EJW
ECON JOURNAL WATCH
Scholarly Comments on
Academic Economics

ECON JOURNAL WATCH 16(1)
March 2019: 84–96

Do Right to Carry Laws Increase Violent Crime? A Comment on Donohue, Aneja, and Weber

Carlisle E. Moody[1] and Thomas B. Marvell

LINK TO ABSTRACT

Along with coauthors, John J. Donohue of the Stanford Law School regularly releases empirical research finding that certain laws concerning the carrying of firearms, those known as right-to-carry or RTC laws, increase violent crime. We have engaged with the previous releases to show that the latest findings are not robust, and here we do so again. This time we address something maintained in the previous releases, namely the weighting of fixed-effects regressions by population, and something that is new to the last two versions, namely a synthetic control procedure. Perhaps after another round or two it will be time to recap all the robustness criticisms to date, but here we confine ourselves to the two new criticisms.

Donohue, Abhay Aneja, and Kyle Weber (hereafter DAW) have produced two revisions (2018a; b) to their 2017 working paper, which itself is part of a series of papers dating back 15 years (Donohue 2003; Aneja, Donohue, and Zhang 2010; 2011; 2012; 2014; see also Aneja, Donohue, Pepper, et al. 2012). That paper adds a synthetic control analysis and finds that RTC laws increase violent crime. They conclude, as they state in the abstract of the latest revision:

> Our preferred panel data regression specification…generates statistically significant estimates showing RTC laws *increase* overall violent crime. Our synthetic control approach also strongly confirms that RTC laws are associated with 13–15 percent *higher* aggregate violent crime rates… [T]he

1. College of William and Mary, Williamsburg, VA 23187.

> average RTC state would need to roughly double its prison population to offset the increase in violent crime caused by RTC adoption. (DAW 2018b, 1, italics in original)

We argue that DAW's results are fragile and most likely incorrect. We investigate two features of the research: the weighting of their fixed-effects regressions by population, and the use of synthetic controls. Both topics apply to numerous other difference-in-differences (DD) policy studies, so the discussion here has relevance beyond the DAW papers.

Weighting by population

DAW (2018a; b) base their conclusions on DD regressions that are weighted by population, a common practice in this literature and one which we have often used unquestioningly. They do not give reasons for weighting, nor to our knowledge have other researchers in this area, even though the results may well depend on the choice of weight.

Donohue has noted that weighting causes large states to dominate DD results (e.g., Ayres and Donohue 2003, 1276), but without stating that such emphasis is the purpose for weighting. Such might be the purpose if the research goal is to estimate the overall national impact of a policy change, because then weighting can be justified by arguing that the impact of laws in large states should be emphasized simply because they affect more people, or more populous states should be weighted more heavily because their data have smaller variance. On the other hand, if policy makers in states without an RTC law want information about whether a new law will have an impact in their state, the results of a weighted regression are not particularly useful unless their state happens to be large. If Donohue and his colleagues desire to emphasize the influence of large states, they should specify as much.

Weighting is also used to address heteroskedasticity, a potential problem in any regression analysis, and perhaps that is what DAW aim to do. Traditionally the standard econometric approach to heteroskedasticity was to test for it, and if found, to weight by some function of the variables in the model. This procedure is known as weighted least squares (WLS) or feasible generalized least squares (FGLS). For the crime equation, one common approach was to assume that the underlying unobserved individual crime equation had independent and identically distributed (i.i.d.) errors, which when summed and divided by the population to produce per capita crime rates within states, generated heteroskedasticity of the form σ^2/pop_i for states i=1…N, so that multiplying the state variances by state population would produce homoskedastic errors. However, it is not clear that this approach is invariably correct. If the individual crime equation errors do not have i.i.d. errors then the population-weighted least squares approach would be incorrect and may exacerbate the problem.

The availability of heteroskedasticity-consistent standard errors (White 1980) has led many to abandon WLS. The Stata command for clustered standard errors includes the White correction, and Donohue and his colleagues cluster by state. In practice researchers never really know the true form of the error variance. Consequently, as Jeffrey Wooldridge notes, "more and more researchers simply use OLS and compute robust standard errors when estimating models using per capita data" (2016, 258–259). Also, according to James Stock and Mark Watson, "it is simple to use heteroskedasticity-robust standard errors, and the resulting inferences are reliable under very general conditions…without needing to specify a functional form for the conditional variance. For these reasons, it is our opinion that, despite the theoretical appeal of WLS, heteroskedasticity-robust standard errors provide a better way to handle potential heteroskedasticity in most applications" (2007, 696). Similarly, Joshua Angrist and Jörn-Steffen Pischke write, "Any efficiency gain from weighting is likely to be modest, and incorrect or poorly estimated weights can do more harm than good" (2009, 94).

On the other hand, Edward Leamer calls this practice "White-washing" and says "we should be doing the hard work of modeling the heteroskedasticity and the time dependence to determine if sensible reweighting of the observations materially changes the locations of the estimates of interest as well as the widths of the confidence intervals" (2010, 43). Joseph Romano and Michael Wolf (2017) propose what they call "adaptive least squares" which requires using a standard test (e.g., White 1980) for heteroskedasticity. If the test does not reject the null of homoskedasticity, use OLS with robust standard errors. If the test rejects, model the variance function and use the inverse of the predicted value as the weight, with inference based on robust standard errors. Wooldridge (2016, 262–264) also recommends combining FGLS with robust standard errors.

Does it make a difference whether one weights the per capita crime equation by population? We use the DAW "preferred" model. We made some slight modifications because we do not have exactly the same data as DAW.[2] For example, DAW use the percent of the population living in MSAs to measure population density; we use the population-to-area ratio. They use lagged total police employment to measure police presence; we use the lagged number of sworn officers. Finally, they use six demographic variables: the proportion of black, white, and other males in two age groups (15–19 and 20–39). We use the percent of the

2. Despite two requests we were unable to get the DAW data and do files.

population 15–39 by five-year intervals and the proportion of the total population that is black. Also, we have data from 1970 as opposed to DAW whose data start in 1977. Like DAW, we include a trend and state and year fixed effects, and also like DAW we use robust clustered standard errors.

DAW estimate two versions of their crime equation. The first is a standard difference in differences analysis using a dummy variable to represent the presence or absence of an RTC law. The second uses a spline (post-law trend) to capture the effects of the law over time. We estimated each model with and without population weights and using FGLS. All models are estimated with errors clustered at the state level. We also tested for heteroskedasticity using the finite version of the White (1980) test. For all the models, the tests were invariably highly significant, indicating the presence of heteroskedasticity, despite the weighting.[3]

DAW, like many other researchers, assume that the variance function is $Var(u_{it}) = (1/pop_{it})\sigma^2$, where σ^2 is the homoskedastic error variance. If that assumption is incorrect then the resulting estimators are inefficient. We can investigate this assumption by estimating a more general form of the variance function including population:

$$u_{it}^2 = X_{it}\beta + \lambda(1/pop_{it}) + \alpha_i + \gamma_t + \upsilon_{it} \tag{1}$$

where u_{it}^2 is the squared regression error, the matrix X_{it} includes all the control variables in the DAW "preferred" model including the time trend, and α_i and γ_t are state and year fixed effects. If λ is significant and equal to one, and all of the other variables are jointly insignificant, then DAW's assumption is correct. We estimated equation (1) using the squared residuals from the OLS regression as the dependent variable. Although not reported to conserve space, there were many significant variables in these regressions, but the inverse of population was significant only in the violent crime models and it was never the only significant coefficient.

Following Leamer (2010), we use FGLS to estimate the appropriate weights for the crime equation, replacing the squared errors in (1) with the corresponding squared residuals and taking logs to guarantee positive variance estimates.

$$\log(\hat{u}_{it}^2) = X_{it}\beta + \lambda(1/pop_{it}) + \alpha_i + \gamma_t + \omega_{it} \tag{2}$$

The FGLS weight is the inverse of the antilog of the predicted value from (2). We also estimated an unweighted OLS model.

The results of this experiment for the murder rate and the violent crime rate

3. All our Stata do files, log files and data are available at the *Econ Journal Watch* website **(link)**.

equations are shown in Table 1, where the coefficients refer to the RTC dummy and spline variables.

TABLE 1. Fixed-effects models

	Murder Weighted (1)	Murder OLS (2)	Murder FGLS (3)	Violent Weighted (4)	Violent OLS (5)	Violent FGLS (6)
Dummy						
Coefficient	−0.466	−1.754	0.808	11.440***	−3.658	−1.440
Std. error	(4.113)	(3.985)	(2.309)	(3.679)	(4.451)	(1.979)
P-value	0.910	0.662	0.728	0.003	0.415	0.470
Spline						
Coefficient	0.352	−0.283	−0.115	1.148**	−0.201	−0.002
Std. error	(0.414)	(0.454)	(0.253)	(0.436)	(0.569)	(0.240)
P-value	0.399	0.536	0.651	0.012	0.725	0.992

Notes: All models include control variables equivalent to the DAW "preferred" model as well as state and year fixed effects. Coefficients and standard errors are reported in percentage terms. Weighted indicates population-weighted least squares. FGLS uses the inverse of the antilog of the predicted values of the log variance equation (2) as the weight. Standard errors are robust to heteroskedasticity and clustered at the state level for all models. * p < .1, ** p < .05, *** p < .01.

The DAW population-weighted least squares estimates are reported in DAW (2018b, 21, Table 3). For the murder equation, they find no significance. For the violent crime model, they find that only the dummy variable is significantly different from zero.

How one weights the regression does in fact make a difference. Our attempted replication of DAW, using population weights, yields similar results in terms of significance, except that we find that both the dummy and spline are significant in the violent crime equation. The OLS model finds no significant coefficients. The FGLS model, which has the most precise estimates, also finds no significant coefficients.[4]

Thus, we corroborate the DAW result that the RTC law has no significant effect on murder, but we also find that their conclusion that violent crime is increased by the passage of a RTC law is fragile insofar as it is dependent on the use of population weights.

4. We also tried using population instead of the inverse of population in the variance functions for both the squared residuals and the log of the squared residuals, the results reported in Table 1 were unchanged.

Fixed effects vs. synthetic controls

In the second part of their paper, DAW use the synthetic control (SC) method which also finds that RTC laws increase violent crime rates. The SC analysis also confirms that RTC laws do not significantly increase murder.

The synthetic control method compares the outcome (e.g., violent crime rate) of a treated state (i.e., a state with an RTC law) to a weighted average of the outcomes of the control states (i.e., states with no RTC law). The resulting weighted average, formed by using a matching algorithm (Abadie, Diamond, and Hainmueller 2010), is called the "synthetic control state." The result is the difference or "gap" between ($y_t^{treated}$) the outcome of the treated state and the weighted average of the outcomes of the control states ($y_t^{synthetic}$):

$$gap_t = y_t^{treated} - y_t^{synthetic} \tag{3}$$

Assume that the outcome is a function of the policy being investigated, measured by z_{it} (i=1,…,N; t=1,…,T) which could be a continuous or a dummy variable, a set of trends (tr_i), a set of control variables (X_{kit}, k=1,…,K), a set of state fixed effects (α_i), a set of year fixed effects (δ_t), and a random i.i.d. error term (ε_{it}). The outcomes of the control states are also functions of z_{it}, trends, control variables, state and year fixed effects and a random error term.

Therefore,

$$y_{it} = f(z_{it}, X_{kit}, \alpha_i, \delta_t, tr_i, \varepsilon_{it}) \tag{4}$$

where y_{it} is the outcome for state i in year t. Denoting the treated state as i=1 and the control states as i=2,…, N* < N,

$$gap_t = f(z_{1t}, X_{k1t}, \alpha_1, \delta_t, tr_1, \varepsilon_{1t}) - \textstyle\sum_{i=2}^{N*} c_i f(z_{it}, X_{kit}, \alpha_i, \delta_t, tr_i, \varepsilon_{it}) \tag{5}$$

where the weights, c_i, are nonnegative and sum to one.

The state and year fixed effects as well as the control variables and trends are in both the treated and control outcomes. Thus the gap is a function of all those factors. For the gap to measure the effect of the policy requires one to assume that all of the variables except for the policy dummy are constant in the treatment period, a very strong assumption.

DAW implicitly recognize this limitation when they argue that their gap measurements are a function of at least three exogenous control variables that are not held constant: "If one adjusts the synthetic controls estimates to control for the increased rates of police and incarceration that follow RTC adoption, the RTC-induced increases in murder are almost nine percent with a p-value of 0.089" (2018b, 42 n.64). Also, "Specifically, if…the death penalty is a powerful deterrent one might be concerned that Texas's far greater use of the death penalty during the post-passage period than in the states comprising synthetic Texas might bias downward the prediction that RTC laws increased crime by 16.9 percent in Texas. Conversely, the greater increases in incarceration and police in synthetic Texas would lead to the opposite bias" (2018a, 39). Finally, "Figure 3 makes clear what Texas is being compared to, and we can reflect on whether this match is plausible and whether anything other than RTC laws changed in these three states during the post-passage decade that might compromise the validity of the synthetic control estimate of the impact of RTC laws" (DAW 2018b, 30–31). Thus, DAW acknowledge that omitted exogenous control variables in both the treatment and control states could bias the results. In any case, the SC model fails to control for police, prison, and executions, as well as all the other control variables, in the treatment period.

Also, states have certain permanent, or at least long-lived, attributes such as history, tradition, culture, climate, attitudes toward firearms, attitudes toward self-defense, etc., that could affect both crime and the policies adopted to combat crime. Louisiana is fundamentally different from Massachusetts, Utah is very different from Nevada, Texas is different from Wisconsin, Alaska and Hawaii are different from each other and from every other state, and so on. The SC gap is determined in part by the difference in the mean of the dependent variable (e.g., violent crime) for the treated state and the weighted average of the corresponding means of violent crime for the states making up the synthetic control state. Thus the gap is a cross-section difference which is a function of the state fixed effects, implying that the SC technique suffers from unobserved heterogeneity. According to Manuel Arellano,

> [Unobserved heterogeneity] has been a pervasive problem in cross-sectional regression analysis. If characteristics that have a direct effect on both left- and right-hand side variables are omitted, explanatory variables will be correlated with errors and regression coefficients will be biased measures of the structural effects. Thus researchers have often been confronted with massive cross-sectional data sets from which precise correlations can be determined but that, nevertheless, had no information of policy interest. (Arellano 2003, 8)

The only way to avoid this type of omitted variable bias is to use panel data with fixed effects, which gives each cross-section unit its own intercept term, allowing these permanent effects to be partialled out of the parameter estimate corresponding to the policy variable.

According to Angrist and Pischke (2009, 244 n.9), the synthetic control model is a nonparametric version of the lagged dependent variable model:

$$y_{it} = \alpha + \beta_0 z_{it} + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{i=1}^{N-1} \gamma_i tr_i + \sum_{j=1}^{p} \varphi_j y_{t-j} + \varepsilon_{it}$$

where the lagged dependent variables control for unobserved time-varying confounders (Abadie et al. 2010, 495).

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.

On the other hand, the standard fixed-effects model estimates a linear approximation of the function *f*.

$$y_{it} = \alpha_i + \beta_0 z_{it} + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{i=1}^{N-1} \gamma_i tr_i + \varepsilon_{it}$$

The estimate of the effect of the policy is the coefficient $\hat{\beta}_0$, where the effects of the control variables, trends, and state and year fixed effects have been partialled out. These two approaches are very different. The fixed-effects regression model, properly specified, controls for all relevant factors, pre- and post-treatment, including trends and state and year fixed effects.[5] This cannot be said for the synthetic control model.

Nevertheless, we estimate the SC model of the violent crime rate and murder rate. Matching in the pre-treatment period is a function of the same control variables from DAW's "preferred model," that we used in the fixed effects analyses above. Following DAW we use the Stata *synth* program, limiting the post-treatment period to 10 years, and we include as potential control states all those states without RTC laws during the 10-year post-treatment period, including states that later adopted RTC laws. We use the default settings, except that we use both the nested matching option and the default regression method. We also use the regression option for those states for which the nested option fails. We do not force the gap to be zero for the year in which the policy was implemented, because that makes that one year more important than all the other years. Also, according to DAW, "As it turns out, the choice we made to subtract off the initial-year crime discrepancy

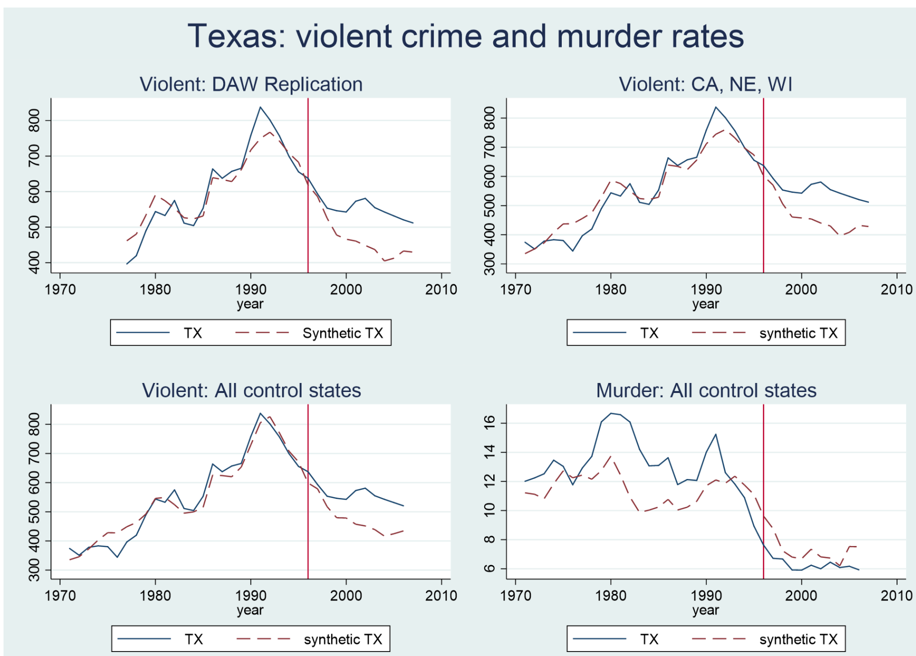
5. Including lagged dependent variables in the fixed effects model will control for unobserved time-varying confounders, but will cause bias (Nickell 1981). However, the bias is small for panels with N>20 and T>30 (Judson and Owen 1999).

is a conservative one, in that the estimated crime increases from RTC laws would be *greater* without subtraction" (2018b, 35 n.55; italics in original). Finally, as noted above, we have seven more years of data in the pre-treatment period.

We determine whether the gap is significant for each state by using a t-test of the null hypothesis that the mean gap over the 10 years of the post-treatment period is equal to zero. We use a t-test on the mean of the 10-year sums over the 33 treated states to determine the overall effect of the law.

As a first step, we do a SC analysis of violent crime for the state of Texas, an example featured in DAW (2018b, 28–30) so we can easily compare our results to theirs. We replicate DAW by limiting the analysis to 1977–2007. The results are shown in Figure 1.

Figure 1. Synthetic control applied to Texas



The top left graph shows the Texas violent crime rate (solid line) and the weighted average of the violent crime rates for the three states chosen by the SC matching algorithm, using the weights reported by DAW (2018b, 30): California (.577), Nebraska (.097), and Wisconsin (.326).[6] This appears to be exactly the figure

6. Even though both California and Wisconsin differ from Texas in terms of history, tradition, culture, climate, attitudes toward firearms, attitudes toward self-defense, etc.

presented by DAW (2018b, 30, 110).

The top right graph shows the result of our application of the SC method to the Texas data where we use all the data from 1970 to 2007 but limited to the three control states chosen by DAW. The two graphs are virtually identical despite the fact that the SC program chose a slightly different synthetic Texas (California .578, Wisconsin .422). In the lower left graph we expand the number of control states to include all the states without RTC laws until 2007. Again the graph is similar to the first two. All three show a large positive gap after 1996 apparently indicating that the Texas RTC law increased violent crime. A robust t-test reveals that the mean over the post-treatment period is significantly positive. This confirms that our analyses are equivalent to those of DAW.

The situation is much different for the murder rate, shown in the lower right-hand corner. The rate for the real Texas is below that of synthetic Texas for all post-treatment years. The robust t-test indicates that the gap is significantly negative. The murder rate for the synthetic Texas was below that for real Texas until 1993, before the passage of its RTC law, after which the real Texas murder rate fell below the synthetic Texas rate, indicating that the RTC law is unlikely to be entirely responsible for the decline. However, there was no increase in the murder rate after 1996 that could be attributed to the Texas RTC law.

The results of the full SC analyses are reported in Table 2.

TABLE 2. Synthetic control analyses of violent crime and murder

	Positive	Negative	Positive Significant	Negative Significant	Mean Gap	T-statistic
Nested						
Violent	15	18	14	12	122.6	0.55
Murder	13	20	8	11	0.85	0.38
Default						
Violent	16	17	13	16	−157.7	−0.45
Murder	17	16	13	11	0.505	0.14

Notes: "Significant" means $p < .05$ for the robust t-statistic across the post-treatment period. The mean gap is the mean of the sums of the 10 yearly post-treatment gaps for each state across all 33 states. The T-statistic refers to the robust t-test of the hypothesis that the net effect across the 33 treated states is zero.

There is no obvious significance here. For violent crime, the number of negative results (35 reductions in crime) is greater than the number of positive results (31 increases in crime) as a result of the RTC law. However, with respect to significant results it is a virtual tie (27 significantly positive and 28 significantly negative). The results are similar for murder, the number of negative results (36) is again greater than the number of positive effects (30), and the number of significant effects is almost the same (21 positive, 22 negative). The one consistent result across all four models is that the net effect across states is not significantly different from zero. In all, the SC analysis suggests that the RTC law cannot be shown to have a significant effect on either violent crime or murder.

Summary and conclusion

Our examination of the evidence indicates that there is no significant relationship between RTC laws and either murder or violent crime. The fixed-effects models estimated using our version of the DAW "preferred model" is fragile. Without weighting by population, there is no significant effect of RTC laws on violent crime. Using feasible generalized least squares confirms the OLS results that RTC laws do not increase violent crime or murder.

We also have a problem with the SC model that DAW use to buttress their population-weighted FE model. In the critical treatment period, the SC model fails to control for any of the major factors that cause crime rates to vary. It also suffers from unobserved heterogeneity. Nevertheless, when we use the SC model we find that the claim that RTC laws increase either murder or violent crime is not supported. We find states where crime increased after the implementation of the RTC law and we find more states in which crime decreased after the law. Our tests reveal that there is no significant overall net effect of the RTC laws on murder or violent crime across all 33 states that have implemented such laws.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effects of California's Tobacco Control Program. *Journal of the American Statistical Association* 105(490): 493–505.

Aneja, Abhay, John J. Donohue, John V. Pepper, Charles F. Wellford, and Alexandria Zhang. 2012. Erratum. *American Law and Economics Review* 14(2): 601–602. **Link**

Aneja, Abhay, John J. Donohue, and Alexandria Zhang. 2010. The Impact of Right-to-Carry Laws and the NRC Report: Lessons for the Empirical Evaluation of Law and Policy. Presented at the 5th Annual Conference on Empirical Legal Studies, Yale Law School (New Haven, Conn.), November. **Link**

Aneja, Abhay, John J. Donohue, 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.

Aneja, Abhay, John J. Donohue, and Alexandria Zhang. 2012. The Impact of Right to Carry Laws and the NRC Report: The Latest Lessons for the Empirical Evaluation of

Law and Policy. *NBER Working Paper* 18294. August. National Bureau of Economic Research (Cambridge, Mass.). **Link**

Aneja, Abhay, John J. Donohue, and Alexandria Zhang. 2014. The Impact of Right to Carry Laws and the NRC Report: The Latest Lessons for the Empirical Evaluation of Law and Policy. *NBER Working Paper* 18294 [revised]. November. National Bureau of Economic Research (Cambridge, Mass.). **Link**

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, N.J.: Princeton University Press.

Arellano, Manuel. 2003. *Panel Data Econometrics*. Oxford: Oxford University Press.

Ayres, Ian, and John J. Donohue. 2003. Shooting Down the "More Guns Less Crime" Hypothesis. *Stanford Law Review* 55: 1193–1312.

Donohue, John J.. 2003. The Impact of Concealed-Carry Laws. In *Evaluating Gun Policy: Effects on Crime and Violence*, eds. Jens Ludwig and Philip J. Cook, 287–324. Washington, D.C.: Brookings Institution Press.

Donohue, John J., Abhay Aneja, and Kyle D. Weber (DAW). 2017. Right-to-Carry Laws and Violent Crime: A Comprehensive Assessment Using Panel Data and a State-Level Synthetic Controls Analysis. *NBER Working Paper* 23510. June. **Link**

Donohue, John J., Abhay Aneja, and Kyle D. Weber (DAW). 2018a. Right-to-Carry Laws and Violent Crime: A Comprehensive Assessment Using Panel Data, the LASSO, and a State-Level Synthetic Controls Analysis. *NBER Working Paper* 23510 [revised]. January. **Link**

Donohue, John J., Abhay Aneja, and Kyle D. Weber (DAW). 2018b. Right-to-Carry Laws and Violent Crime: A Comprehensive Assessment Using Panel Data and a State-Level Synthetic Control Analysis. *NBER Working Paper* 23510 [revised]. November. **Link**

Judson, Ruth A., and Ann L. Owen. 1999. Estimating Dynamic Panel Data Models: A Guide for Macroeconomists. *Economics Letters* 65: 9–15.

Leamer, Edward E. 2010. Tantalus on the Road to Asymptopia. *Journal of Economic Perspectives* 24(2): 31–46.

Nickell, Stephen. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* 49(6): 1417–1426.

Romano, Joseph P., and Michael Wolf. 2017. Resurrecting Weighted Least Squares. *Journal of Econometrics* 197: 1–19.

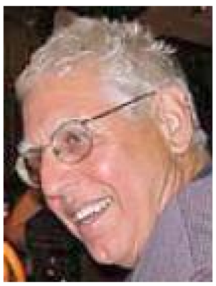
Stock, James H., and Mark W. Watson. 2007. *Introduction to Econometrics*, 2nd ed. Boston: Pearson/Addison-Wesley.

White, Halbert. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 48(4): 817–838.

Wooldridge, Jeffrey M. 2016. *Introductory Econometrics: A Modern Approach*, 5th ed. Boston: Cengage Learning.

About the Authors

Carlisle E. Moody is Professor of Economics at the College of William & Mary, where he teaches mathematical economics, econometrics, and time series analysis. His research is primarily in the economics of crime, especially the relationship between guns and crime. His email address is cemood@wm.edu.



Thomas B. Marvell is a lawyer-sociologist. His email address is marvell@cox.net.

Donohue, Aneja, and Weber's reply to this article
Go to archive of Comments section
Go to March 2019 issue

Discuss this article at Journaltalk:
https://journaltalk.net/articles/5982/