


Do execution moratoriums increase homicide? Re-examining evidence from Illinois

A. Ahrens, T. V. Kovandzic & L. M. Vieraitis


To cite this article: A. Ahrens, T. V. Kovandzic & L. M. Vieraitis (2015) Do execution moratoriums increase homicide? Re-examining evidence from Illinois, *Applied Economics*, 47:31, 3243-3257, DOI: [10.1080/00036846.2015.1013613](https://doi.org/10.1080/00036846.2015.1013613)

To link to this article: <https://doi.org/10.1080/00036846.2015.1013613>

 [View supplementary material](#) 

 Published online: 23 Feb 2015.

 [Submit your article to this journal](#) 

 Article views: 319

 [View related articles](#) 

 [View Crossmark data](#) 

 Citing articles: 1 [View citing articles](#) 

Do execution moratoriums increase homicide? Re-examining evidence from Illinois

A. Ahrens^{a,*}, T. V. Kovandzic^b and L. M. Vieraitis^b

^a*Department of Accountancy, Economics & Finance, Heriot-Watt University, Edinburgh EH14 4AS, UK*

^b*Program in Criminology, University of Texas at Dallas, Richardson, TX 75083-0688, USA*

This article revisits the event study by Cloninger and Marchesini (2006), who find that the declaration of the Illinois' death penalty moratorium on 31 January 2000 had a homicide-promoting effect and resulted in 150 additional homicides over the period 2000–2003. We reassess the author's identification strategy, which they refer to as 'portfolio approach' and which draws upon event studies in finance research. We argue that their methodology is not applicable in crime studies. Instead, we apply univariate time-series methods to test for a structural break at a known and unknown break date. We allow for unknown break points as the structural break might have occurred slightly earlier (criminals might have anticipated the moratorium) or later (due to persistence in criminal behaviour). In addition, we implement the synthetic control estimator which approximates the counterfactual homicide series by a weighted average of homicide outcomes in other US states. Based on various testing methods and two distinct data sets, we conclude that there is no empirical evidence to support the hypothesis that the Illinois' execution moratorium significantly increased homicides.

Keywords: capital punishment; event studies; structural breaks; homicides

JEL Classification: C22; K42

I. Introduction

Since the early 2000s, there has been a resurgence of academic studies, primarily conducted by economists, estimating the deterrent effect of capital punishment (CP) on homicide rates (e.g. Dezhbakhsh

et al., 2003; Mocan and Gittings, 2003; Shepherd, 2004; Zimmerman, 2004). With few exceptions (e.g. Hjalmarsson, 2009), these studies rely on panel models using the fixed effects or first difference estimators, span the period since the reinstatement of the death penalty in the United States following the

*Corresponding author. E-mail: aa1266@hw.ac.uk

1976 Supreme Court decision in *Gregg v. Georgia*, and operate within the OLS or IV/GMM estimation framework. Despite a large number of empirical studies, there is still disagreement regarding whether CP has a deterrent effect on homicides (Dezhbakhsh *et al.*, 2003; Mocan and Gittings, 2003; Donohue and Wolfers, 2006; Kovandzic *et al.*, 2009). A recent assessment of research by the National Research Council concluded that ‘research to date on the effect of capital punishment on homicide is not informative about whether capital punishment decreases, increases, or has no effect on homicide rates. Therefore, the committee recommends that these studies not be used to inform deliberations requiring judgments about the effect of the death penalty on homicide’ (National Research Council, 2012, p. 2). Chalfin *et al.* (2013) state that all of the recent panel studies examining the deterrent effect of the death penalty are inconclusive at best, uninformative at worst. A review by Donohue and Wolfers (2006) suggests that support for the CP deterrence hypothesis is largely confined to studies using instruments for execution risk that fail to meet the basic relevancy and validity requirements of instrumental variables (see also Kovandzic *et al.*, 2009).

We contribute to the debate by reviewing and reassessing an event study by Cloninger and Marchesini (2006, in the following referred to as CM) who exploit a natural experiment in Illinois to test the effects on homicide counts of a death penalty moratorium established by ex-Illinois Governor George Ryan. On 31 January 2000, the ex-governor declared a moratorium on executions and announced the establishment of a commission to review the death penalty system (IGNN, 2000). CM’s identification strategy, which they refer to as a ‘portfolio approach’, draws upon event studies in finance research Cloninger and Marchesini (1995, 2001). The authors examine the relationship between US and Illinois’ monthly homicide series from 1994 to 2003 and conclude that the moratorium resulted in an additional 150 homicides in Illinois in the 48-month post-intervention period. In this article, we argue that the identification strategy used by CM has no evidentiary value and is not appropriate for drawing conclusions about the effects of naturally occurring quasi-experiments on crime rates.

Instead of CM’s portfolio approach, we apply univariate time-series methods to test for

structural breaks occurring at both known and unknown dates. Specifically, we apply Chow-type tests to monthly homicide data in Illinois from 1994 to 2010 to look for a structural break beginning in February 2000. As the structural break might have occurred slightly earlier (criminals might have anticipated the moratorium) or later (due to persistence in criminal behaviour), we also consider the Quandt Likelihood Ratio test (Quandt, 1960; Andrews, 1993) which allows us to test for a structural break in an arbitrary interval around February 2000. We also conduct CUSUM-type tests for parameter stability which do not require us to specify a break date. Furthermore, we employ the synthetic control method due to Abadie and Gardeazabal (2003) and Abadie *et al.* (2010, forthcoming) to an annual US state-level panel data set covering 1977 to 2006. The synthetic control estimator approximates the counterfactual outcome by a weighted average of untreated US states. These methods provide consistent results in that we cannot find any evidence of a homicide-promoting effect of ex-Governor Ryan’s decision to halt executions or commute previously imposed death sentences.

In the following section, we outline and discuss the identification strategy employed by CM. We explain why their identification strategy cannot be used to justify claims of policy intervention effectiveness (or unintended consequences) on crime rates, never mind the impact of an execution moratorium on homicide rates. In Section III, we describe the univariate time-series procedures used to identify structural breaks at known and unknown break dates, including the results when such procedures are applied to monthly data on Illinois homicides from 1994 to 2010. Section IV considers an alternative approach, the synthetic control method, which is applied to an annual state-level panel data set. Section V concludes the article.

II. Cloninger and Marchesini (2006)

This section replicates and reassesses CM’s event study of the Illinois’ moratorium. Data on monthly homicides in Illinois from 1994 to 2003 come from the Federal Bureau of Investigation’s Uniform Crime Reporting (UCR) Program and were graciously

provided by CM to Donohue and Wolfers (2006) for reanalysis in their article.¹ CM examine the extent to which the annual homicide change rates in Illinois are correlated with annual homicide change rates for the entire United States before and after the implementation of the execution moratorium. The pre-event period is referred to as the 'control period' and is defined as the 60 months prior to the governor's declared moratorium in January 2000, that is, January 1995 to December 1999. The post-event period is referred to as the 'experimental period' and is defined as the 48 months following the execution moratorium, that is, January 2000 to December 2003. We should note that CM mistakenly assigned the entire month of January 2000 to the post-intervention period even though the moratorium officially began on 31 January 2000, that is, the moratorium was announced by Governor Ryan on the last day in January 2000.

The regression model used by CM is defined as follows:

$$R_{jt} = \alpha_j + \beta_j H_t + \mu_{jt} \quad (1)$$

where R_{jt} and H_t is the percentage change in homicide counts compared to the same month in the previous year for Illinois (i.e. state j) and for the entire United States, respectively. CM refer to R and H as 'homicide returns'. β is the slope coefficient for US 'homicide returns' and is interpreted by the authors as 'all those forces that relate changes in the number of homicides in a particular state with those of a broad national index' (p. 969). The authors maintain that any significant increase in the post-moratorium homicide β or significantly positive post-event residuals can be viewed as evidence in support of the death penalty deterrence efficacy hypothesis:

It is hypothesized that in the presence of a deterrence effect, the post-event residuals will be, in the main, significantly positive and the absolute size of the homicide beta will significantly increase. The former may be interpreted as evidence in support of the deterrence hypothesis while the latter could be similarly interpreted as well as viewed as an increase in

the systematic risk (the variation in state-specific changes in homicides associated with similar changes in national homicides) of homicide. (p. 968)

CM report a pre-moratorium national homicide β of 0.39 with an SE of 0.31. Because the pre-event β is not significantly different from zero, they conclude that changes in monthly Illinois homicide levels from 1995 to 1999 were operating independently of those occurring at the national level. On the other hand, the authors report a post-moratorium national homicide β of 0.97 with an SE of 0.44 (based on the period 2000–2003). Because the post-event homicide β is significantly different from zero, coupled with the author's statement that it is significantly different from the pre-moratorium β as determined by a Chow test, the authors conclude that Illinois homicides are now directly linked with 'homicide returns' nationally (at least from 2000 to 2003) and attribute this newly found association to the execution moratorium. CM also compare the mean of 'homicide returns' in the control and experimental period and find 'homicide returns' to be significantly higher in the post-event period, resulting in 150 additional homicides in 2000–2003 (pp. 970–1). Lastly, the authors state that 'the citizens of Illinois experienced increased systematic risk of homicide as a result of the executive actions that imposed a moratorium' (p. 972).

The methodology employed by CM is borrowed from event studies in finance. In the finance literature, Equation 1 is referred to as the single-index or market model (Fama *et al.*, 1969; see also MacKinlay, 1997; Corrado, 2011) and expresses the relationship between the (excess) return on firm j 's security, R_{jt} , and H_t is a measure of 'general market conditions' (Fama *et al.*, 1969, p. 4). The single-index model is a popular and well-established base model for event studies in finance. H_t is usually proxied by a stock market index, for example, S&P 500. The single-index model decomposes the returns of security j into a firm-specific time-invariant component (α_j), a systematic factor ($\beta_j H_t$) and a firm-specific time-varying factor (μ_{jt}). β_j is a measure of the sensitivity of stock j 's returns to general market conditions.

¹ We obtained these data from Justin Wolfers's personal website. See <http://users.nber.org/~jwolfers/deathpenalty.php> (Accessed on 4 July 2013).

While the relationship in Equation 1 has a theoretical foundation in finance, CM fail to justify the use of Equation 1 in the crime context and do not discuss the rationale for the existence of a long-run relationship between national ‘homicide returns’ and the change in homicide counts in Illinois. Indeed, the fact that β is not significant in the pre-event period indicates that the relationship between US-wide and Illinois homicide counts is, if anything, unstable and does not suffice as a basis for an event study.

Let us assume for a moment that, as claimed by CM, the introduction of the moratorium caused β to switch from zero to a significantly positive value. It is important to stress that this does not imply that the moratorium caused an increase in homicides. A positive β implies that Illinois’ ‘homicide returns’ and national ‘homicide returns’ tend to move together. Hence, whether Illinois’ homicide counts decrease or increase then depends on whether national homicide counts decrease or increase. CM argue that if β were still zero after 2000, then the number of homicides in 2000–2003 would be lower by approximately 150. This however is *conditional on* an increase of national homicide counts in the post-event period (the sample mean of H_t in the period 2000–2003 is 0.006). In an alternative scenario,

the national index could have decreased after January 2000 and homicides in Illinois would have decreased as well. In that case, CM’s approach would lead to a fundamentally different conclusion.

Although we have criticized the use of the single-market model and the interpretation of the β coefficient, we assume for the moment that Equation 1 is a legitimate model for examining the deterrence effect of the moratorium. Hence, we assume that CM’s statement, that significantly positive post-event residuals and a significantly greater post-event β provide evidence for the deterrence hypothesis, is correct. Models 1, 2 and 3 in Table 1 replicate the results in Tables 1–3 in CM.² Models 4 and 5 use the sample until 2002:12 and 2003:12, but add a moratorium dummy, D_t , which is defined such that $D_t = 1$ if $t \geq 2000:1$, 0 otherwise, as well as an interaction term, $H_t \times D_t$. Testing the null hypothesis that the pre-moratorium β is not significantly different from the post-moratorium β is equivalent to testing whether the coefficient on $H_t \times D_t$ is significantly different from zero. Analogously, testing the null that the constant is the same in the pre- and post-event period is equivalent to testing whether D_t is significant. This is referred to as the Chow test (Chow, 1960) and will be discussed

Table 1. Replication of Tables 1–3 in CM

	(1)	(2)	(3)	(4)	(5)
H_t	0.391 (1.25)	0.811 (1.63)	0.968* (2.20)	0.391 (1.08)	0.391 (1.05)
D_t				0.0443 (0.93)	0.0259 (0.56)
$H_t D_t$				0.420 (0.76)	0.577 (1.09)
Constant	−0.0402 (−1.30)	0.00409 (0.11)	−0.0143 (−0.45)	−0.0402 (−1.12)	−0.0402 (−1.09)
N	60	36	48	96	108
Box–Ljung [†]	0.0195	0.00290	0.00169	0.00477	0.00280
From	1994:1	2000:1	2000:1	1994:1	1994:1
To	1999:12	2002:12	2003:12	2002:12	2003:12

Notes: t -Statistics in parentheses.

[†] p -values are shown. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

² All calculations and estimations, if not mentioned otherwise, were performed in *Stata 12.0*.

Table 2. Abnormal returns corresponding to 2000:1–2000:12

Date	Abnormal returns	<i>t</i> -Statistic
2000:1	−0.117	(−0.76)
2000:2	0.0629	(0.41)
2000:3	0.0550	(0.36)
2000:4	−0.281	(−1.83)
2000:5	−0.173	(−1.11)
2000:6	−0.243	(−1.47)
2000:7	0.450**	(2.86)
2000:8	−0.0583	(−0.37)
2000:9	0.182	(1.17)
2000:10	0.240	(1.49)
2000:11	0.0854	(0.56)
2000:12	0.0544	(0.35)

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3. Summary statistics for the Illinois' monthly homicide series, 1994–2010

	Obs.	Mean	SD	Min	Max
Full sample	204	59.14	16.55	22	104
$t < 2000:2$	73	69.78	15.20	43	104
$t \geq 2000:2$	130	53.21	14.17	22	84

in more detail in Section III. Both the post-moratorium dummy and the interaction term are not significant (when using the sample until 2002:12 and until 2003:12), indicating that β is not significantly different after the event. This is in contrast to CM's statement that 'the statistical difference between pre- and post-event β s, as measured by a standard *t*-test of means, is confirmed by the Chow test' (p. 970). As shown in the table, all models suffer from residual correlation as indicated by the Box–Ljung statistic, implying that the models are misspecified.

Next, we consider the post-event residuals, in finance also referred to as abnormal returns. The abnormal returns are the difference between the expected returns and the actual returns where the expected returns are the predicted values obtained using the pre-event model. The abnormal returns corresponding to the period 2000:1 to 2000:12 are shown in Table 2. Only the abnormal returns corresponding to 2000:4 and 2000:7 are statistically different from zero at the 10% level. The former,

however, has a negative sign. We cannot reject that the sum of the abnormal returns over the period 2000:1 to 2000:12 is equal to zero (p -value is 0.70). Therefore, even under the assumption that Equation 1 is an appropriate model for examining the impact of the moratorium, there is no evidence that the moratorium led to an increase in 'homicide returns'.

CM conduct a comparison-of-means analysis where they compare the mean of Illinois' 'homicide returns' in the pre- and post-event period. The sample mean equals −0.702 and 0.006 in the pre- and post-event period, respectively. CM conclude, based on the finding that the mean of 'homicide returns' in the post-event period is significantly greater, that due to the moratorium 150 additional homicides occurred. Interestingly, the mean of Illinois' 'homicide returns' in 2000, which is the first year after the moratorium, is negative and not statistically different from the pre-event mean ($t = 1.07$). Clearly, we would expect to find significantly higher and positive 'homicide returns' in 2000 if the moratorium had a homicide-promoting effect. However, we maintain that a simple comparison-of-means analysis in a noncontrolled experiment is not appropriate for making causal inferences. A comparison-of-means analysis does not capture the overtime dynamics of the series and is likely to suffer from omitted variable bias.

III. Re-Examining Illinois' Monthly Homicide Counts

For our analysis, we update the monthly UCR homicide series through 2010 using a data query tool available through the National Consortium on Violence Research's data center.³ The homicide series is displayed in Fig. 1 and descriptive statistics for the full, the pre- and post-moratorium period are shown in Table 3. The introduction of the moratorium is indicated by the vertical dashed line in Fig. 1. In our analysis, we treat February 2000 as the first full month in which the execution moratorium was in effect.

Donohue and Wolfers (2006) critically discuss CM's use of year-on-year growth in homicides instead of using homicides in levels, arguing that the interest lies in homicide counts. In time-series analysis, it is common practice to apply first or

³ The data query tool is available at <http://www.ncovr.heinz.cmu.edu/Docs/datacenter.htm> (Accessed on 4 July 2013). The updated data set is available at the link provided in the supplementary data section of this paper.

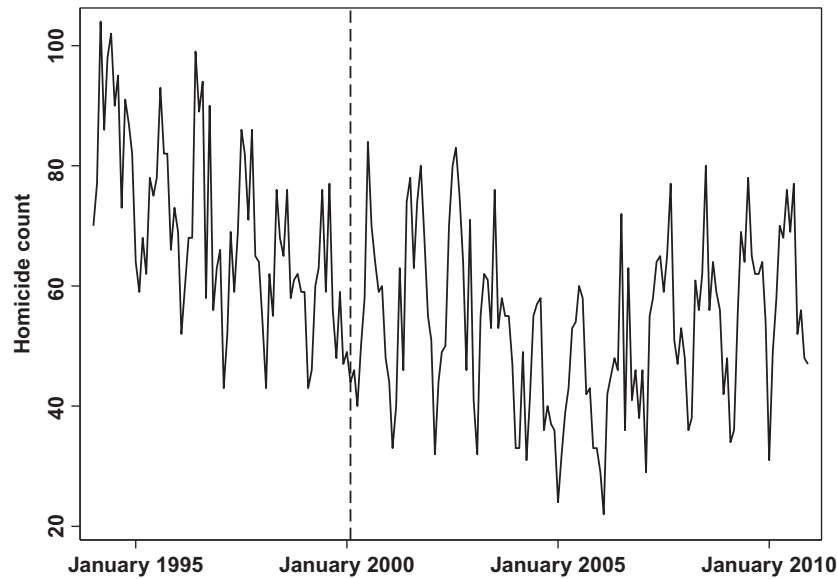


Fig. 1. Monthly homicide series in Illinois, 1994–2010

seasonal differencing if a series contains unit roots. Therefore, we formally test for the presence of a unit root in the Illinois' monthly homicide series using the widely popular ADF test (Dickey and Fuller, 1979). We extend the classical ADF regression by seasonal dummy variables due to strong seasonality (see Fig. 1), that is,

$$\Delta y_t = a_0 + \sum_{j=1}^{11} \alpha_j M_{jt} + \gamma y_{t-1} + a_2 t + \sum_{i=1}^p \theta_i \Delta y_{t-i} + \varepsilon_t$$

The regression model includes a deterministic drift (a_0), a deterministic seasonal component ($\sum_j \alpha_j M_{jt}$), a deterministic trend term ($a_2 t$) and is augmented by lagged first differences of the

series to account for serial correlation. To not affect inference, the monthly dummy variables are defined such that $M_{jt} = 1/12$, if t corresponds to month j , $M_{jt} = -1/12$ otherwise (Enders, 2009). Under the null hypothesis, $\gamma = 0$ and the series contains a unit root. If we reject the null hypothesis in favour of $\gamma < 0$, then we conclude that the series is stationary and conduct hypothesis tests using standard asymptotic distribution theory. The ADF test has a non-standard test statistic distribution that depends on the sample size and the model specification (with/without deterministic drift and trend). Since we want to test for a structural break, it should be mentioned that the ADF is biased towards the non-rejection of the null hypothesis if structural breaks are present (Perron, 1989).

Table 4 presents the results of the ADF tests applied to the entire monthly homicide series covering the period 1994 to 2010 using the ADF specification

Table 4. ADF test results for unit roots in the Illinois' monthly homicide series, 1994–2010

Test	Version	Statistic	Critical values		
			1%	5%	10%
ADF (2 lags)	Drift, time-trend	-4.15	-4.04	-3.45	-3.15
Significance test	Time-trend	9.10	8.73	6.49	5.47

Note: Critical values from Enders (2009, pp. 488–9) for $T = 100$.

with deterministic drift and trend.⁴ Based on the critical values in the right-most columns of the table, the test statistic of -4.15 does reject the null hypothesis that the homicide series has a unit root at the 1% level. Also, as shown in the table, we conclude that the deterministic trend is significant at the 1% level. We also considered the seasonal unit root tests proposed by Osborn *et al.* (1988) and Canova and Hansen (1995). Both tests indicate that the series has no seasonal unit roots.⁵

Testing for structural break in February 2000

To detect whether the death penalty moratorium had a significant impact on homicide counts in Illinois, we estimate the monthly homicide series as an autoregressive process. Consider for simplicity the two-regime AR(1) model

$$y_t = \begin{cases} \delta_0 + \delta_1 y_{t-1} + \epsilon_t & t < 2000:2 \\ \pi_0 + \pi_1 y_{t-1} + \epsilon_t & t \geq 2000:2 \end{cases}$$

where y_t is the monthly homicide counts in Illinois in levels. The two-regime model allows the drift parameter as well as the autoregressive parameter to be different after the putative structural break. If $\delta_0 = \pi_0$ and $\delta_1 = \pi_1$, the introduction of the moratorium had no impact on the data-generating process governing homicide counts over time. We can define the intervention dummy variable D_t which equals 1 if $t \geq 2000:2$, $D_t = 0$ otherwise, and rewrite the two-regime model in one equation as

$$y_t = \delta_0 + \phi_0 D_t + \delta_1 y_{t-1} + \phi_1 D_t y_{t-1} + \epsilon_t$$

where we define $\phi_0 \equiv (\pi_0 - \delta_0)$ and $\phi_1 \equiv (\pi_1 - \delta_1)$. We can test the null hypothesis of no structural break at time $t = 2000:2$, that is, $\phi_0 = \phi_1 = 0$, using a standard F -test. This test is often referred to as the Chow test for a structural break at a known date (Chow, 1960).

For our analysis, we consider the more general model

$$y_t = \delta_0 + \sum_{i=1}^q \delta_i y_{t-i} + \phi_0 D_t + \sum_{i=1}^q \phi_i y_{t-i} D_t + \psi P_t + \sum_{j=1}^{11} \mu_j M_{jt} + \epsilon_t \quad (2)$$

We allow for higher-order lags of y_t on the right-hand side as well as a deterministic seasonal component in order to capture seasonal fluctuations. M_{1t} to M_{11t} are the monthly dummy variables defined above. In addition, we define a pulse dummy variable, P_t , which equals 1 if $t = 2000:2$, 0 otherwise. The rationale for including a pulse dummy variable is that the moratorium might have induced a temporary (1 month) change in homicide levels in Illinois. Under the null hypothesis of no structural break, $\phi_i = 0$ for $i = 0, 1, \dots, q$ and $\psi = 0$. We do not test whether the coefficients on the monthly dummies changed after the break as changes in the seasonal pattern are not relevant to whether the event had a significant impact on the general level of homicide counts.

In order to select the lag order in Equation 2, we consider the model under the alternative as well as under the null hypothesis. In both cases, $q = 3$ achieves the lowest Bayesian information criterion (BIC) score which indicates best fit. After checking the Box–Ljung statistic, the residual correlogram as well as the White–Koenker test for homoscedasticity, we conclude that the residuals do not exhibit any systematic pattern and are homoscedastic. Estimation results are shown in Table 5 and results for different flavours of the structural break test can be found at the bottom of the table. Model 5 corresponds to Equation 2. All coefficients involving D_t and P_t are insignificant and the joint test does not allow us to reject the null hypothesis of no structural break after February 2000. The results do not change if we exclude the pulse dummy variable (model 4) or impose $\phi_i = 0$ for all i (model 3).

⁴ To select p , that is, the number of lags of Δy_t on the right-hand side, we consider the BIC as well as the general-to-specific approach, where the lag length is successively reduced using t - and F -tests. Both BIC and the general-to-specific methodology indicate that $p = 2$ provides best fit. Diagnostic checks suggest that the model is adequate in that the Box–Ljung statistics for residual serial correlation indicate no serial correlation and the White–Koenker test indicates homoscedasticity. The White–Koenker test was performed using *ivhetttest* by Schaffer (2010).

⁵ These tests were carried out using the *R* packages *forecast* and *urca* (Hyndman and Khandakar, 2008; Pfaff, 2008; Hyndman, 2014).

Table 5. Results of tests for structural break at February, 2000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
y_{t-1}	0.156* (2.27)	0.126 (1.81)	0.154* (2.25)	0.0888 (0.88)	0.0874 (0.87)	0.109 (1.53)	0.131 (1.85)	0.0827 (0.82)	0.0813 (0.80)
y_{t-2}	0.391*** (6.20)	0.353*** (5.42)	0.395*** (6.23)	0.403*** (4.03)	0.403*** (4.03)	0.334*** (4.97)	0.369*** (5.61)	0.389*** (3.83)	0.389*** (3.83)
y_{t-3}	0.253*** (3.75)	0.217** (3.16)	0.251*** (3.72)	0.265** (2.64)	0.264** (2.63)	0.227** (3.23)	0.260*** (3.80)	0.258* (2.51)	0.256* (2.50)
D_t		-3.884* (-2.18)		2.999 (0.42)	2.969 (0.41)	-3.490 (-1.90)		2.812 (0.38)	2.692 (0.37)
P_t			7.934 (0.85)		9.149 (0.99)		6.758 (0.72)		8.180 (0.87)
$y_{t-1}D_t$				0.0510 (0.44)	0.0496 (0.43)			0.0372 (0.32)	0.0368 (0.31)
$y_{t-2}D_t$				-0.0874 (-0.68)	-0.0829 (-0.65)			-0.0916 (-0.71)	-0.0874 (-0.68)
$y_{t-3}D_t$				-0.0728 (-0.62)	-0.0770 (-0.66)			-0.0457 (-0.38)	-0.0497 (-0.41)
Indiana						0.160 (1.55)	0.191 (1.86)	0.152 (1.46)	0.143 (1.36)
Missouri						0.0506 (0.57)	0.0437 (0.49)	0.0369 (0.41)	0.0409 (0.45)
Iowa						0.0737 (0.24)	0.0555 (0.18)	0.0585 (0.19)	0.0712 (0.23)
$H_0 : \phi_0 = 0$		0.0306		0.676	0.679	0.0586		0.702	0.714
$H_0 : \phi_i = 0$ for all i				0.180	0.165			0.310	0.286
$H_0 : \phi_i = \psi = 0$ for all i		0.0306	0.396	0.180	0.204	0.0586	0.472	0.310	0.354
White-Koenker	0.540	0.415	0.576	0.443	0.475	0.314	0.607	0.387	0.423
Box-Ljung	0.779	0.722	0.814	0.726	0.760	0.796	0.837	0.800	0.815

Notes: t -Statistics in parentheses.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.Results for seasonal dummies and constant are omitted. Bottom five rows show p -values. Box-Ljung test is based on 12 lags.

Model 2 includes the level dummy variable only. The coefficient estimate for the moratorium level dummy is statistically different from zero at the 5% level, but the coefficient has an unexpected negative sign. Lastly, we extended the previous model specifications by adding homicides in neighbouring states⁶ to account for the potential suppression of any ‘homicide promoting’ effects of the execution moratorium due to omitted variable bias. If, for example, other historical events occurring in the midwest region of the United States in or around the timing of the moratorium were responsible for significant reductions in homicide levels, then failing to control for such confounding factors could suppress any deleterious effects of the execution moratorium on homicide. The results of this specification are reported in the last four columns of Table 5. In sum, there is no evidence of a temporary or permanent break in the data-generating process of homicides in February 2000 when ex-Governor Ryan declared the nation’s first moratorium on state executions. Contrary to the findings reported by CM, we find no evidence that the implementation of an execution moratorium had a permanent effect on the levels of homicides in Illinois. While not shown, the results remain largely unchanged when the Chow test is carried out using first differences of homicide counts. This finding also provides some additional evidence that homicide counts are stationary in levels.

Testing for structural break at an unknown date

The analysis so far has assumed that the putative structural break occurred exactly with the introduction of the moratorium. We tested the null of no structural break against the alternative of a structural break at $t = 2000:2$. However, the structural break might have occurred earlier (criminals might have anticipated the moratorium) or later (due to persistence in criminal behaviour). It seems natural to test the null hypothesis

of no structural break against alternative break points around 2000:2. We consider the interval [1999:2, 2001:2]. That is, we allow the break point to be up to 12 months before or after the introduction of the moratorium. Let F_j be the statistic obtained under the null hypothesis $\phi_i = 0$ for $i = 0, 1, \dots, 3$ where the break point is at $t = j$ under the alternative and j is in the interval [1999:2, 2001:2]. The asymptotic distribution of the supremum of the set of F -statistics $\{F_{1999:2}, \dots, F_{2001:2}\}$ is considered by Andrews (1993). Critical values depend on the number of restrictions and the size of the chosen interval and are tabulated in Andrews (2003).⁷ We cannot reject the null of no structural break as the highest F -statistic (2.29) is considerably smaller than the 10% critical value (10.41). Although the alternative hypothesis is an abrupt parameter change, the above test is shown to have power against more complex structural changes, such as changes involving a transition period (Andrews, 1993).

Next, we consider the CUSUM test which allows exploring parameter stability and identifying structural breaks without requiring us to specify a break date (Brown *et al.*, 1975). The approach is based on the cumulative sum of scaled one-step-ahead prediction errors. Figure 2 depicts the CUSUM series and the corresponding 95% confidence bands. A significant departure of the series from the zero-mean line indicates parameter instability. Thus, there is no evidence of parameter instability in the series until the end of 2004. The CUSUM-style test proposed by Ploberger and Krämer (1992) is based on the cumulative sum of OLS residuals and provides no evidence of a structural change either.⁸

IV. Synthetic Control Method

The univariate time-series methods discussed in the previous section suggest that there is no evidence of a

⁶ Homicide counts for Kentucky and Wisconsin could not be included as controls due to missing data points during the time period studied.

⁷ More specifically, critical values depend on $\lambda = \pi_2(1 - \pi_1)/(\pi_1(1 - \pi_2))$ where $\pi_1 = t_1/T$ and $\pi_2 = t_2/T$. T is the number of observations and t_1 and t_2 is the time index corresponding to the start and the end of the interval, respectively. In our case, $T = 204$ and 1999:2 corresponds to $t_1 = 62$ and 2000:2 corresponds to $t_2 = 86$. Therefore, $\lambda = 1.67$. Since we test $\phi_i = 0$ for $i = 0, 1, \dots, 3$, the number of restrictions is 4. The closest critical value at the 10% level tabulated in Andrews (2003) is for $\lambda = 1.49$ and is equal to 10.41.

⁸ The CUSUM tests were performed using the *strucchange* package by Zeileis (2006) in *R 2.15.1* and the *cusum6* command by Baum (2000) in *Stata 12.0*.

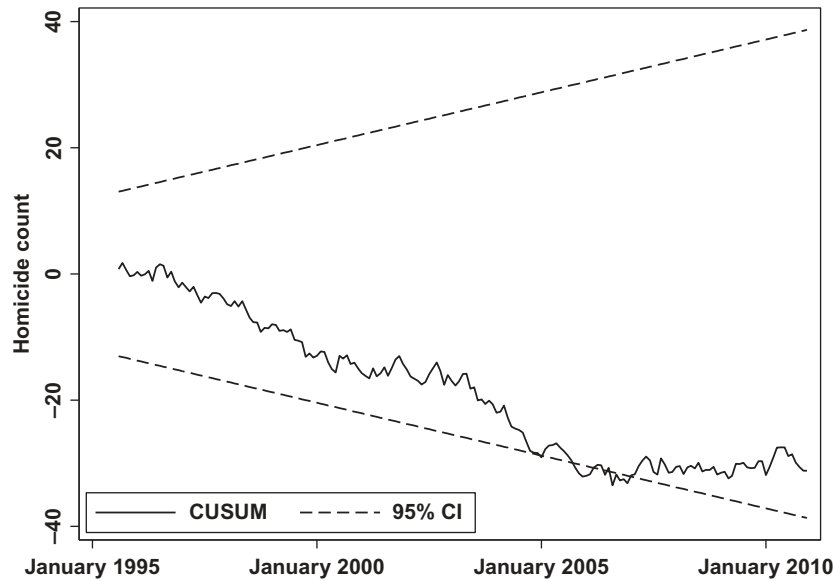


Fig. 2. CUSUM of Illinois' monthly homicide series

structural break or parameter instability, indicating that the introduction of the moratorium had no significant impact on homicides in Illinois. However, while we have included homicides in neighbouring states as control variables, there remain doubts as to whether the univariate model adequately controls for confounding factors which might have countervailed the impact of the moratorium. For this reason, we consider an alternative estimator: the synthetic control estimator due to Abadie and Gardeazabal (2003) and Abadie *et al.* (2010, forthcoming).⁹

Let y_{it} denote the observed Illinois' homicide count and y_{it}^N denote the unobserved homicide count if the moratorium had not been introduced. The effect of the Illinois' moratorium on homicide counts is equal to $y_{it} - y_{it}^N$. The synthetic control estimator approximates the counterfactual outcome, y_{it}^N , as a weighted average of outcomes in donor pool units. That is, the synthetic control is generated as $\sum_{j \neq i} \omega_j y_{jt}$ (subject to $\sum_{j \neq i} \omega_j = 1$ and $\omega_j \geq 0$). In our application, the donor pool includes all US states for which CP was legal throughout the sample period (i.e. following the *Gregg v. Georgia* decision and until 2006).¹⁰ In contrast to difference-in-differences

methods, the synthetic control estimator does not require specifying the control group, but infers the optimal control group from the data. Intuitively, the synthetic control weights are chosen such that states which exhibit a relatively high degree of similarity to Illinois prior to the event receive large weights. The synthetic control estimator requires an appropriate number of predictor variables for the treated unit and the donor pool units in order to assess the similarity across states. Since data for most known homicide predictors is not available at a monthly or quarterly frequency, we exploit an annual state-level US panel data set that covers 1977–2006 and contains an extensive list of predictor variables including socioeconomic, demographic, policy and deterrence control variables such as the number of executions. The data is taken from Kovandzic *et al.* (2009), who provide an extensive discussion of data sources. This data set is available at the link provided in the supplementary data section of this paper. A list of control variables and summary statistics is provided in Table 6.

The synthetic control estimation includes all variables listed in Table 6, as well as three lags of the homicide rate (1980, 1990 and 1999). Figure 3 shows

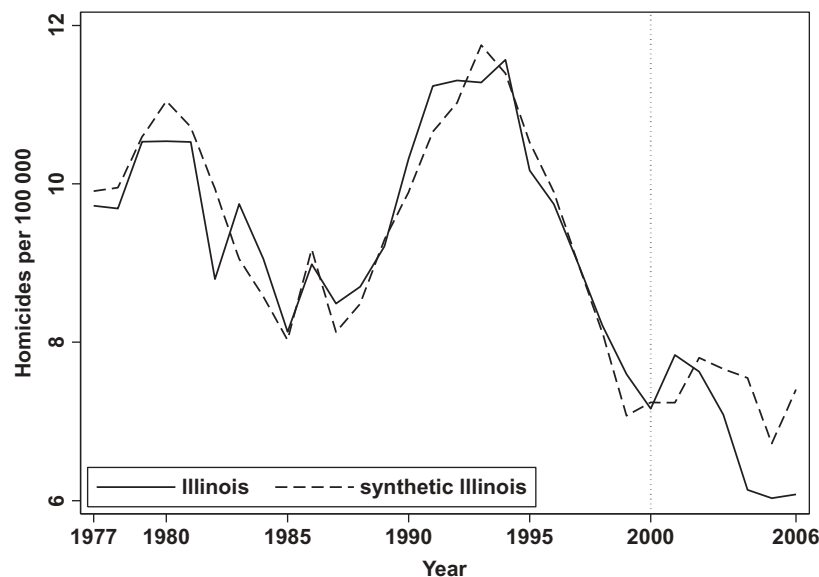
⁹ We are grateful to an anonymous referee who suggested implementing the synthetic control method.

¹⁰ The donor pool includes Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Idaho, Indiana, Kentucky, Louisiana, Mississippi, Missouri, Montana, Nebraska, Nevada, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, Washington and Wyoming (32 states).

Table 6. Descriptive statistics of state-level data set

	Mean	SD
<i>Dependent variable</i>		
Homicides per 100 000	7.161	3.465
<i>Death penalty</i>		
Death penalty dummy variable	0.989	0.109
Executions	1.060	3.396
<i>Policy control variable</i>		
Right-to-carry concealed law dummy variable	0.280	0.444
3X law dummy variable	0.219	0.410
<i>Sociodemographic and economic controls</i>		
Unemployment rate (in %)	5.962	1.897
Employment (in % of total population)	55.025	5.643
Poverty rate	13.779	4.023
Real per capita income	4.621	0.863
Total population	5 662 143	6 022 861
Per cent residents age 15–24	15.737	2.085
Per cent residents age 25–34	15.565	1.842
Per cent residents age 35–44	14.124	1.937
Per cent persons with college degree	20.187	5.173
Per cent persons residing in metropolitan areas	66.451	19.912
Beer consumption (31-gallon barrels per 100 000)	74 774.2	13 170.7
<i>Deterrence controls</i>		
Prisoners per 100 000	291.580	154.028
Police officers per 100 000	265.296	42.578

Notes: Data set covers the time period 1977–2006. Descriptive statistics are for Illinois and 32 US states in the donor pool.

**Fig. 3. Annual homicide rate in Illinois and synthetic control, 1977–2006**

the realized homicide series and the synthetic control for Illinois. The synthetic control method assigns positive weights to Arizona (0.018), California (0.122), Colorado (0.246), Connecticut (0.017),

Indiana (0.239), Louisiana (0.371) and Pennsylvania (0.042). The difference between observed homicides and synthetic control is depicted in Fig. 4 in black. The difference is positive in 2001,

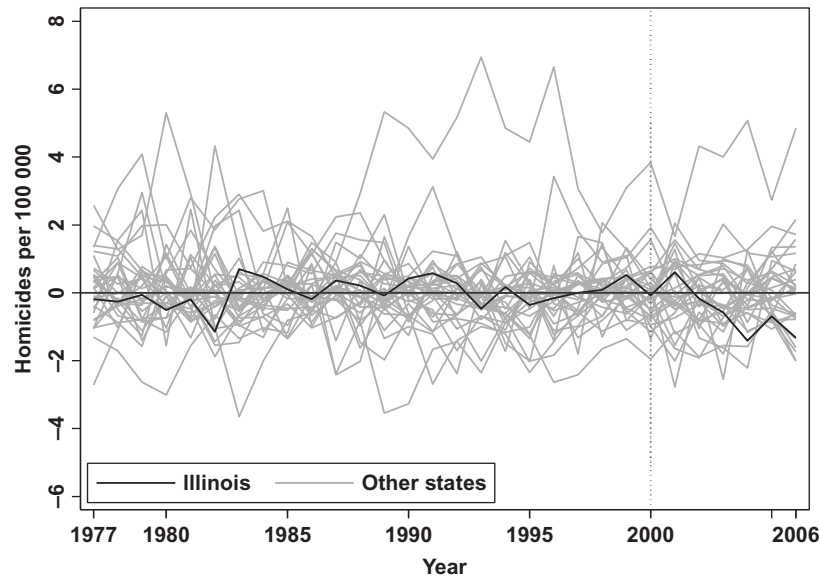


Fig. 4. Prediction error for Illinois and other US states

which would suggest a positive treatment effect, but drops to below zero in 2002. Also, at first glance, the spike in 2001 does not seem unusually large relative to the prediction errors before 2000.

To adequately assess whether the prediction errors deviate from zero by chance or whether the deviations reflect a significant impact of the moratorium on homicide outcomes, we follow Abadie *et al.* (2010) and apply placebo experiments across states. For each US state in the donor pool, we obtain the synthetic control series,

treating the state as if a moratorium had been introduced in 2000. Post-event prediction errors that are extreme relative to other states would indicate a significant impact of the moratorium. Figure 4 shows the difference between observed homicides and synthetic controls for all states from the donor pool in grey. As can be seen from the figure, the fit varies substantially across states and, hence, not all states provide information about the distribution of Illinois' prediction errors. For this reason, Fig. 5 restricts the set of

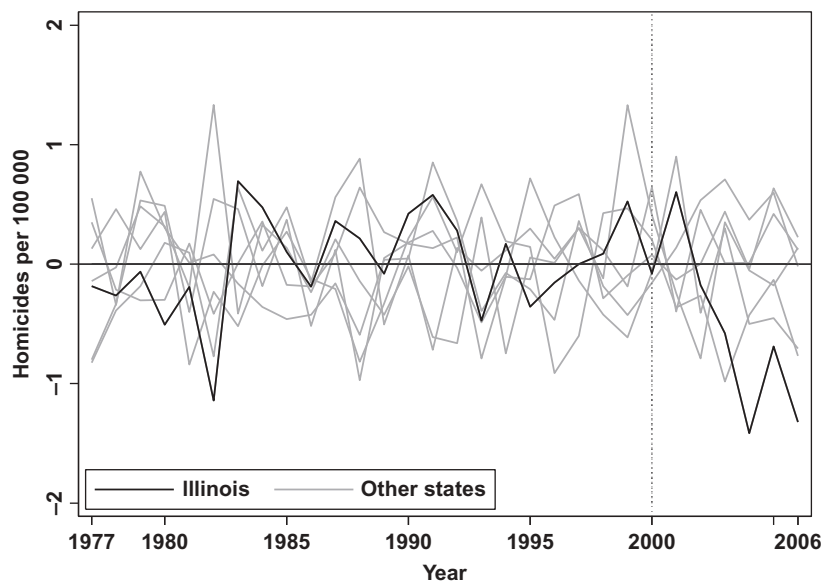


Fig. 5. Prediction error for Illinois and reduced reference group

comparison states to those states for which the pre-intervention mean squared prediction error (PPE) is lower than the Illinois' PPE.¹¹ The cut-off excludes 25 out of 32 states from the reference group. Note that by reducing the reference group to states that have a better fit than Illinois we lower the threshold for detecting significant treatment effects. The comparison across states shows that the spike in 2001 is not unusually extreme and also reveals the relatively large negative deviations in 2003–2006.

The synthetic control estimator relies on the assumption that homicide outcomes in donor states are not influenced by the Illinois' moratorium. We argue that this assumption is reasonable in this context as the Illinois moratorium does not directly affect the deterrent potential of the CP in other states. However, to empirically check this assumption and to verify the robustness of the above results, we iteratively run the synthetic control estimation, but in each step exclude one of the states with positive weights. Figure 6 shows the difference prediction errors for the unrestricted synthetic control estimation (in black) and the leave-one-out estimation results (in grey). Clearly, the fit deteriorates when one state is excluded, but

the overall pattern remains the same. We conclude that, consistent with the univariate methods in the previous section, the synthetic control method does not provide any evidence for an increase in homicides caused by the introduction of the moratorium. In fact, the only unusual deviations appear after 2002 and are negative.

V. Conclusion

We have argued that CM's identification strategy, although well-established in finance, does not have a theoretical justification in the study of homicides. While returns on securities move together due to market forces, state-level and national homicide counts are potentially unrelated. Thus, the relationship between state-level and national homicide is not adequate as the basis for an event study. More importantly, there is no reason to believe that national homicide counts have a *causal* effect on state-level homicide. Furthermore, CM's central assumption that 'any significant increase in beta would be consistent with the deterrence hypothesis' is again without theoretical justification (p. 969).

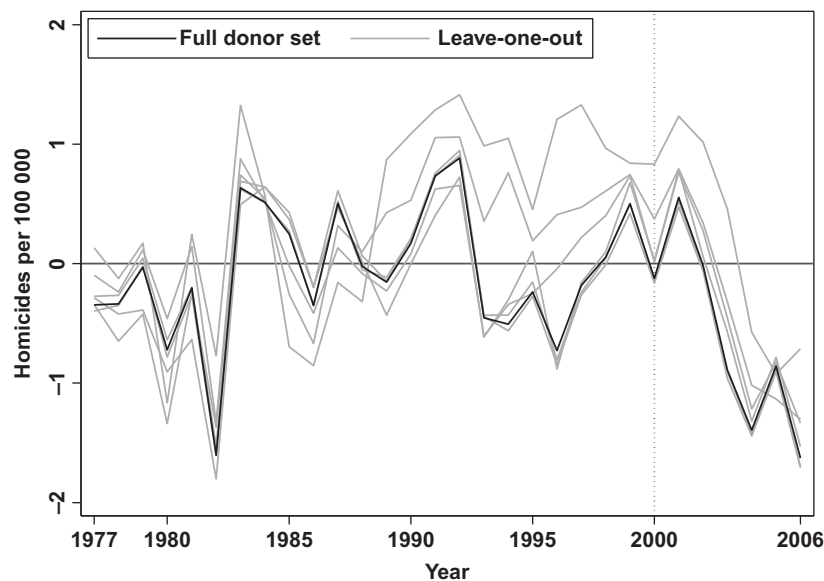


Fig. 6. Prediction errors from leave-one-out estimations

¹¹ The pre-intervention mean squared error for state i is defined as $1/T_0 \sum_{t=1}^{T_0} (y_{it} - \sum_{j \neq i} \hat{\omega}_j y_{jt})^2$ where $\hat{\omega}_j$ are the estimated synthetic weights and T_0 corresponds to the introduction of the moratorium in 2000.

A positive significant beta merely suggests that Illinois' and national homicide counts are positively correlated. CM's central finding that the moratorium and the subsequent commutation lead to 150 additional homicides over the 4-year post-moratorium period is therefore misleading. It relies on the assumption that national homicides have a causal impact on Illinois' homicides and is conditional on an increase in national homicides in the post-moratorium period. If national homicides had decreased over the post-moratorium period, a positive significant beta would have implied a decrease in Illinois' homicides, leading to a fundamentally different conclusion. Beyond the theoretical discussion, we have shown that, in fact, the pre- and post-event betas are not significantly different and that CM's portfolio model is misspecified as indicated by serial correlation in the residuals.

We apply well-known methods of univariate time-series econometrics to Illinois' monthly homicide series in order to test for a structural break at the introduction of the moratorium in February 2000. We also allow for an unknown break date to account for the fact that the structural break might have occurred slightly earlier or later. Furthermore, we exploit a state-level panel data set and, inspired by comparative case studies, implement the synthetic control estimator which compares the observed Illinois' homicide series to an optimal control unit generated from other US states.

Across a variety of different time-series methods and using two distinct data sets, we find no empirical evidence to support the hypothesis that ex-Governor Ryan's decision to halt executions in February 2000 significantly increased homicides.

Acknowledgement

The authors would like to thank Mark Schaffer, Justin Wolfers and two anonymous referees for helpful comments.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Supplemental data

Supplemental data for this article can be accessed here. Data were obtained from the web-based Data Center of the National Consortium on Violence Research (NCOVR) at Carnegie Mellon University. <http://www.ncovr.org/>.

The NCOVR Data Center was developed with support by the National Science Foundation under Grant No. SES-0215551 and SBR-9513040.

References

- Abadie, A., Diamond, A. and Hainmueller, J. (2010) Synthetic control methods for comparative case studies: estimating the effect of California's Tobacco Control Program, *Journal of the American Statistical Association*, **105**, 493–505. doi:10.1198/jasa.2009.ap08746
- Abadie, A., Diamond, A. and Hainmueller, J. (forthcoming) Comparative politics and the synthetic control method, *American Journal of Political Science*, doi:10.1111/ajps.12116
- Abadie, A. and Gardeazabal, J. (2003) The economic costs of conflict: a case study of the Basque Country, *The American Economic Review*, **93**, 113–32. doi:10.1257/00028280321455188
- Andrews, D. W. K. (1993) Tests for parameter instability and structural change with unknown change point, *Econometrica*, **61**, 821–56. doi:10.2307/2951764
- Andrews, D. W. K. (2003) Tests for parameter instability and structural change with unknown change point: a corrigendum, *Econometrica*, **71**, 395–7. doi:10.1111/1468-0262.00405
- Baum, C. F. (2000) *CUSUM6: Stata Module to Compute Cusum, Cusum² Stability Tests. Statistical Software Components S408601*, Boston College Department of Economics, Boston, MA.
- Brown, R. L., Durbin, J. and Evans, J. M. (1975) Techniques for testing the constancy of regression relationships over time, *Journal of the Royal Statistical Society. Series B (Methodological)*, **37**, 149–92.
- Canova, F. and Hansen, B. E. (1995) Are seasonal patterns constant over time? A test for seasonal stability, *Journal of Business and Economic Statistics*, **13**, 237–52.
- Chalfin, A., Haviland, A. M. and Raphael, S. (2013) What do panel studies tell us about a deterrent effect of capital punishment? A critique of the literature, *Journal of Quantitative Criminology*, **29**, 5–43. doi:10.1007/s10940-012-9168-8
- Chow, G. C. (1960) Tests of equality between sets of coefficients in two linear regressions, *Econometrica*, **28**, 591–605. doi:10.2307/1910133
- Cloninger, D. O. and Marchesini, R. (1995) Crime betas: a portfolio measure of criminal activity, *Social Science Quarterly*, **76**, 643–7.

- Cloninger, D. O. and Marchesini, R. (2001) Execution and deterrence: a quasi-controlled group experiment, *Applied Economics*, **33**, 569–76. doi:10.1080/00036840122871
- Cloninger, D. O. and Marchesini, R. (2006) Execution moratoriums, commutations and deterrence: the case of Illinois, *Applied Economics*, **38**, 967–73. doi:10.1080/00036840500462020
- Corrado, C. J. (2011) Event studies: a methodology review, *Accounting and Finance*, **51**, 207–34. doi:10.1111/j.1467-629X.2010.00375.x
- Dezhbakhsh, H., Rubin, P. H. and Shepherd, J. M. (2003) Does capital punishment have a deterrent effect? New evidence from postmoratorium panel data, *American Law and Economics Review*, **5**, 344–76. doi:10.1093/aler/ahg021
- Dickey, D. A. and Fuller, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427–31.
- Donohue, J. J. and Wolfers, J. (2006) Uses and abuses of empirical evidence in the death penalty debate, *Stanford Law Review*, **58**, 791–845.
- Enders, W. (2009) *Applied Econometric Time Series*, 3rd edn, Wiley (Wiley Series in Probability and Statistics), Hoboken, NJ.
- Fama, E. F., Fisher, L., Jensen, M. C. *et al.* (1969) The adjustment of stock prices to new information, *International Economic Review*, **10**, 1–21. doi:10.2307/2525569
- Hjalmarsson, R. (2009) Does capital punishment have a “local” deterrent effect on homicides?, *American Law and Economics Review*, **11**, 310–34. doi:10.1093/aler/ahn004
- Hyndman, R. J. (2014) Forecast: forecasting functions for time series and linear models, R package version 5.1. Available at <http://CRAN.R-project.org/package=forecast> (accessed 24 March 2014).
- Hyndman, R. J. and Khandakar, Y. (2008) Automatic time series forecasting: the forecast package for R, *Journal of Statistical Software*, **27**, 1–22.
- IGNN. (2000) Governor Ryan declares moratorium on executions, will appoint commission to review capital punishment system. Governor’s Office Press Release. Available at <http://www3.illinois.gov/PressReleases/showpressrelease.cfm?subjectid=3&recnum=359> (accessed 12 December 2013).
- Kovandzic, T. V., Vieraitis, L. M. and Boots, D. P. (2009) Does the death penalty save lives?, *Criminology and Public Policy*, **8**, 803–43. doi:10.1111/j.1745-9133.2009.00596.x
- MacKinlay, A. C. (1997) Event studies in economics and finance, *Journal of Economic Literature*, **35**, 13–39.
- Mocan, N. H. and Gittings, R. K. (2003) Getting off death row: commuted sentences and the deterrent effect of capital punishment, *The Journal of Law and Economics*, **46**, 453–78. doi:10.1086/382603
- Nagin, D. S. (2012) Scientific evidence and public policy, *The Criminologist*, **37**, 1–6.
- National Research Council. (2012) Deterrence and the death penalty, in *Committee on Deterrence and the Death Penalty*, Nagin, D. S. and Pepper, J. V. (Eds), The National Academies Press, Washington, DC.
- Osborn, D. R., Chui, A. P. L., Smith, J. *et al.* (1988) Seasonality and the order of integration for consumption, *Oxford Bulletin of Economics and Statistics*, **50**, 361–77. doi:10.1111/j.1468-0084.1988.mp50004002.x
- Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, **57**, 1361–401. doi:10.2307/1913712
- Pfaff, B. (2008) *Analysis of Integrated and Cointegrated Time Series with R*, 2nd edn, Springer, New York.
- Ploberger, W. and Krämer, W. (1992) The CUSUM test with OLS residuals, *Econometrica*, **60**, 271–85. doi:10.2307/2951597
- Quandt, R. E. (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes, *Journal of the American Statistical Association*, **55**, 324–30. doi:10.1080/01621459.1960.10482067
- Schaffer, M. E. (2010) *IVHETTEST: Stata Module to Perform Pagan-Hall and Related Heteroskedasticity Tests after IV*, Statistical Software Components, S428801, Boston College Department of Economics, Boston, MA.
- Shepherd, J. M. (2004) Murders of passion, execution delays, and the deterrence of capital punishment, *The Journal of Legal Studies*, **33**, 283–321. doi:10.1086/421571
- Zeileis, A. (2006) Implementing a class of structural change tests: an econometric computing approach, *Computational Statistics and Data Analysis*, **50**, 2987–3008. doi:10.1016/j.csda.2005.07.001
- Zimmerman, P. R. (2004) State executions, deterrence, and the incidence of murder, *Journal of Applied Economics*, **7**, 163–93.