12 states will full data, 10 states with extensive data, and 20 states with sparse data, but with at least one data point for each state, comprising 408 state-years out of 1,067. We pursue two strategies for estimating missing data. The first leverages the fact that carry permit rates in states with full or extensive data can be accurately modeled as a simple function of logistical growth over time. The second utilizes the statistical package Amelia II to impute missing data in cross section time series designs conditional on covariates. Although independent approaches, both yield similar estimates and analytic results.

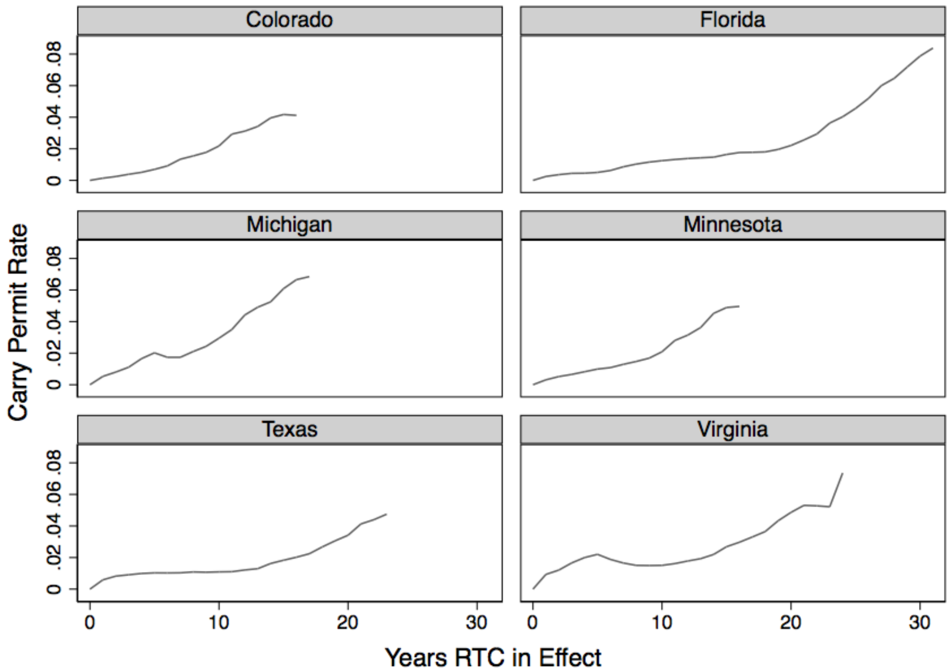


Figure 7. : Growth in Carry Permits Over Time (Six States with Longest Reporting Period)

Figure 7 plots the growth in total valid carry permits by year (after RTC passage) for the six states with longest, complete reporting records through 2018: Florida, Virginia, Texas, Colorado, Michigan, and Minnesota. The growth in permits follows a similar pattern across all states, consistent with standard logistic growth models. The initial number of permit seekers is low, but accelerates over time, and eventually must level out as the potential pool of permit holders is exhausted. Note if the mechanism through which RTC laws affect crime involves the actual ability of individuals to carry a firearm, this varies dramatically over time. Models that do not take into account the growth of carry permit rates will not be accurately assessing the impact of the ability to carry concealed on crime.

We first examine whether logistical growth curves do a good job of describing the growth in permit rates as a function of time. We fit a logistic curve to the data for these six states with the following functional form:

$$y = \frac{a}{1 + e^{-\frac{x-b}{c}}}$$

Where x indicates years after RTC passage, a represents the upper limit on permit rate growth, b represents the inflection point at which the permit rate reaches half of a, and c is a parameter scaling the growth rate. We further allow a and c to be state specific by including a random intercept U_j in the equation for a and c for the jth state:

$$a = a_j = \beta_1 + U_j$$
$$b = \beta_2$$
$$c = c_j = \beta_3 + U_j$$

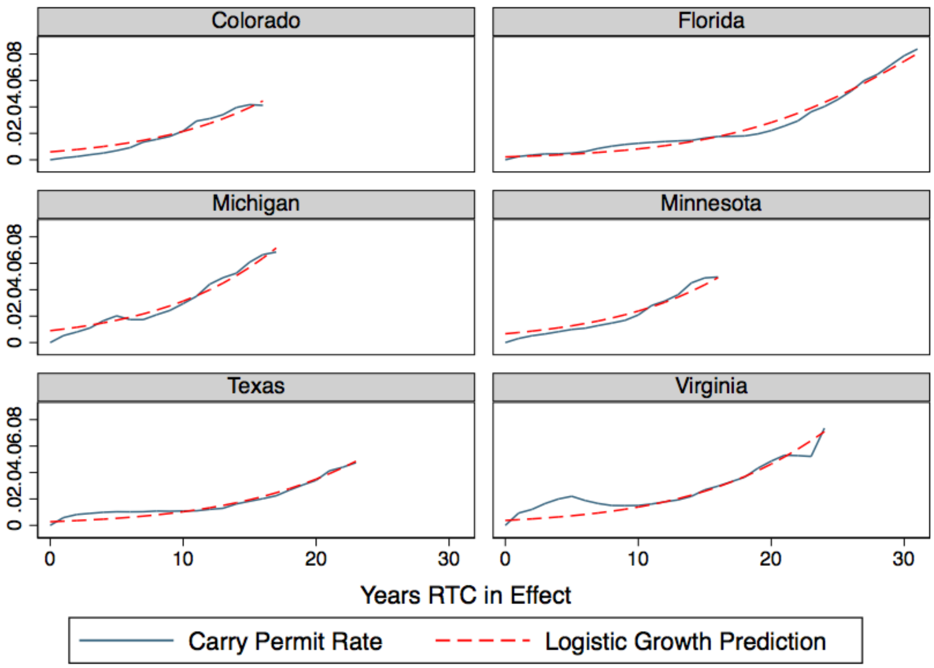


Figure 8. : Logistic Growth Model Approximates Carry Permit Rate Growth

Figure 8 plots carry permit rate data along with the rates predicted by the fitted logistic growth function. Visually, the logistic growth curves fit the data well, and goodness of fit statistics confirm this. The calculated R-squared is .952, and the estimated parameter coefficients have small standard errors (indicated in parentheses):β_1=.35 (.085), β_2=32.05 (2.12), β_3=7.54 (.33).

We proceed to fit the same logistic growth curve using data for all 408 state-years and use the results to generate predicted carry permit rate values. Figure 9 shows carry permit rate data along with the predicted values for 42 states. [12] The calculated R-squared is .953, and the estimated parameter coefficients again have small standard errors:β_1=.29 (.047), β_2=32.0 (1.43), β_3=8.08 (.28). Given the goodness-of-fit, we replace missing carry permit rates with the values predicted by the logistic model. Overall, these estimates suggest that, nationwide, there were about 569,000 permit holders in 1990, 2.5 million in 2000, 7.5 million in 2010, and 11.9 million in 2014.

[12]Note that, for purposes of visualization, display of Illinois is excluded because its RTC did not begin until 2014.

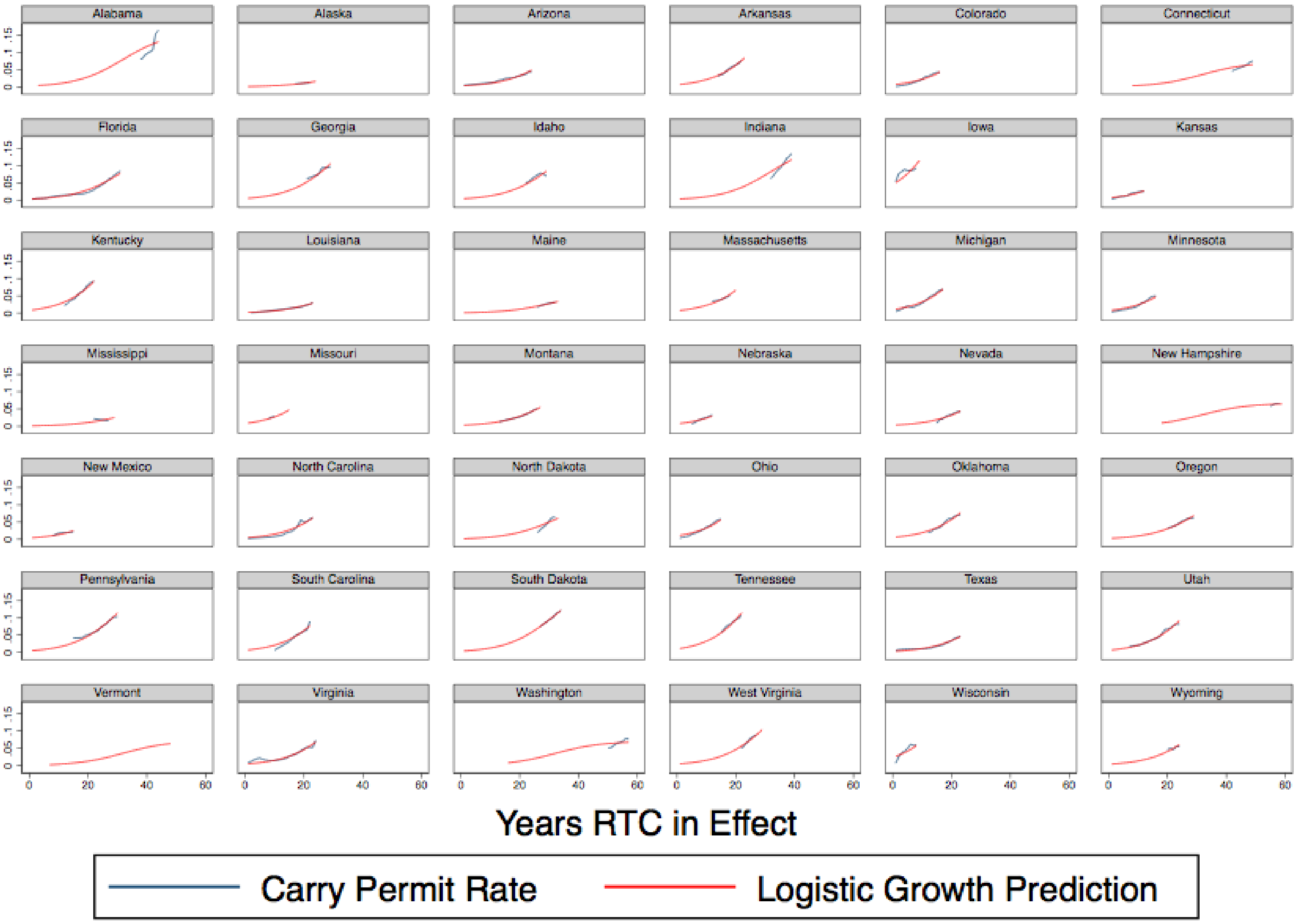


Figure 9. : Logistic Growth Model predictions

With the original data used by Donohue, Aneja and Weber (2019*a*) augmented by this carry permit rate variable, we analyze the effect of carry permit rates on crime using the same panel model employed by Donohue et al., estimating an equation of the general form:

$$y_{st} = \alpha_t + \delta_s + \beta X_{st} + \phi c_{st} + \epsilon_{st}$$

Where y is the rate of crime being investigated in state s and year t, α_t is a vector of dummy variables indicating year t, δ_s are state dummies for state fixed effects s, ϵ_{st} is the error term, the matrix X_{st} contains covariates or demographic controls for state s in year t, and c_{st} is the carry permit rate in state s at year t. ϕ is the coefficient reflecting the average estimated impact of one unit change in the carry permit rate on crime.[13] We also follow Donohue et al. and use robust standard errors clustered by state accompanying state and year fixed effects.

Table 3 reports the coefficients and standard errors for the estimated effect of carry permit rate on the five categories of crime. The regressions for first two rows include the covariates selected earlier for the synthetic control analysis, while the last two rows employ Donohue et al.'s covariates. For each set, we analyze the effect on the crime rate as well as on the natural log of the crime rate, the latter of which is indicative of the effect of a unit change in carry permit rate on the percent change in crime. As a robustness exercise, Appendix A contains additional tables showing results with Vermont, Iowa, Massachusetts, and DC excluded, all of which yield the same substantial results.

Table 3—: Panel Data Estimates using Carry Rates as Predicted by Logistic Growth with State- & Year-Fixed Effects & Alternative Covariates, 1977–2014

	(1) Murder Rate	(2) Firearm Murder Rate	(3) Nonfirearm Murder Rate	(4) Violent Crime Rate	(5) Property Crime Rate
Synthetic Control Covariates w/ Crime Rate DV	1.48 (12.08)	-0.56 (11.26)	1.44 (3.05)	298.47 (661.56)	4757.47† (2561.82)
Synthetic Control Covariates w/ Ln(Crime Rate) DV	-0.16 (1.20)	-0.58 (1.85)	-0.15 (1.16)	0.06 (1.32)	0.93 (0.76)
Donohue et al. Covariates w/ Crime Rate DV	1.45 (13.37)	-1.31 (12.11)	1.50 (3.38)	436.72 (751.13)	5067.77† (2788.59)
Donohue et al. Covariates w/ Ln(Crime Rate) DV	-0.05 (1.31)	-0.54 (2.00)	-0.15 (1.24)	0.36 (1.51)	1.04 (0.76)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Clustered standard errors in parentheses. All models include year- and state-fixed effects, and employ clustered errors by state.

This analysis reveals that carry permit rates have no significant association with crime. There is modest evidence that higher carry permit rates trend towards an association with property crime rates, consistent with the initial thesis of Lott and Mustard (1997). However, the standard errors for the coefficient estimates for homicide (murder) rate, firearms homicide rate, nonfirearm homicide rate, and violent crime rate are very large, and some coefficients change sign between different specifications. This analysis further confirms that, contrary to the suggestions made by Donohue et al., carry is not associated with a rise in violent crime rates.

[13]See Raffalovich and Chung (2015) a discussion of this common technique

B. Imputing Missing Data Using AMELIA

Although the logistic model for predicting missing carry permit rate data fits the overall growth of carry permit rates well, one concern is that it may not capture small variations from the overall growth trend that could be generated by local or national shocks. For example, comparatively more individuals might seek permits in in response to a spike in local crime or in response to a local or national election. As a robustness exercise, we take a second approach to estimating missing carry permit rate data that is better equipped to model such variations.

The Amelia (II) package developed by Honaker et al. (2011) is specifically formulated to aid in multiple imputation of missing data in cross-section, time-series models, using an algorithm that employs expectation-maximization with bootstrapping. The first step involves identifying all variables to include in the imputation model, which must include (at least) all variables used in the analysis model. Second, known parameters (structure of time series, ordinal/nominal variables, bounds etc.) can be specified.

While Amelia permits the logistic transformation of a variable, this is not a feasible approach for the carry permit rate variable, which takes the value of 0 in no-carry states and before RTC laws are enacted. Amelia does, however, allow us to specify 0 and 1 as bounds for the carry permit rate variable and choose specifications for patterns across time that can approximate logistic growth if indeed present. We specify a first-order polynomial of time (polytime = 1) and include splines of time with one knot (splinetime = 1). Year is set as the time series variable, while the number of years that an RTC has been in effect in a given state is specified as an ordinal variable. Given the relatively high degree of missingness, we follow the recommendation of Honaker et al. (2011) and add a ridge prior, which helps with numerical stability by "shrinking the covariances among the variables toward zero without changing the means or variances." We use their specific recommendation of approximately 1% of the number of observations (empri=20). However, our substantive results are robust to different specifications, including the elimination of a ridge prior altogether. We impute 10 datasets, which then form the basis for a comprehensive analysis.

Figure 10 plots the average imputed values for carry permit rates in addition to initial rate data across 42 states with RTC regimes (Illinois is again excluded for visualization because its RTC began in 2014 and the permit rate is known). The growth in carry permit rates resembles the logistical growth curves estimated above, but contains greater variation. Note that this presentation of average imputed values is done for visualization purposes only, and that each imputed data set is analyzed separately and the results averaged using the "mi estimate" Stata command, which is based on Schafer (1997).

As a diagnostic exercise to test the accuracy of the imputation algorithm, we can "overimpute" data that is not missing and compare predictions to the actual, known data. Figure 11 shows the

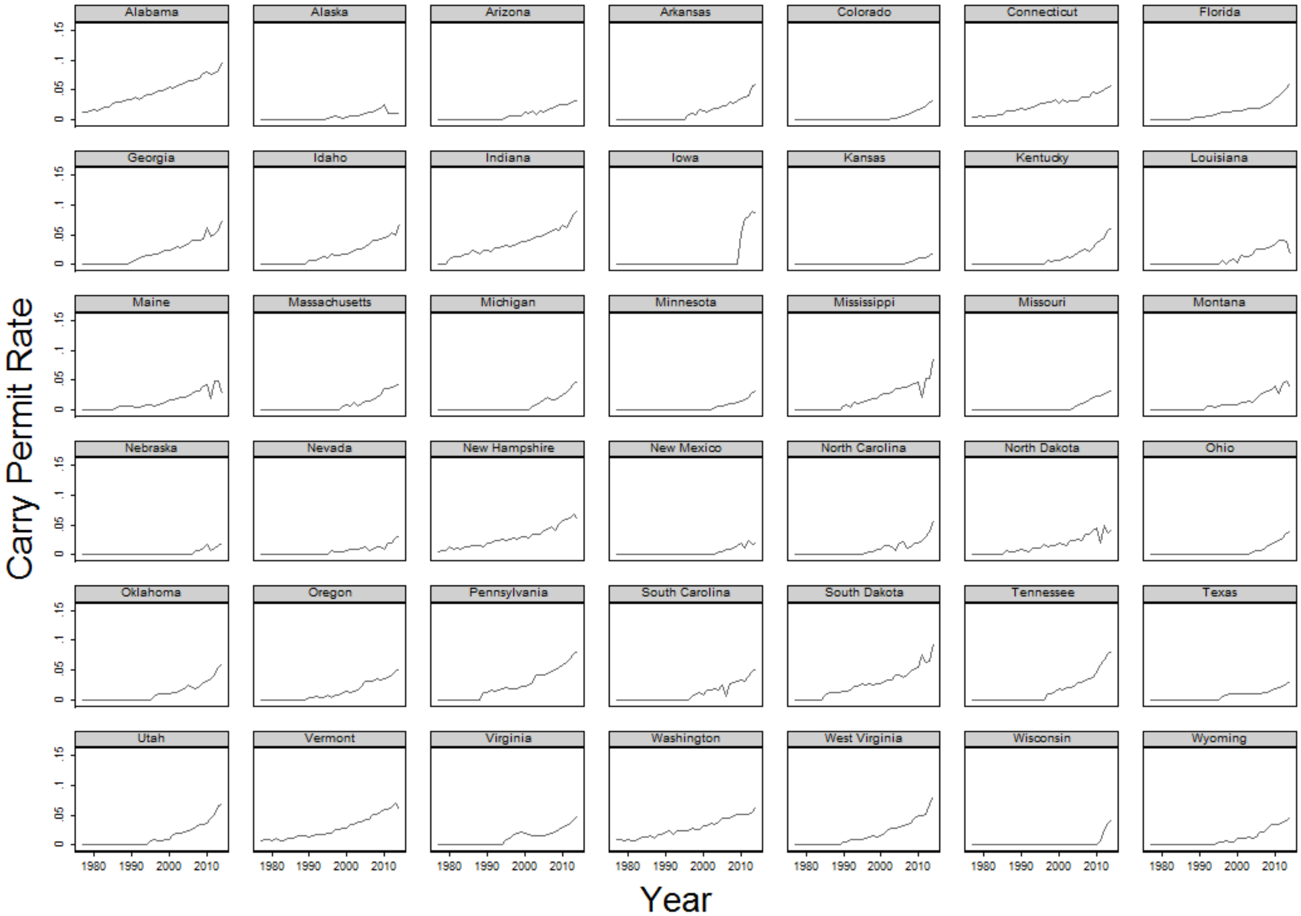


Figure 10. : Average Carry Permit Rate Imputed by Amelia

overimputation diagnostic graph, which plots ninety percent confidence intervals showing where an observed value would have been imputed had it been missing from the dataset, given the imputation model, with dot representing the mean imputation. Over 95% of these confidence intervals contain the y = x line, which means that the true observed value falls within this range. This suggests our imputation model has a high level of accuracy.

Using the 10 imputed datasets, we analyze and average their results using the mi estimate command with the same panel model as above with state and year fixed effects and robust standard errors clustered by state. As expected, the results displayed in Table 4 closely resemble the results reported in Table 4, and reveal no significant association between carry and violent crime rates, homicide rates, firearm homicide rates, non-firearm homicide rates or property crime rates. However, the pattern is again suggestive of Lott and Mustard (1997)'s original thesis, with property crime trending up, with relatively small standard errors, and violent crime and homicide rates trending down. As a robustness exercise, Appendix B contains additional tables showing results

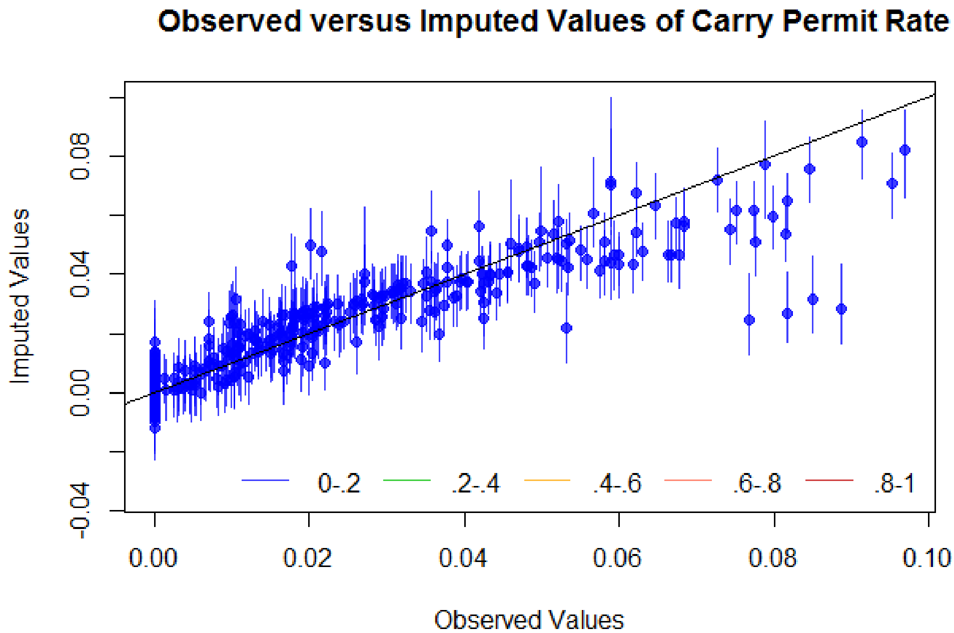


Figure 11. : Average Carry Permit Rate Imputed by Amelia

with Vermont, Iowa, Massachusetts, and DC excluded, all of which yield the same substantial results.[14]

In sum, the alternative method for estimating missing data using Amelia II, which demonstrates a high level of predictive accuracy using "overimpute" diagnostics, again reveals that rising carry permit rates are not associated with a rise in violent crime rates or homicide rates.

IV. Conclusion

Using data on the number of concealed carry permits issued each year by states with RTC regimes, along with two different methods for estimating missing data, we analyzed how the percent of a state's population licensed to carry affects crime rates, employing panel models and covariates widely used in this literature. This approach is theoretically more valid and econometrically more powerful than previous approaches that model a one-time regime change with a binary dummy variable. The results, which are consistent across both approaches, support the consensus view in the literature that more permissive carry has no significant relationship with violent crime rates or homicide rates.

Moreover, we show that the recent outlier result reported by Donohue, Aneja and Weber (2019*a*), which suggests that RTC laws have been associated with a significant rise in violent crime rates

[14]As an additional robustness test, we also examine a hybrid model that includes a dummy variable for RTC laws in addition to carry permit rates, which can better capture any immediate deterrent effects that may accompany an RTC law's passage in addition to the long-term effects of a growth in carry permits. The results are substantially the same.

Table 4—: Panel Data Estimates using Carry Rates as Imputed by Amelia with State- & Year-Fixed Effects & Alternative Covariates, 1977–2014

	(1) Murder Rate	(2) Firearm Murder Rate	(3) Nonfirearm Murder Rate	(4) Violent Crime Rate	(5) Property Crime Rate
Synthetic Control Covariates w/ Crime Rate DV	-7.63	-8.70	-0.85	-19.03	3730.65
	(9.35)	(8.28)	(2.67)	(618.14)	(2690.87)
Synthetic Control Covariates w/ Ln(Crime Rate) DV	-0.29	-0.31	-0.63	-0.18	0.87
	(1.31)	(2.00)	(1.18)	(1.23)	(0.73)
Donohue et al. Covariates w/ Crime Rate DV	-0.66	-2.19	-0.20	154.01	4772.65
	(10.43)	(9.11)	(2.60)	(640.09)	(2854.77)
Donohue et al. Covariates w/ Ln(Crime Rate) DV	0.02	0.07	-0.45	-0.11	1.05
	(1.38)	(2.11)	(1.18)	(1.26)	(0.73)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Clustered standard errors in parentheses. All models include year- and state-fixed effects, and employ clustered errors by state.

when examined using synthetic control analysis, is an artifact of poorly supported modeling choices. When corrected to allow covariates to influence the construction of predicted crime rates, synthetic control analysis likewise indicates that RTC laws have had no significant effect on violent crime rates.

While a null result that confirms the prior consensus in the literature may not always constitute a significant contribution, the public policy implications of this research are significant, and this study makes a timely and methodologically substantial contribution that should greatly increase confidence in the robustness of the main findings.

In addition to examining the phenomena of interest in a more direct and tractable manner, a strength of this analysis is that it examines an extended period over which a majority of states switched from restrictive, no-carry regimes to RTC regimes. Beyond this period, only a small number of states continue to operate as restrictive carry states (New York, California, Hawaii, New Jersey, Maryland, Rhode Island, and Delaware) which can limit their power to serve as representative counterfactuals going forward, particularly in a binary difference-in-difference framework.

Using estimates of actual carry rates can aid in extending this research in informative ways for future years. However, the recent proliferation of "constitutional carry laws," which eliminate the need to obtain a permit to carry, may complicate extensions of this analysis beyond the time period examined here. Thus, this analysis is particularly well suited for examining the time period in question, and the time period being examined is the most relevant for evaluating the overall impact of RTC laws.

While this study is informative for evaluating whether RTC laws that have enable higher rates of carry over time have affected crime rates, it could be valuable for future research to leverage carry permit data to investigate more detailed questions. For example, how do rates of carry affect forms of crime during rare periods of widespread social unrest or natural disasters, in which local police capacity is under stress, as compared with jurisdictions where carry is restricted? More generally, how do carry rates respond in a dynamic manner to spikes in local crime, or perceived spikes as measured through media stories or internet searches, and what are the short-term and long-term implications, if any? Also, additional research that examines the effects of carry rates over extended time periods across smaller geographical regions, such as counties, could constitute a significant contribution, although such data may only be obtainable in certain states and require extensive data gathering efforts. Similarly, permit rate data might be leveraged to examine associations with defensive uses of firearms, although again gathering sufficient high quality data may be challenging. Finally, it could be informative to examine how different sorts of requirements for obtaining a permit (fees, training, renewal intervals, etc.) affect permit issuance, the types of people who are able to obtain permits, and crime. In sum, carry permit data may be leveraged to study a number of questions regarding the effects of carry in greater detail in future research.

Overall, the results of this study are encouraging from a public policy perspective. The widespread adoption of RTC laws coupled with sustained growth in the portion of the population seeking carry permits has led to extensive speculation regarding the potential cost or benefits to crime. While the existing consensus in the literature is that permissive carry has no significant effect on crime, this research has been plagued by methodological challenges. The recent claim by Donohue, Aneja and Weber (2019*a*) to have discovered a significant association with violent crime using synthetic control analysis understandably generated concern. However, as we have shown, this result is an artifact of indefensible modeling choices, which, once corrected, indicate that synthetic control analysis likewise finds that RTC regimes have not raised crime rates. Moreover, we have shown that, using the much more detailed measure of carry permit rates by year by state, increasing rates of carry have no significant relationship with violent crime rates or homicide rates. This study thus provides further, strong evidence that the dramatic growth in the ability to carry firearms for self-defense in recent decades has not exacerbated crime.

REFERENCES

Abadie, Alberto. 2021. "Using synthetic controls: Feasibility, data requirements, and methodological aspects." *Journal of Economic Literature*, 59(2): 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." *Journal of the American statistical Association*, 105(490): 493–505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative politics and the synthetic control method." *American Journal of Political Science*, 59(2): 495–510.

Aneja, Abhay, John J Donohue III, and Alexandria Zhang. 2011. "The impact of right-to-carry laws and the NRC report: lessons for the empirical evaluation of law and policy." *American Law and Economics Review*, 13(2): 565–631.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly journal of economics*, 119(1): 249–275.

Council, National Research, et al. 2005. "Firearms and violence: a critical review."

Cramer, Clayton E, and David B Kopel. 1994. "Shall issue: The new wave of concealed handgun permit laws." *Tenn. L. Rev.*, 62: 679.

De Marchi, Scott. 2005. *Computational and mathematical modeling in the social sciences.* Cambridge University Press.

Donohue, John J. 2018. "More gun carrying, more violent crime." *Econ Journal Watch*, 15(1): 67.

Donohue, John J, Abhay Aneja, and Kyle D Weber. 2017*a*. "Right-to-Carry Laws and Violent Crime: A Comprehensive Assessment Using Panel Data and a State-Level Synthetic Control Analysis." National Bureau of Economic Research Working Paper 23510.

Donohue, John J, Abhay Aneja, and Kyle D Weber. 2019*a*. "Right-to-carry laws and violent crime: a comprehensive assessment using panel data and a state-level synthetic control analysis." *Journal of Empirical Legal Studies*, 16(2): 198–247.

Donohue, John J., Abhay Aneja, and Kyle D. Weber. 2019*b*. "RTC Laws Increase Violent Crime: Moody and Marvell Have Missed the Target." *Econ Journal Watch*, 16(1): 97.

Donohue, John J, Abhay Aneja, and Kyle Weber. 2017*b*. "Right-to-carry laws and violent crime: A comprehensive assessment using panel data and a state-level synthetic controls analysis."

Donohue, John J, and Ian Ayres. 2003. "Shooting down the more guns, less crime hypothesis."

Durlauf, Steven N, Salvador Navarro, and David A Rivers. 2016. "Model uncertainty and the effect of shall-issue right-to-carry laws on crime." *European Economic Review*, 81: 32–67.

Galiani, Sebastian, and Brian Quistorff. 2017. "The synth_runner package: Utilities to automate synthetic control estimation using synth." *The Stata Journal*, 17(4): 834–849.

Goodman-Bacon, Andrew. 2018. "Difference-in-differences with variation in treatment timing." National Bureau of Economic Research.

Gresenz, Carole Roan. 2018/2020. "Effects of Concealed-Carry Laws on Violent Crime." *RAND Corporation (first published in 2018, updated in 2020).*

Hamill, Mark E, Matthew C Hernandez, Kent R Bailey, Martin D Zielinski, Miguel A Matos, and Henry J Schiller. 2019. "State level firearm concealed-carry legislation and rates of homicide and other violent crime." *Journal of the American College of Surgeons*, 228(1): 1–8.

Helland, Eric, and Alexander Tabarrok. 2004. "Using placebo laws to test" more guns, less crime"." *Advances in Economic Analysis & Policy*, 4(1): 1182.

Honaker, James, Gary King, Matthew Blackwell, et al. 2011. "Amelia II: A program for missing data." *Journal of statistical software*, 45(7): 1–47.

Kaul, Ashok, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler. 2017. "Synthetic control methods: Never use all pre-intervention outcomes together with covariates."

Klößner, Stefan, Ashok Kaul, Gregor Pfeifer, and Manuel Schieler. 2018. "Comparative politics and the synthetic control method revisited: A note on Abadie et al.(2015)." *Swiss journal of economics and statistics*, 154(1): 1–11.

Lott, John R. 2011. "What a balancing test will show for right-to-carry laws." *Md. L. Rev.*, 71: 1205.

Lott, Jr, John R, and David B Mustard. 1997. "Crime, deterrence, and right-to-carry concealed handguns." *The Journal of Legal Studies*, 26(1): 1–68.

Manski, Charles F, and John V Pepper. 2018. "How do right-to-carry laws affect crime rates? Coping with ambiguity using bounded-variation assumptions." *Review of Economics and Statistics*, 100(2): 232–244.

Moody, Carlisle E, and Thomas B Marvell. 2018. "The impact of right-to-carry laws: A critique of the 2014 version of Aneja, Donohue, and Zhang." *Econ Journal Watch*, 15(1): 51.

Moody, Carlisle E, and Thomas B Marvell. 2019. "Do right to carry laws increase violent crime? a comment on Donohue, Aneja, and Weber." *Econ Journal Watch*, 16(1): 84–96.

Moody, Carlisle E, and Thomas B Marvell. 2020. "Clustering and standard error bias in fixed effects panel data regressions." *Journal of Quantitative Criminology*, 36(2): 347–369.

Morrall, Andrew. 2018. "The science of gun policy: a critical synthesis of research evidence on the effects of gun policies in the United States." *Rand health quarterly*, 8(1).

Office, US Government Accountability. 2012. "Gun Control-States' Laws and Requirements for Concealed Carry Permits Vary Across the Nation."

Plassmann, Florenz, and John Whitley. 2003. "Confirming" more guns, less crime"." *Stanford Law Review*, 1313–1369.

Raffalovich, Lawrence E, and Rakkoo Chung. 2015. "Models for pooled time-series cross-section data." *International Journal of Conflict and Violence (IJCV)*, 8(2): 209–221.

Schafer, Joseph L. 1997. *Analysis of incomplete multivariate data.* CRC press.

Schell, Terry. 2018. "Methodological Challenges to Identifying the Effects of Gun Policies." *Research Review Essay.*

Siegel, Michael, Ziming Xuan, Craig S Ross, Sandro Galea, Bindu Kalesan, Eric Fleegler, and Kristin A Goss. 2017. "Easiness of legal access to concealed firearm permits and homicide rates in the United States." *American journal of public health*, 107(12): 1923–1929.

Wolfers, Justin. 2006. "Did unilateral divorce laws raise divorce rates? A reconciliation and new results." *American Economic Review*, 96(5): 1802–1820.

Table A1—: Logistic Growth Model, Panel Data Estimates with State- & Year-Fixed Effects, Synthetic Control Covariates, 1979–2014, Robust to Dropping Contested States

	Murder Rate	Firearm Murder Rate	Nonfirearm Murder Rate	Violent Crime Rate	Property Crime Rate
DC Dropped	(1)	(2)	(3)	(4)	(5)
DV=Crime Rate	3.69	1.23	0.56	282.91	4365.47
	(7.46)	(6.30)	(2.84)	(658.71)	(2745.54)
DV=Ln(Crime Rate)	0.04	-0.01	-0.54	-0.23	0.82
	(1.14)	(1.61)	(1.16)	(1.41)	(0.79)
VT Dropped					
DV=Crime Rate	1.39	-0.42	1.41	328.67	5274.43*
	(12.30)	(11.35)	(3.07)	(671.29)	(2543.34)
DV=Ln(Crime Rate)	-0.09	-0.38	-0.01	0.24	1.09
	(1.22)	(1.85)	(1.16)	(1.32)	(0.76)
IA Dropped					
DV=Crime Rate	1.56	-0.80	1.56	332.56	5157.16†
	(13.29)	(12.50)	(3.32)	(724.18)	(2741.74)
DV=Ln(Crime Rate)	-0.29	-0.92	-0.17	0.21	1.04
	(1.31)	(2.01)	(1.26)	(1.43)	(0.82)
MA Dropped					
DV=Crime Rate	1.12	-1.22	1.52	306.12	4769.56†
	(12.08)	(11.28)	(3.03)	(664.95)	(2570.55)
DV=Ln(Crime Rate)	-0.16	-0.66	-0.06	0.09	0.94
	(1.20)	(1.85)	(1.15)	(1.32)	(0.77)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Robust clustered standard errors in parentheses. All models use Donohue et al. (2019)'s data, include year- and state-fixed effects, and employ clustered errors by state.

Table B1—: Amelia Model, Panel Data Estimates with State- & Year-Fixed Effects, Synthetic Control Covariates, 1979–2014, Robust to Dropping Contested States

	Murder Rate	Firearm Murder Rate	Nonfirearm Murder Rate	Violent Crime Rate	Property Crime Rate
DC Dropped	(1)	(2)	(3)	(4)	(5)
DV=Crime Rate	2.96	2.48	-1.23	14.82	3200.90
	(7.77)	(6.09)	(2.70)	(645.42)	(2959.26)
DV=Ln(Crime Rate)	0.18	0.83	-1.20	-0.66	0.68
	(1.23)	(1.72)	(1.22)	(1.29)	(0.77)
VT Dropped					
DV=Crime Rate	-7.94	-8.81	-0.89	-10.00	4253.13
	(9.69)	(8.58)	(2.74)	(629.07)	(2684.86)
DV=Ln(Crime Rate)	-0.24	-0.20	-0.56	-0.07	1.02
	(1.32)	(2.04)	(1.24)	(1.29)	(0.73)
IA Dropped					
DV=Crime Rate	-8.44	-9.70	-0.97	-23.67	4063.57
	(10.42)	(9.20)	(2.94)	(688.37)	(2942.80)
DV=Ln(Crime Rate)	-0.45	-0.63	-0.71	-0.06	0.98
	(1.44)	(2.20)	(1.29)	(1.34)	(0.80)
MA Dropped					
DV=Crime Rate	-8.13	-9.27	-0.78	-8.55	3723.31
	(9.42)	(8.36)	(2.66)	(618.92)	(2697.87)
DV=Ln(Crime Rate)	-0.30	-0.39	-0.54	-0.15	0.87
	(1.31)	(2.01)	(1.19)	(1.23)	(0.73)

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Robust clustered standard errors in parentheses. All models use Donohue et al. (2019)'s data, include year- and state-fixed effects, and employ clustered errors by state.

Appendix C

Figure C1 plots histograms of homicide rates by state, illustrating that this data is not highly skewed. Conducting a formal test of skewedness, 74.5% of states do not have skewed homicide data. Moreover, natural log transformation of homicide data yields non-normal distributions in 27.5% of states, compared with 29.4% of states when examining raw rate data (as calculated using a joint test of skewness and kurtosis).

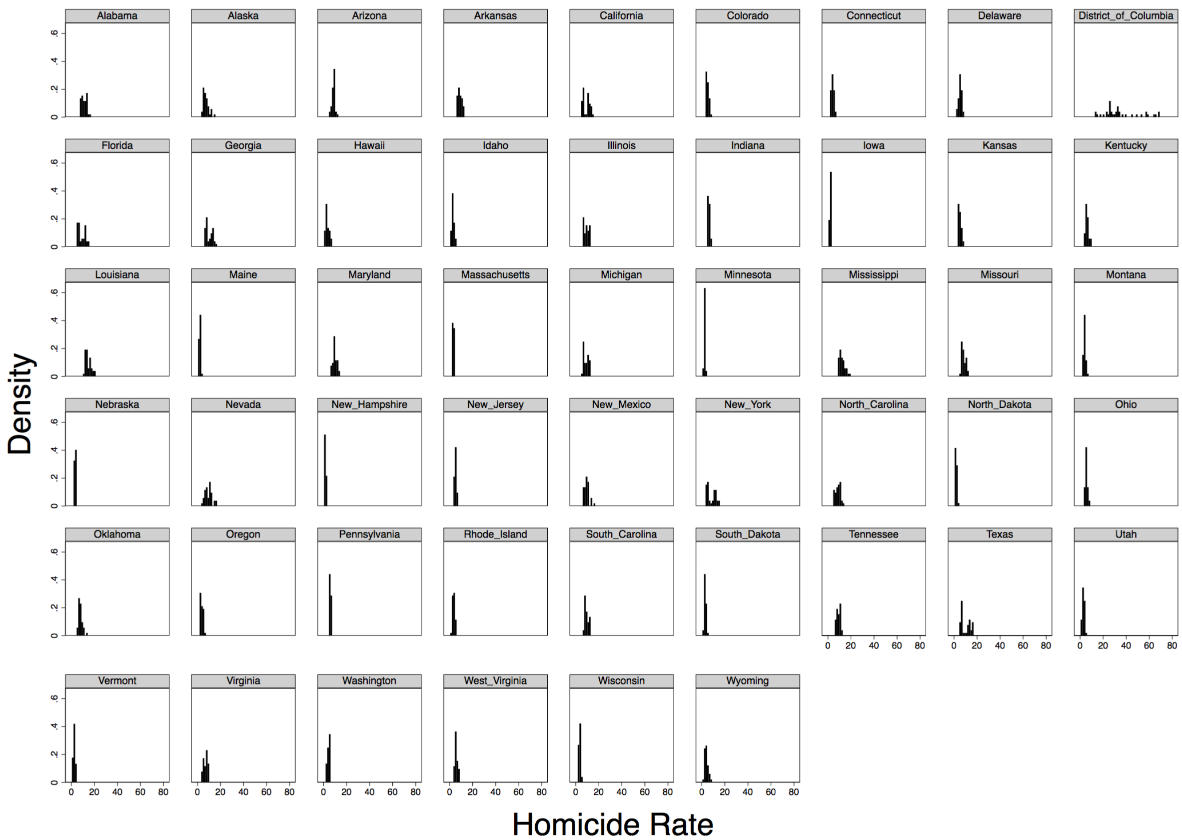


Figure C1. : Homicide Rate Histograms by State

Appendix D

Figure D1 plots histograms of violent crime rates by state, illustrating that this data is not highly skewed. Conducting a formal test of skewedness, 70.6% of states do not have skewed violent crime rate data. Moreover, natural log transformation of violent crime rate data yields non-normal distributions in 17.6% of states, compared with 35.3% of states when examining raw rate data (as calculated using a joint test of skewness and kurtosis).

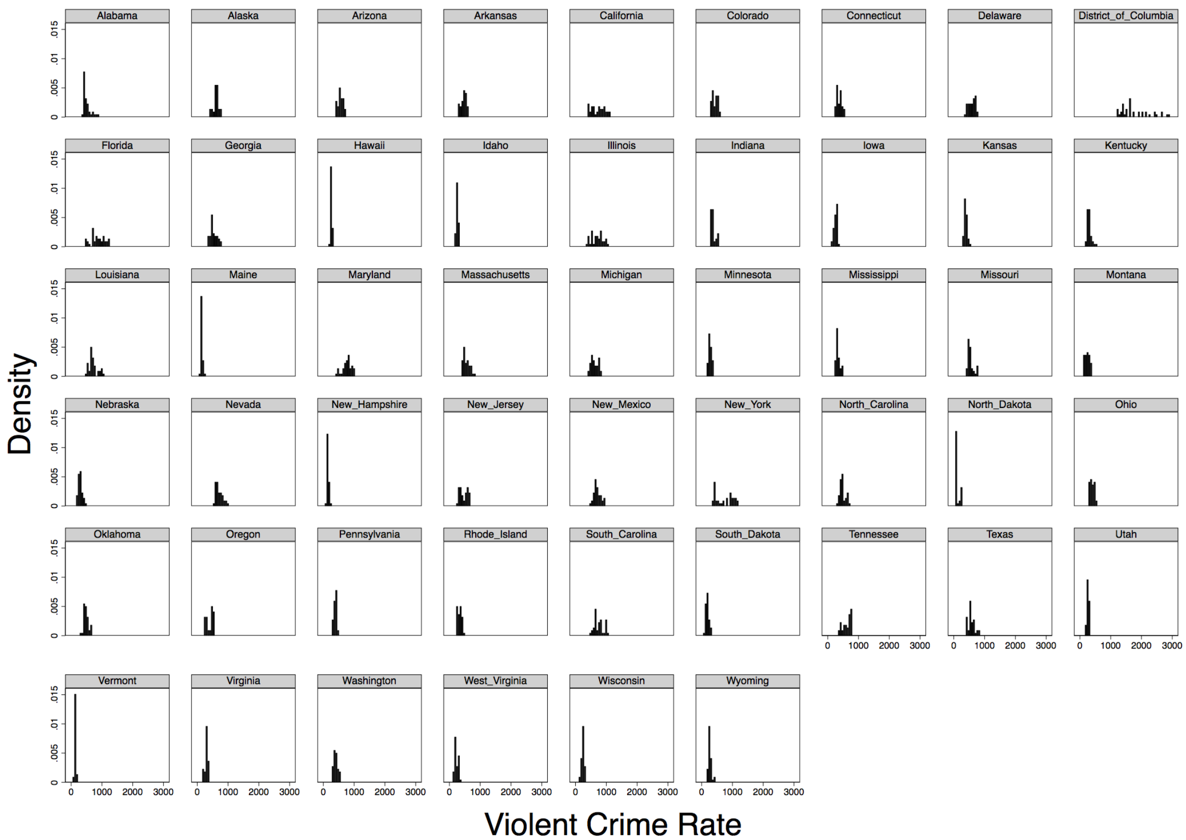


Figure D1. : Violent Crime Rate Histograms by State

Appendix E

Figure E1 plots histograms of property crime rates by state, illustrating that this data is not highly skewed. Conducting a formal test of skewedness, 98.0% of states do not have skewed property crime rate data. Moreover, natural log transformation of property crime rate data yields non-normal distributions in 29.4% of states, compared with 41.2% of states when examining raw rate data (as calculated using a joint test of skewness and kurtosis).

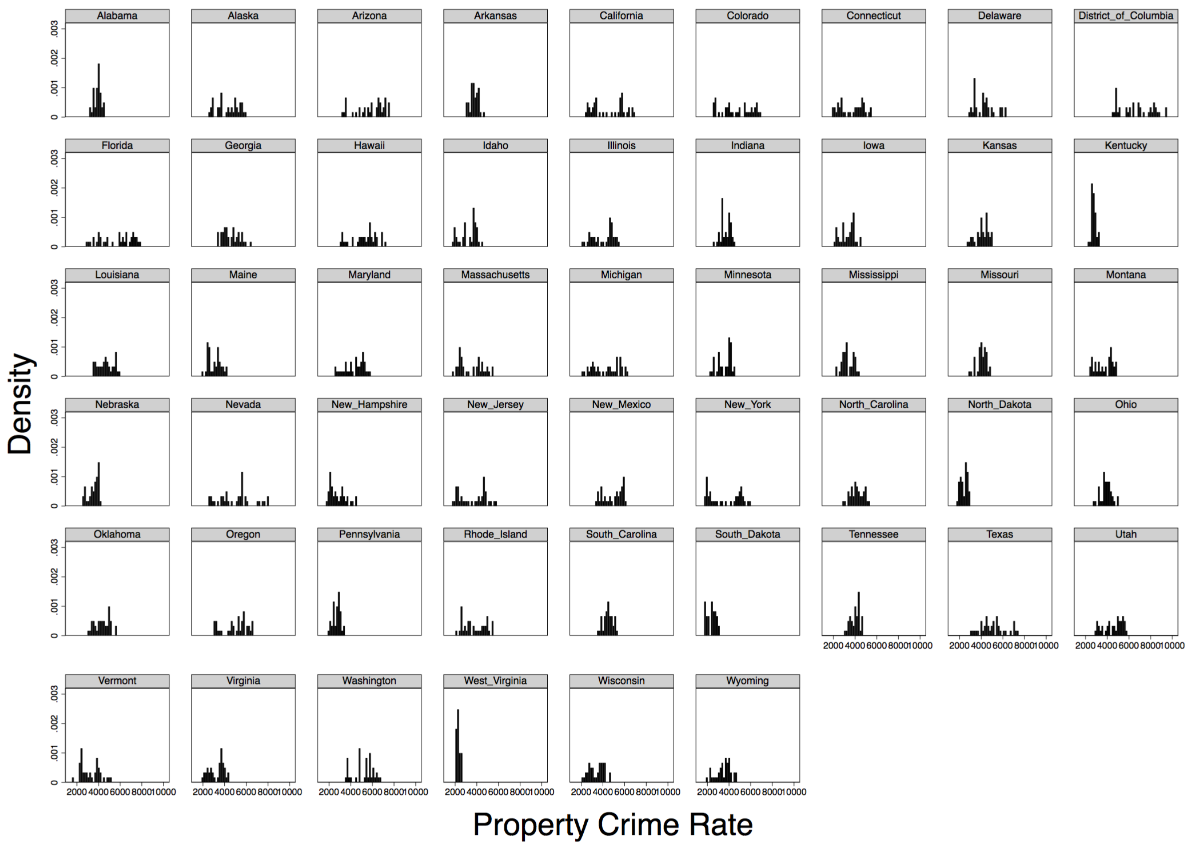


Figure E1. : Property Crime Rate Histograms by State