

Lott and Mustard Replication

Abby Johnson

5/6/2022

1 Introduction

This paper explores the Lott and Mustard 1997 study titled “Crime, Deterrence, and Right-to-Carry Concealed Handguns”¹. Specifically, this paper aims to replicate the 1997 analysis using state-level data and extend the causal inference methods using more modern techniques. By comparing the 1997 results to what we find using state-level data and modern methodology, we can assess the validity and accuracy of the original study’s findings.

In their study, Lott and Mustard employed Twoway Fixed Effects (TWFE) models to control for bias in their estimates. This model provides important predictive power in causal inference studies exploring multiple groups across multiple years. However, new techniques provide a deeper understanding of the time and group effects at play within TWFE models. We extend the original study using three additional methods. First, we employ the Bacon Decomposition², which dissects the different weights and pieces of the original coefficients. Then, we explore the Callaway and Sant’Anna³ estimator and Sun and Abraham⁴ event study, which allow us to further analyze group-time average treatment effects and treatment effect heterogeneity.

In the following sections, we review a brief background on the context and economic theory of the 1997 study before describing the data used in this replication. Then, we explore, in detail, the original methodology along with the three new techniques. Lastly, we conclude our findings and summarize how the new data and methods compare to the original study.

¹Lott, Jr., John R., and David B. Mustard. “Crime, Deterrence, and Right-to-Carry Concealed Handguns.” *The Journal of Legal Studies* 26, no. 1 (1997): 1–68. <https://doi.org/10.1086/467988>.

²Goodman-Bacon, Andrew. 2019. “Difference-in-Differences with Variation in Treatment Timing.”

³B. Callaway and P.H.C. Sant’Anna, Difference-in-Differences with multiple time periods. *Journal of Econometrics* (2020), <https://doi.org/10.1016/j.jeconom.2020.12.001>

⁴Liyang Sun, Sarah Abraham, Estimating dynamic treatment effects in event studies with heterogeneous treatment effects, *Journal of Econometrics*, Volume 225, Issue 2, 2021, Pages 175-199, ISSN 0304-4076, <https://doi.org/10.1016/j.jeconom.2020.09.006>

2 Background and Economic Theory

The 1997 study focused primarily on the impact of changes in concealed carry laws in the United States. Specifically, the authors explored how passing a concealed carry law may affect different crime rates in the United States. The authors used county-level data for 18 different states with information on arrest rates, crime rates, and various demographic and economic information. The data also identified whether or not a state had an active “shall issue” law. The “shall issue” law requires officials to only issue concealed carry gun permits to individuals that passed basic criminal background and mental health screenings. This “shall issue” law is a useful indicator for change in concealed carry laws, therefore, the assigned treatment in this study was each state’s rollout of the “shall issue” law. Table 1 shows when each state implemented the concealed carry law during the 1977-1992 time period.

Table 1: State Rollout of “shall issue”

State	Year
Alabama	Pre-1977
Connecticut	Pre-1977
Indiana	Pre-1977
New Hampshire	Pre-1977
North Dakota	Pre-1977
South Dakota	Pre-1977
Vermont	Pre-1977
Washington	Pre-1977
Maine	1985
Florida	1987
Virginia	1988
Georgia	1989
Pennsylvania	1989
West Virginia	1989
Idaho	1990
Mississippi	1990
Oregon	1990
Montana	1991

The authors used TWFE models to mitigate bias across county and year. They regressed the natural log of various crime rates onto the shall issue binary indicator variable, each crime arrest rate, and various economic and demographic controls. The authors studied several different categories of crime, including: murder, rape, aggravated assault, robbery, burglary, larceny, and auto theft. Additionally, murder, rape, aggravated assault, and robbery were grouped into one category called “violent crimes”, while burglary, larceny, and auto theft were grouped into “property crimes”. This makes up a total of nine different crime categories that were included in this study.

After estimating the effect of the shall issue law on each of the nine crime rates, the authors’ main results found a negative effect for both violent crimes and property crimes. These results imply that concealed carry laws are correlated with less overall crime in the United States.

3 Data

The data used for this study includes information on the same 18 states as the original study, just at the state level. In addition to each state's shall issue law, the data includes crime and arrest rates for each of the nine crime categories, as well as economic and demographic information such as population, income, age, gender, and race. Because this data is at the state level, there are less observations than the original study. However, the average sample size for each variable is about 800, which should be sufficient to provide accurate results.

Table 2.1 summarizes the average crime rates and arrest rates for each crime category. Additionally, this table summarizes average population and economic characteristics, such as population density, personal income, and unemployment insurance. We can see the most common crimes, based on average crime rates, tend to be property crimes. Specifically, larceny and burglary are the most common crimes committed in this data.

Table 2.1: National Sample Means and Standard Deviations

Variable	N	Mean	S.D.
Gun ownership information:			
Shall issue dummy	816	.191	0.393
Arrest rates (ratio of arrests to offenses) for a particular crime category:			
Violent crimes	802	41.091	22.204
Property crimes	809	16.918	4.677
Murder	806	91.299	55.943
Rape	799	41.023	17.389
Aggravated assault	809	44.625	16.978
Robbery	808	31.458	13.593
Burglary	809	13.804	4.571
Larceny	809	18.537	5.196
Auto theft	808	22.345	37.611
Crime rates are defined per 100,000 people:			
Violent crimes	816	483.926	318.943
Property crimes	816	4,618.339	1,210.465
Murder	816	7.768	6.882
Rape	816	33.982	15.072
Robbery	816	163.421	176.251
Aggravated assault	816	278.755	159.650
Burglary	816	1,239.336	417.758
Larceny	816	2,968.708	751.023
Auto theft	816	410.295	231.154
Real per capita income data (in real 1983 dollars):			
Personal income	816	9,351.821	4,689.701
Unemployment insurance	816	50.019	38.081
Income maintenance	816	115.276	70.953
Retirement payments per person over 65	816	1,002.226	546.468
Population characteristics:			
State population per square mile	816	355.973	1,408.250
State population	816	4,646,787.3	5,010,349.87

Table 2.2 summarizes the remaining demographic information on age, race, and gender. We can see that, on average, the highest population group is comprised of white males and females between the ages 20-29.

Table 2.2: Demographics Summary

Variable	<i>N</i>	Mean	S.D.
Race and age data (% of population):			
Black male 10-19	816	.01	0.011
Black female 10-19	816	.01	0.011
White male 10-19	816	.067	0.015
White female 10-19	816	.064	0.015
Other male 10-19	816	.004	0.008
Other female 10-19	816	.004	0.007
Black male 20-29	816	.01	0.010
Black female 20-29	816	.01	0.012
White male 20-29	816	.074	0.012
White female 20-29	816	.073	0.012
Other male 20-29	816	.004	0.007
Other female 20-29	816	.004	0.007
Black male 30-39	816	.007	0.008
Black female 30-39	816	.008	0.010
White male 30-39	816	.066	0.012
White female 30-39	816	.066	0.012
Other male 30-39	816	.003	0.007
Other female 30-39	816	.003	0.007
Black male 40-49	816	.005	0.006
Black female 40-49	816	.005	0.007
White male 40-49	816	.048	0.009
White female 40-49	816	.048	0.009
Other male 40-49	816	.002	0.005
Other female 40-49	816	.002	0.005
Black male 50-64	816	.005	0.007
Black female 50-64	816	.006	0.009
White male 50-64	816	.058	0.010
White female 50-64	816	.062	0.012
Other male 50-64	816	.002	0.006
Other female 50-64	816	.002	0.007
Black male over 65	816	.004	0.005
Black female over 65	816	.005	0.008
White male over 65	816	.043	0.011
White female over 65	816	.062	0.017
Other male over 65	816	.001	0.005
Other female over 65	816	.001	0.005

4 Empirical Model and Estimation

This section explores in detail the advantages and limitations of each estimation technique. First, we analyze the TWFE model employed in the original study. Then, we explore the three new methods: Bacon Decomposition, Callaway and Sant'Anna estimator, and Sun and Abraham event study.

4.1 Twoway Fixed Effects

The TWFE model works to mitigate bias across multiple groups over multiple years. For this replication study, our model specification is as similar to the original study as possible. We regress the natural log of each crime rate onto its respective arrest rate and the remaining demographic control variables, while controlling for state and year effects⁵. The main difference between this model and original model, is we use state data instead of county data.

Table 3 summarizes the results with the main coefficients of interest.

Table 3: Two Way Fixed-Effects Shall Issue Coefficients

Variable	Coefficient	T-Statistic
log(Violent Crime Rate)	-0.0978	-4.37
log(Murder Rate)	-0.0507	-1.26
log(Rape Rate)	-0.0340	-1.22
log(Aggravated Assault Rate)	-0.1004	-3.46
log(Robbery Rate)	-0.0532	-1.71
log(Property Crime Rate)	-0.0072	-0.55
log(Burglary Rate)	-0.0461	-2.43
log(Larceny Rate)	0.0033	0.27
log(Auto Theft Rate)	-0.0090	-0.32

Similar to the original study, all coefficients, except for larceny, had an estimated negative relationship with the shall issue law. However, the only coefficients that were estimated at a significant level were violent crime, aggravated assault, and burglary. The two coefficients with the highest magnitude are aggravated assault and violent crime, about -0.1000 and -0.098, respectively. This implies that concealed carry laws may have the largest and most significant impact on reducing the rate of aggravated assault and violent crimes.

4.2 Bacon Decomposition

The Bacon Decomposition provides a deeper understanding of the different weights and parts of the TWFE estimator. The TWFE coefficient is a weighted average of all potential 2x2 DD estimates, which are based on group size and variance. The TWFE estimator is a weighted average of all possible ATTs, under the assumption of variance weighted common trends (VWCT) and time invariant treatment effects. However, when there are in fact time-variant treatment effects (as in this study) the TWFE estimator may be severely biased.

The Bacon Decomposition breaks down the TWFE estimator into the different 2x2 coefficients for each comparison group, and provides weights to each group. The TWFE estimator ultimately aggregates over these weights and coefficients to output the one coefficient we see. Therefore, the TWFE coefficient relies on group variation to calculate its parameter.

⁵See Appendix 6.1 for full regression output

It is important to note that none of the Bacon Decompositions in this study include controls, so the coefficient estimates won't directly reflect the coefficients seen in Table 3.

Table 4.1 shows the Bacon Decomposition for the violent crime rate TWFE coefficient. We can see that the Treated vs. Untreated group has an estimate of about -0.112 and has the largest weight, making up about 75% of the total estimator. From these estimates and weights (despite not including controls), we can better understand how the TWFE coefficient was calculated.

However, the Later Treatment vs. Earlier Comparison group presents a potential problem in the 2x2 Bacon Decomposition. This group relies on the assumption of parallel trends and homogeneous treatment effects across time for the earlier treated group. This assumption often won't hold when considering differential treatment effects. Thankfully, this 2x2 group only makes up about 2% of the TWFE estimator, so it may only present a marginal amount of bias.

Table 4.1: Bacon Decomposition log(Violent Crime Rate)

Total TWFE: -0.085		
DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	0.076
Later T vs. Earlier C	0.023	-0.076
T vs. Never treated	0.749	-0.112
T vs. Already treated	0.159	-0.028
T = Treatment; C = Comparison		

Table 4.2 shows the Bacon Decomposition for the murder crime rate TWFE coefficient⁶. Note that the group weights remain the same in each decomposition, because they are dependent on groups over a specific time period. Because all coefficients are estimated over the same time period, the group weights remain the same. Therefore, the Treated vs. Never Treated group once again has the largest weight with an estimate of about -0.056.

Table 4.2: Bacon Decomposition log(Murder Rate)

Total TWFE: -0.037		
DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	0.080
Later T vs. Earlier C	0.023	0.002
T vs. Never treated	0.749	-0.056
T vs. Already treated	0.159	-0.008
T = Treatment; C = Comparison		

4.3 Callaway and Sant'anna

Another approach to solving the weighting problem of the TWFE estimator, is the Callaway and Sant'Anna estimator. This approach estimates group-time average treatment effects, which is just the ATT for a given group at any point

⁶See Appendix 6.2 for remaining Bacon Decompositions

in time. However, this group-time ATT assumes parallel trends conditional on time-invariant covariates and weights. Because this approach calculates group ATT by time, you can end up with several treatment effect parameters in each group. Callaway and Sant'Anna solves this issue by taking all of the treatment effects and collapsing them into one interpretable parameter for each group.

Table 5 shows these collapsed overall ATTs for each group. Some of these estimates are significantly different than the initial TWFE estimates we found. For example, the violent crime rate estimate went from about -0.098 to about -0.010, which is almost a tenth of the initial magnitude. Moreover, many parameters switched signs from their initial TWFE. For example, the estimate for the rape crime rate went from about -0.034 to about 0.013.

Table 5: Callaway and Sant'anna Overall ATTs

Variable	Overall ATT	SE
log(Violent Crime Rate)	-0.010	0.026
log(Murder Rate)	-0.083	0.027
log(Rape Rate)	0.013	0.022
log(Aggravated Assault Rate)	-0.036	0.046
log(Robbery Rate)	0.039	0.037
log(Property Crime Rate)	0.029	0.019
log(Burglary Rate)	-0.016	0.016
log(Larceny Rate)	0.030	0.015
log(Auto Theft Rate)	0.011	0.040

These large changes in magnitude and direction illuminate the importance of group-time treatment effects, and how vastly they can differ from the original TWFE estimates. Moreover, the significant changes in the Callaway and Sant'Anna estimates could be due to a violation of the parallel trends assumption, which would greatly bias our estimates.

4.4 Event Study: Sun and Abraham

For our third model extension, we explore how Sun and Abraham address some of the biases brought on by heterogeneous treatment effects. Specifically, we explore how Sun and Abraham address some of the issues caused by event studies with differential timing.

As previously uncovered in the Bacon Decomposition and the Callaway and Sant'Anna estimator, leads and lags are often used to measure the treatment in an event study with differential timing. However, these can produce causally uninterpretable results if there are heterogeneous treatment effects over time. So, the Sun and Abraham event study figures help us visualize if the pretrends assumption holds and the average treatment effects over time.

We use the Sun and Abraham function on treatment year and year for each outcome variable, and we include state and year fixed effects controls in our models⁷. Figure 1 shows the event study figure for the violent crime rate. We can see

⁷See Appendix 6.3 for remaining Sun and Abraham figures

that the pre-treatment estimates are statistically different than zero for most of the pre-treatment period, which implies the pretrend assumptions do not hold for this model. Meanwhile, the post-treatment estimates are only statistically different than zero for a couple years after treatment before returning to zero. These results provide some concerns to the validity and longevity of the treatment effect for violent crimes.

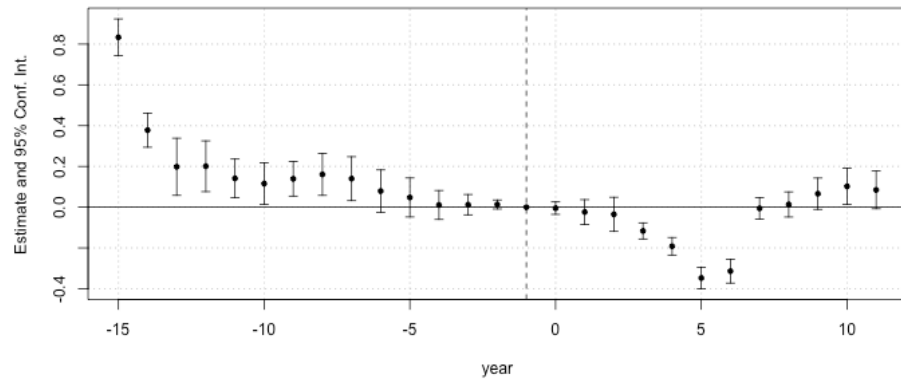


Figure 1: Sun and Abraham: $\log(\text{Violent Crime Rate})$

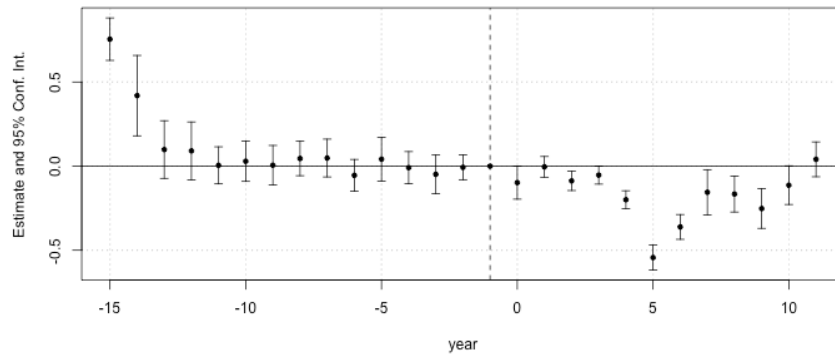


Figure 2: Sun and Abraham: $\log(\text{Murder Rate})$

Figure 2 shows the Sun and Abraham event study for the murder rate. Unlike Figure 1, there does appear to be evidence of pretrend assumptions holding, as most of the pre-treatment estimates are not statistically different than zero. However, the post-treatment period experiences only a brief treatment effect, similar to that in Figure 1. It appears that there is no statistically significant effect for the first couple years after treatment, then there is a significant treatment effect for a few years before returning to zero again. Once again, this trend raises concern for the longevity of these treatment effects.

5 Conclusion

In their 1997 study, Lott and Mustard developed a TWFE model to identify the effect of concealed carry gun laws on various crime rates in the United States. In an effort to compare and extend their results, we replicated their original TWFE model using state-level data. Then, we extended their analysis with three more modern approaches. First, we used the Bacon Decomposition to break down the TWFE coefficient into different comparison groups and developed a deeper understanding of the TWFE coefficient composition. Next, we calculated Callaway and Sant'Anna estimators to better understand the group-time average treatment effects. Lastly, we employed the Sun and Abraham event study to better understand the average treatment effect over time and identify potential bias from heterogeneous treatment effects.

While our replicative TWFE model found similar results as the original study, the three additional models provided further insight to the validity of the TWFE estimates. In studies regarding differential timing and potential heterogeneous treatment effects, it is important to understand the effect of leads and lags and group-time treatment effects. These sources of bias can confuse the true causal implications of treatments, so it is important to be vigilant in identifying potential sources of bias. Moreover, this study brings to light the effectiveness of using several different models. Each model has its own strengths and weaknesses in identifying treatment effects and biases. Therefore, employing a few different models can bring strengths in each area of weakness, and provide a fuller understanding of the true treatment effects and causal implications.

6 Appendix

6.1 Twoway Fixed Effects: Extended Regression Output

Table 3.1: Two-Way Fixed Effects Model (extended)

	log(Violent Crime Rate)	log(Murder Rate)	log(Rape Rate)	log(Aggravated Assault Rate)	log(Robbery Rate)
	(1)	(2)	(3)	(4)	(5)
<i>Variables</i>					
Shall issue dummy	-0.0978*** (-4.37)	-0.0507 (-1.26)	-0.0340 (-1.22)	-0.1004*** (-3.46)	-0.0532 (-1.71)
Crime arrest rate	-0.0003 (-1.69)	-0.0004*** (-3.35)	-0.0006 (-1.80)	-0.0028*** (-7.12)	-0.0014** (-2.73)
<i>Fixed effects</i>					
State	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
<i>Fit Statistics</i>					
N	802	806	799	809	808
F	27.779	7.541	32.584	29.594	17.763
R-squared	0.7031	0.3900	0.7362	0.7141	0.6002

T-statistic in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Omitted control variables: density, rpcpi, rpcui, rpcim, rpcrpo, popstate, ppbm, ppbf*, ppwm*, ppwf*, ppnm*, ppnf**

Table 3.2: Two-Way Fixed Effects Model (extended)

	log(Property Crime Rate)	log(Burglary Rate)	log(Larceny Rate)	log(Auto Theft Rate)
	(6)	(7)	(8)	(9)
<i>Variables</i>				
Shall issue dummy	-0.0072 (-0.55)	-0.0461* (-2.43)	0.0033 (0.27)	-0.0090 (-0.32)
Crime arrest rate	-0.0020** (-2.85)	-0.0052*** (-4.30)	-0.0011 (-1.91)	-0.003 (-1.35)
<i>Fixed effects</i>				
State	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
<i>Fit Statistics</i>				
N	809	809	809	808
F	26.987	33.616	31.426	28.340
R-squared	0.6949	0.7394	0.7262	0.7055

T-statistic in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Omitted control variables: density, rpcpi, rpcui, rpcim, rpcrpo, popstate, ppbm, ppbf*, ppwm*, ppwf*, ppnm*, ppnf**

6.2 Bacon Decomposition: Remaining Tables

Table 4.3: Bacon Decomposition log(Rape Rate)

Total TWFE: : -0.032

DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	-0.039
Later T vs. Earlier C	0.023	-0.082
T vs. Never treated	0.749	0.004
T vs. Already treated	0.159	-0.192
T = Treatment; C = Comparison		

Table 4.4: Bacon Decomposition log(Aggravated Assault Rate)

Total TWFE: -0.132

DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	0.116
Later T vs. Earlier C	0.023	-0.147
T vs. Never treated	0.749	-0.184
T vs. Already treated	0.159	-0.006
T = Treatment; C = Comparison		

Table 4.5: Bacon Decomposition log(Robbery Rate)

Total TWFE: 0.017

DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	0.108
Later T vs. Earlier C	0.023	0.090
T vs. Never treated	0.749	-0.013
T vs. Already treated	0.159	0.110
T = Treatment; C = Comparison		

Table 4.6: Bacon Decomposition log(Property Crime Rate)

Total TWFE: 0.029

DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	-0.011
Later T vs. Earlier C	0.023	0.006
T vs. Never treated	0.749	0.028
T vs. Already treated	0.159	0.054
T = Treatment; C = Comparison		

Table 4.7: Bacon Decomposition log(Burglary Rate)

Total TWFE: 0.029

DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	-0.011
Later T vs. Earlier C	0.023	0.006
T vs. Never treated	0.749	0.028
T vs. Already treated	0.159	0.054
T = Treatment; C = Comparison		

Table 4.8: Bacon Decomposition log(Larceny Rate)

Total TWFE: 0.036

DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	-0.006
Later T vs. Earlier C	0.023	0.021
T vs. Never treated	0.749	0.037
T vs. Already treated	0.159	0.050
T = Treatment; C = Comparison		

Table 4.9: Bacon Decomposition log(Auto Theft Crime Rate)

Total TWFE: 0.068

DD Comparison	Weight	Avg DD Estimate
Earlier T vs. Later C	0.068	0.083
Later T vs. Earlier C	0.023	0.087
T vs. Never treated	0.749	0.035
T vs. Already treated	0.159	0.211
T = Treatment; C = Comparison		

6.3 Sun and Abraham: Remaining Figures

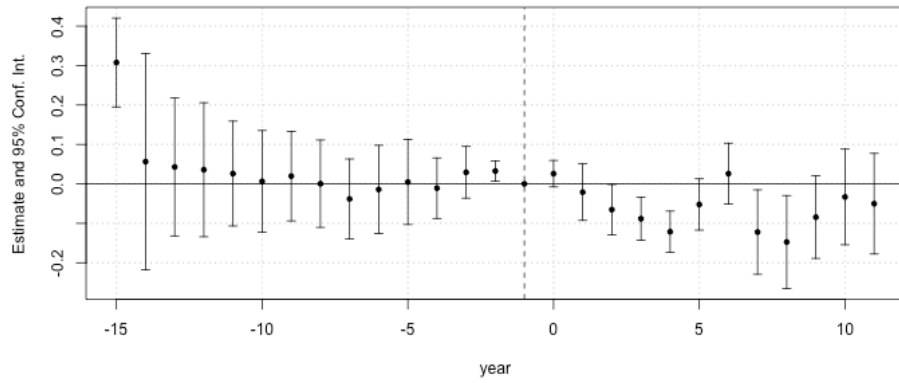


Figure 3: Sun and Abraham: $\log(\text{Rape Rate})$

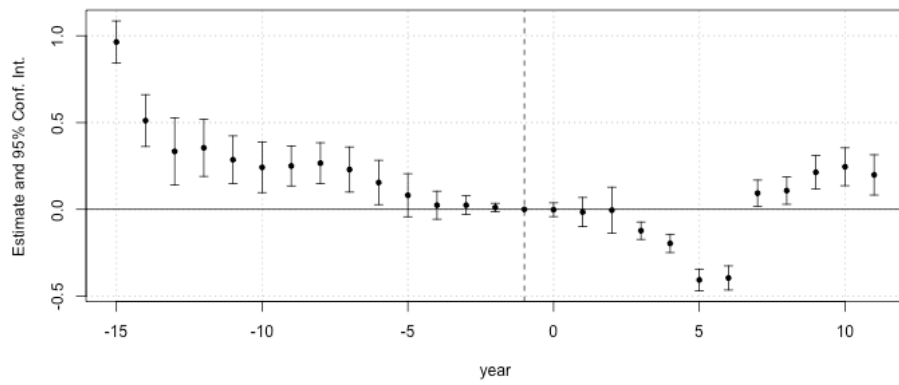


Figure 4: Sun and Abraham: $\log(\text{Aggravated Assault Rate})$

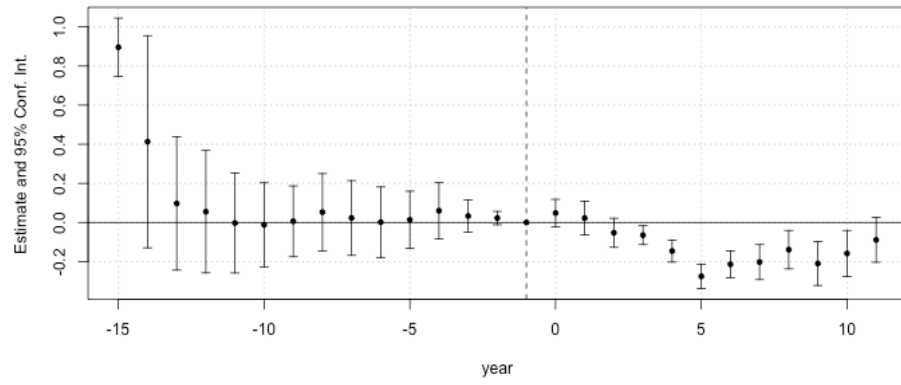


Figure 5: Sun and Abraham: log(Robbery Rate)

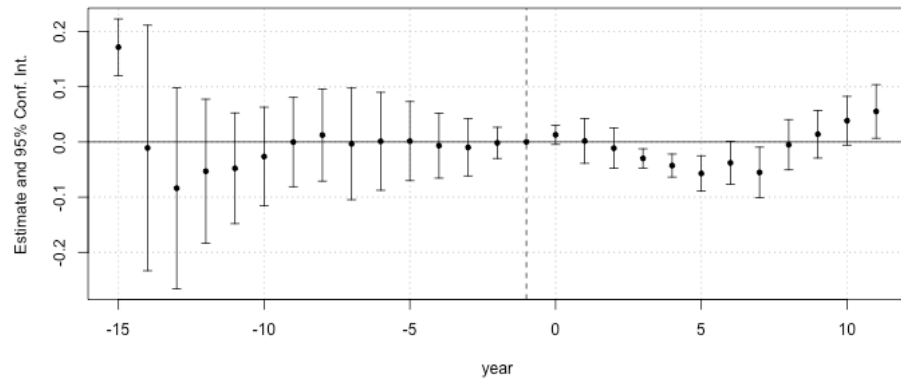


Figure 6: Sun and Abraham: log(Property Crime Rate)

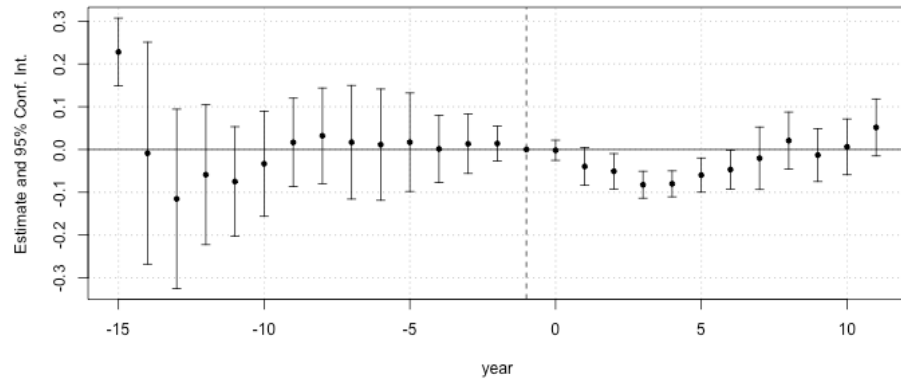


Figure 7: Sun and Abraham: log(Burglary Rate)

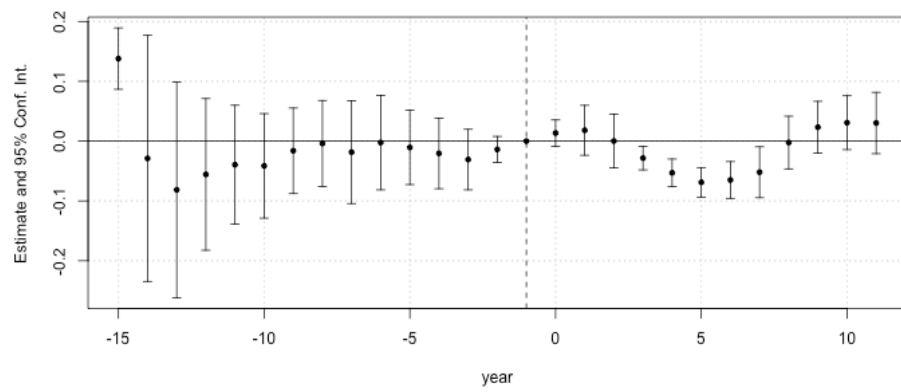


Figure 8: Sun and Abraham: log(Larceny Rate)

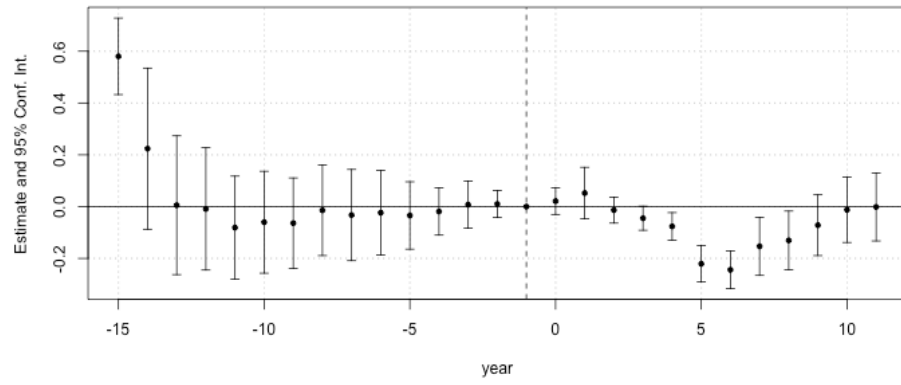


Figure 9: Sun and Abraham: $\log(\text{Auto Theft Rate})$