Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment

By JUSTIN MCCRARY*

In an influential paper in the June 1997 *American Economic Review,* Steven Levitt argues that there is an electoral cycle in police hiring, with faster hiring in election years and slower hiring in other years. He then uses elections as an instrument for police hiring to estimate the causal effect of police on crime. This comment points out that a weighting error in Levitt's estimation procedure led to incorrect inferences for the key results of the paper.

Levitt presents a series of regression models explaining changes in crime rates in different cities over time, including ordinary least squares (OLS) and two-stage least squares (2SLS) specifications. He draws two main conclusions. First, police substantially reduce violent crime, but have a smaller effect on property crime. Second, 2SLS estimates are consistently more negative than OLS estimates.

Levitt's 2SLS results for violent crime are driven by a large, apparently precise estimate of the effect of police on murder. This is surprising since among the seven categories of crime considered, murder exhibits the *greatest* year-to-year variability. It turns out that the precision of the murder estimate is due to a weighting error. The weighting procedure was designed to give relatively more weight to crimes with lower year-to-year variability. However, an error in Levitt's computer program accomplished exactly the opposite, giving highly variable crimes the most weight in the estimation, and severely biasing all standard errors. To demonstrate the substantive implications of this error, I present replication estimates that use the correct (and intended) weighting scheme.

When weights are employed correctly, the data support neither of Levitt's main conclusions. First, correctly weighted 2SLS estimates show no significant effect of police on *any* of the crime categories under consideration. Pooled 2SLS estimates for violent crime (the estimates that Levitt emphasized in his discussion and that are cited in the literature) are half the published magnitude and statistically indistinguishable from zero. Pooled 2SLS property crime estimates, while more precise when correctly weighted than when not, are also indistinct from zero. Second, 2SLS estimates are sometimes more negative and sometimes more positive than the OLS estimates, and the two are never statistically distinguishable when correctly weighted.

The weighting error arose in the attempt to gain efficiency. Since covariates and instruments are the same for all crime categories, estimation separately for each crime category is best, barring coefficient restrictions across crime categories (Arnold Zellner, 1962 p. 351). If the estimation were performed separately for each crime category, then no weighting would be necessary. However, Levitt imposes coefficient restrictions across crime categories throughout, necessitating joint estimation. Analyzed separately, the largest 2SLS t ratio is 1.4. When analyzed jointly and weighted correctly, the largest 2SLS t ratio is 1.5. Analyzed jointly and weighted incorrectly, the largest 2SLS t ratio increases to 3.4.

In the spirit of replication, I attempted recollection of each series used in Levitt. For the most part, the data replication effort was successful.[1] The primary correction I report is to

* Department of Economics, 549 Evans Hall #3880, University of California, Berkeley, CA 94720 (e-mail: jmccrary@econ.berkeley.edu). I thank David Card, John DiNardo, Ken Chay, and Margaret McConnell for detailed comments on earlier drafts of the paper. Steven Levitt provided both data and computer code. Any errors are my own. Data and computer programs are available at ⟨http://elsa.berkeley.edu/replications/mccrary/index.html⟩.

[1] I was able to accurately replicate Levitt's gubernatorial election data using a combination of web search (for 1991–1992) and *Candidate and Constituency Statistics of Elections in the United States, 1788–1990,* an electronic file available from the Inter-university Consortium for Political and Social Research (ICPSR) (1994). For 1975–1992, Levitt's (hand-entered) data on police and crime differ in only minor respects from electronic data available from ICPSR. A small random sample of data on police and crimes for 1969–1974 were verified against *Crime in the United States*

the mayoral election-year indicator, the source of which is not reported by Levitt. I collected information on mayoral election dates from two published sources, obtaining a measure substantially different from Levitt's and (moderately) more predictive of police hiring. Given this stronger first-stage relationship, one might expect that use of this new measure would lead to greater precision of the 2SLS estimates. However, 2SLS estimates based on my mayoral election-year indicator are actually less precise than the correctly weighted estimates based on Levitt's original election data.

In summary, municipal police force size does appear to vary over state and local electoral cycles. This is an interesting finding in its own right. However, elections do not induce enough variation in police hiring to generate informative estimates of the effect of police on crime.

I. Published Estimates

Levitt (1997) models year-to-year city-level growth rates in crime per capita as a function of two lags in the growth rate of a city's police force size per capita.[2] The coefficient of interest is the elasticity of crime with respect to police; it is estimated by the sum of the two lag coefficients.[3] He argues that cities hire additional police officers in anticipation of projected crime waves, leading OLS estimates of the effect of police on crime to exhibit positive bias.[4]

To overcome this simultaneity bias, Levitt proposes to identify the police effect using only the variation in police hiring induced by the electoral cycle. Given his choice of lag structure, he instruments the lagged police growth rates with lagged indicators of mayoral and gubernatorial election years. While the growth rate in police per capita is significantly faster in election years than in nonelection years, the predictive power of elections is low, as will be discussed further in Section III, below.

Seven crime categories are considered. Although Levitt presents separate estimates for each crime, all specifications impose restrictions on coefficients across the different crimes, as noted in the introduction. Specifically, city effects are constrained to be equal across the seven crime categories, and six state- and MSA-level covariates are constrained to have the same effect among violent crimes, and among property crimes.[5] To impose these restrictions, Levitt estimates the police coefficients jointly, introducing heteroscedasticity due to the different variances of the crime growth rates. Striving for efficiency, Levitt employs a two-step weighting procedure for both OLS and 2SLS. In the first step, he estimates the crime categories jointly without weights, and calculates the variance of the residuals separately for each crime category. In the second step, he again estimates the crime categories jointly, but weights observations for different crimes by a factor reflecting the variability of the different crimes' growth rates. The appropriate weight is the *inverse* of the residual variance. In the bulk of the estimation, however, Levitt weights each crime category by its residual standard deviation. This appears to be a mistake in his computer program, rather than a conscious choice.

If the residual standard deviations were approximately equal across crime categories, then weighting (and thus the weighting error) would be of minor consequence. Column (1) of Table 1 shows the standard deviations of the crime growth rates, along with standard deviations of their (first-step) OLS and 2SLS fitted residuals. For each crime, the three quantities are of similar magnitude. Rare crimes such as murder have highly variable growth rates (standard deviation = 26 percent) compared to common crimes such as larceny (standard deviation = 10 percent).[6] Thus, the weighting error is potentially important.

(Federal Bureau of Investigation, 1969–1974), the source for Levitt's hand entry. However, replication of Levitt's state- and MSA-level covariates was abandoned after failure to reproduce his measure of state and local education and welfare spending. To minimize discrepancies, I utilize Levitt's data with no alterations.

[2] Levitt analyzes data on 59 large U.S. cities with directly elected mayors, 1970–1992.

[3] Levitt's lag structure implies that normalizing crime and police by population does not lead to OLS division bias.

[4] This is also the view of many criminologists. See, for example, Daniel Nagin (1978, but especially 1998). Other types of bias of the OLS estimator and alternative causes of simultaneity bias are not discussed.

[5] The remaining covariates are crime-specific year, region, and city-size indicators.

[6] In Levitt's sample, there are roughly 19 murders and 4,400 larcenies per 100,000 population (Levitt's [1997] table 1).

TABLE 1—ESTIMATES OF THE ELASTICITY OF CRIME WITH RESPECT TO POLICE

Crime type	Standard deviations: Unconditional (OLS residuals) {2SLS residuals} (1)	Published OLS (2)	Published 2SLS (3)	Replication OLS (4)	Replication 2SLS (5)	New mayoral elections measure 2SLS (6)
A. *Separate Estimates for Seven Crime Categories:*						
Violent crimes						
Murder	0.26	−0.60	−3.05	−0.56	−3.03	−2.69
	(0.25)	(0.19)	(0.91)	(0.19)	(2.03)	(2.07)
	{0.29}					
Rape	0.17	−0.06	0.67	0.00	0.74	0.79
	(0.15)	(0.13)	(1.22)	(0.12)	(1.19)	(1.25)
	{0.17}					
Robbery	0.16	−0.31	−1.20	−0.28	−1.39	−0.98
	(0.13)	(0.10)	(1.31)	(0.11)	(1.00)	(1.09)
	{0.14}					
Aggravated assault	0.17	0.11	−0.82	0.17	−0.58	−0.90
	(0.16)	(0.13)	(1.20)	(0.12)	(1.16)	(1.32)
	{0.17}					
Property crimes						
Burglary	0.12	−0.25	−0.58	−0.20	−0.55	−0.47
	(0.10)	(0.08)	(1.55)	(0.08)	(0.67)	(0.77)
	{0.10}					
Larceny[a]	0.10	−0.10	0.26	−0.05	0.53	0.80
	(0.08)	(0.06)	(1.66)	(0.06)	(0.58)	(0.68)
	{0.08}					
Motor vehicle theft	0.15	−0.29	−0.61	−0.24	−0.44	−0.77
	(0.14)	(0.10)	(1.31)	(0.11)	(0.98)	(1.08)
	{0.14}					
Type of weights:		correct	incorrect	correct	correct	correct
B. *Pooled Estimates:*						
All violent crimes		−0.27	−1.39	−0.12	−0.79	−0.66
		(0.06)	(0.55)	(0.06)	(0.61)	(0.65)
All property crimes		−0.23	−0.38	−0.13	0.00	0.11
		(0.09)	(0.83)	(0.04)	(0.34)	(0.43)
Type of weights:		incorrect		correct		correct
Source of mayoral instrument:		Levitt		Levitt		Author
Numbers based on:	Author's calculations	Levitt (1997)		Author's calculations		Author's calculations

Notes: The table presents estimates of the elasticity of crime with respect to police. Column (1) gives standard deviations of the growth rates of the seven crime categories considered (first row), and standard deviations of the first-step OLS (second row, parentheses) and 2SLS (third row, braces) residuals. Columns (2) and (3) present Levitt's estimates. Estimates in the top panel of column (2) are from a weighted, joint regression of the seven growth rates in crime per capita on growth rates in police per capita. Restrictions across crime categories are imposed for unreported coefficients. Specifically, city effects are constrained to be equal for all seven crime categories, and six state- and MSA-level covariates are constrained to have the same effect among violent crimes, and among property crimes. The remaining covariates are year, region, and city-size indicators, which are all allowed to have a different effect on each crime. Weights based on the OLS standard deviations in column (1) were employed to correct for the different variances of the crime growth rates. The pooled estimates in the bottom panel of column (2) impose the further restriction that the effect of police on crime is equal among violent crimes and among property crimes. The weighting procedure used in producing the pooled OLS estimates and *all* 2SLS estimates is incorrect, and gave crime categories with higher variance more weight. Column (3) instruments police growth rates with election-year indicators and the covariates described above. Weights for column (3) are based on the 2SLS standard deviations in column (1). Columns (4) and (5) replicate Levitt's estimates using correct weights. Column (6) replaces Levitt's mayoral election year indicator with my own. For all models, there are 1,129 observations on rape and 1,136 observations for each of the other crime categories, for a total of 7,945 observations.

[a] Ex motor vehicle theft.

Columns (2) and (3) of Table 1 show Levitt's OLS and 2SLS estimates. The top panel gives estimates for each of the seven crime categories, and the bottom panel gives pooled estimates of the effect of police on violent and property crimes. The pooled estimates constrain the elasticity of crime with respect to police to be equal among violent crimes and among property crimes. Of the estimates shown in columns (2) and (3), only the OLS estimates fitted separately by crime category use a correct weighting procedure. The pooled OLS estimates and *all* the 2SLS estimates are weighted incorrectly.

Looking at the OLS estimates in column (2) by crime category, most of the elasticities are negative and in the range of -0.1 to -0.3. Several of the elasticities are statistically significant. In particular, the OLS elasticity for homicides has a t ratio of about 3, as do the elasticities for robbery, burglary, and motor vehicle theft. The pooled estimates for violent and property crime are both near -0.25, and have t ratios above 2.

Compared to the OLS estimates, the 2SLS estimates in column (3) are more negative for all crime categories except rape and larceny. For several of the crimes, the 2SLS estimates are substantially larger in magnitude than their OLS counterparts. For example, the murder elasticity is around -3, with a t ratio of about the same magnitude as the OLS estimate ($t = 3.4$). Taken seriously, this estimate implies that a 10-percent increase in police per capita would reduce murders per capita by 30 percent. The 2SLS estimates for robbery and aggravated assault are also much more negative than the OLS estimates. Nonetheless, murder is the only crime for which the 2SLS estimate is distinct either from zero or from OLS (Hausman statistic = 7.58 [see Jerry Hausman, 1978]).

Both in his presentation of estimates and in his discussion, Levitt emphasizes the pooled 2SLS specifications that group the three violent crimes and the four property crimes.[7] Only the violent crime estimate is statistically significant, leading Levitt to conclude that police reduce violent crime but not property crime. It is important to note the extent to which the magnitude and precision of the violent crime elasticity is driven by the magnitude and precision of the murder elasticity.[8] Aside from any weighting issues, the heavy reliance of the violent crime elasticity on the murder results is potentially troublesome, as this seems to be a priori the crime least likely to respond to marginal changes in police activity.

Before turning to the replication estimates, it is interesting to note that the weighting error could have been inferred from the published estimates: Levitt's OLS and 2SLS standard errors exhibit extreme negative correlation (correlation coefficient = -0.98). Since covariates and instruments are the same for all crimes, the correlation should be very close to $+1$.[9] By a similar logic, standard errors of correctly weighted estimates should be almost perfectly (positively) correlated with the standard deviations of the crime growth rates. Since these were not reported by Levitt, I have included them in column (1). The correlation between the correctly weighted unpooled estimates' standard errors [columns (2), (4), (5), and (6)] and the standard deviations in column (1) is in each instance above 0.97. The analogous correlation for Levitt's 2SLS standard errors [column (3)] is -0.96.

II. Replication Estimates

How do Levitt's conclusions—that police reduce violent crime but not property crime, and that there is positive bias in OLS estimates of the effect of police on crime—hold up to implementation of a correct weighting scheme? To answer this question, I reestimated the elasticities from columns (2) and (3) using a data set

[7] Indeed, unpooled estimates only appear in the final table of the paper.

[8] To make this relationship explicit, note that since the correlation between the estimates for the individual crimes is small (between -0.04 and 0.02), one may approximate well the pooled estimate by a weighted average of the unpooled estimates, with weights summing to one and proportional to the squared inverse of the standard errors (i.e., diagonal minimum distance).

[9] Let $\mathbf{s}$ and $\mathbf{t}$ denote the 7-vectors of standard errors of the OLS and 2SLS estimates, respectively. Since covariates and instruments are the same for all crime categories, and the predictive power of the models does not vary much by crime category, approximate $\mathbf{s} = c\sigma$ and $\mathbf{t} = d\sigma$, where σ is the 7-vector of crime-growth-rate standard deviations and c and d are positive constants. Then $\mathbf{s}$ is in the column space of $\mathbf{t}$ and the correlation is $+1$ since c and d share sign. The correlation between the correctly weighted standard errors in column (4) and (5) is 1.00.

provided by Levitt, but correcting the weighting error.[10] These corrected estimates are given in columns (4) and (5) of Table 1.

The unpooled OLS estimates in column (4) should be identical to those published since Levitt weighted those estimates correctly. There are nonetheless some differences. I believe these are due to minor changes in the data set supplied by Levitt relative to the one he used in producing his published estimates, and/or to differences between the specification described in the text of his paper and that used in producing his estimates. Overall, however, the unpooled OLS estimates in column (4) are very close to those in the original paper.

By comparison, Levitt's pooled OLS estimates use an incorrect weighting procedure, and the replication estimates are both less than half those published. Both are near −0.12 and have *t* ratios above 2. The smaller size of the correctly weighted pooled estimates reflects a general pattern in the estimates: crime categories with greatest year-to-year variability exhibit the largest effects.

This tendency is even more pronounced among the unpooled 2SLS estimates, presented in column (5). As would be expected, correcting the weights alters the point estimates little. However, the effect on the standard errors is substantial. The rank order of the standard errors is the reverse of that of the published, and none of the unpooled estimates are distinct from zero. The murder *t* ratio is 1.5, and the remaining unpooled *t* ratios are all below 1.

As noted above, the published pooled violent crime estimate relies heavily on both the large magnitude and apparent precision of the murder estimate. This reliance is made clear by the replication estimates. The correctly weighted pooled violent crime estimate discounts the large magnitude of the murder coefficient because of the variability of murder growth rates, leading to an estimate just over half the published value. Coupled with the larger standard error, this results in a wide confidence region of (−2.0, 0.4). The pooled property crime estimate is also less negative than the published magnitude and has a confidence region of (−0.68, 0.68).[11] Thus, correctly implemented, Levitt's identification strategy does not allow statistical rejection of most economically meaningful hypotheses.

Levitt's second conclusion, that the OLS estimates exhibit positive bias, is also without statistical justification. When correctly weighted, none of the nine OLS–2SLS comparisons are significant at even the 10-percent level. On the other hand, it is true that for the five categories of crime excepting rape and larceny, the 2SLS estimates are more negative than the OLS estimates. Perhaps greater precision of the 2SLS estimates would strengthen our confidence that OLS estimates exhibit positive bias.[12]

III. Can Improved Dating of Mayoral Elections Increase Precision?

A potential explanation for the imprecision of the correctly weighted 2SLS estimates is the presence of errors in the dating of local election cycles. While gubernatorial elections are measured quite well, there is some measurement error in Levitt's mayoral election-year indicator. As part of my replication effort, I recollected data on mayoral elections for Levitt's 59 cities from the *World Almanac* (Newspaper Enterprise Association, 1960–1998) and the *Municipal Yearbook* (International City Managers' Association, 1960–1998).[13] For 23 of the cities, Levitt's measure and my measure are identical. For 33 cities, the measures are in substantial disagreement, and for three cities the measures are in moderate disagreement.

[10] In addition to his data set, Levitt provided me with a SAS computer program which (almost) produces his published estimates. It is only through inspection of this program that I recognized the weighting error.

[11] Although not addressed here, confidence regions should be even larger due to the weak correlation between police hiring and elections. For tests of correct size and references to the weak instruments literature, see Marcelo Jovita Moreira (2001).

[12] For his original paper, Levitt (1995) presented three other sets of estimates I have not discussed: (1) 2SLS using elections interacted with city-size indicators as instruments for police hiring; (2) 2SLS using elections interacted with region indicators as instruments; and (3) LIML using elections interacted with region indicators as instruments. Replications of these specifications, presented in Table 3 of McCrary (2001), reflect the conclusions already drawn. Of the 72 replication estimates shown there, two are (marginally) significant at the 5-percent level, and none is different from the corresponding OLS estimates.

[13] The data on mayoral elections are described in more detail in McCrary (2001).

TABLE 2—ESTIMATES OF THE ELECTORAL CYCLE IN POLICE HIRING

	Levitt measure of mayoral elections		New measure of mayoral elections	
Election-year indicator	Δ ln Police$_{t-1}$ (1)	Δ ln Police$_{t-2}$ (2)	Δ ln Police$_{t-1}$ (3)	Δ ln Police$_{t-2}$ (4)
Mayor$_{t-1}$	0.0091	0.0053	0.0143	−0.0098
	(0.0049)	(0.0050)	(0.0048)	(0.0049)
Mayor$_{t-2}$	−0.0037	0.0149	0.0037	0.0065
	(0.0049)	(0.0050)	(0.0048)	(0.0049)
Governor$_{t-1}$	0.0262	−0.0078	0.0248	−0.0070
	(0.0068)	(0.0070)	(0.0068)	(0.0069)
Governor$_{t-2}$	−0.0010	0.0259	0.0001	0.0242
	(0.0069)	(0.0070)	(0.0068)	(0.0070)
R^2:	0.1131	0.1083	0.1157	0.1110
Number of observations:	1,136	1,136	1,136	1,136
F test on exclusion of all four election-year indicators:	5.07	6.09	5.84	6.91
	(p = 0.00)	(p = 0.00)	(p = 0.00)	(p = 0.00)
F test on Mayor$_{t-1}$ + Governor$_{t-1}$ = 0:	16.02	0.08	21.09	3.77
	(p = 0.00)	(p = 0.78)	(p = 0.00)	(p = 0.05)
F test on Mayor$_{t-2}$ + Governor$_{t-2}$ = 0:	0.28	20.49	0.20	12.92
	(p = 0.60)	(p = 0.00)	(p = 0.65)	(p = 0.00)
Source of mayoral instrument:	Levitt	Levitt	Author	Author

Notes: Table presents OLS estimates from a regression of growth rates of police per capita on mayoral and gubernatorial election-year indicators. Also included in the estimation are year, city, and city-size indicators, and six state- and MSA-level covariates. Region indicators, included in Levitt's 2SLS specification, are absorbed by city indicators in the first stage. In contrast, city-size indicators as defined by Levitt vary over time and are not absorbed by the city indicators. In all of Levitt's 2SLS specifications, both lags of police growth rates are deemed endogenous; as such, two lags of each instrument are used. Columns (1) and (2) utilize Levitt's measure of mayoral elections, while columns (3) and (4) use my measure. The number of observations here differs from Levitt's table 2 because the results here rely only on the observations utilized in the 2SLS regressions. Strictly speaking, the coefficients reported here apply only to the crime categories excepting rape, but first-stage results for the 1,129 observations on rape are quite similar.

Table 2 presents the first-stage regressions for both Levitt's electoral measure and my measure. Specifically, the table shows coefficients from a regression of once- and twice-lagged growth rates in police per capita on once- and twice-lagged mayoral and gubernatorial election-year indicators. Also included in the specification (but not shown) are the exogenous regressors used in Table 1. The first stage is complicated by the use of two lags of each of the election indicators. Unfortunately, Levitt's choice of lag structure (current crime growth rates are modeled as a function of once- and twice-lagged growth rates in police) renders this complication unavoidable.

Columns (1) and (2) use Levitt's mayoral election-year indicator, while columns (3) and (4) employ my measure. The two measures appear to have very similar effects on police hiring. The F statistic on the exclusion of the four election indicators are stronger using my measure, but the differences are minor. Perhaps the most interesting pattern in Table 2 is that mayors have a smaller effect on police hiring than governors. It is possible that this pattern is attributable to measurement error in the mayoral election-year indicators (both Levitt's and my own).[14]

Levitt's specification makes it difficult to accurately summarize the effect of elections on growth rates in police per capita. Heuristically, however, it is useful to consider the implications of the estimates for a city with a four-year

[14] In addition, the significance of gubernatorial elections appears to be overstated by about 10 percent due to a Brent J. Moulton (1986) effect. There are eight cities in California and Texas, four cities in Florida and Ohio, three cities in Arizona, and two cities in New Jersey, Tennessee, Pennsylvania, Missouri, New York, and Oklahoma.

mayoral and gubernatorial election cycle in which the elections are held in the same year. According to the estimates in Table 2, such a city would exhibit no growth in nonelection years, contrasted with 3–4 percent growth in election years. In the context of police officers per capita, this is relatively rapid growth.[15] However, the variation in police hiring induced by elections is small. The F statistics on the exclusion of the four election-year indicators suggest that only 2 percent of the growth rate in police per capita may be explained by the electoral cycle.

Column (6) of Table 1 gives 2SLS estimates that result from replacing Levitt's mayoral election-year measure with my own. The point estimates are slightly different than those in column (5), but are qualitatively similar. Five of the seven estimates are less negative than the corresponding estimates that use Levitt's measure. However, use of my measure *increases* the standard errors for every estimate, despite the slightly stronger relationship between elections and police hiring. Following the pattern of the estimates reported in Section II, none of the estimates using my mayoral election-year indicator (either unpooled or pooled) is significantly different from zero or from OLS. Thus, even with a somewhat stronger first stage, it does not appear possible to obtain precise estimates of the effect of police on crime using elections as instruments.

IV. Conclusion and Discussion

Although Levitt's weighting error led to mistaken inferences, his article makes at least two contributions that should not be overlooked. First, he appears to be only the second researcher to collect city-level data on crime and police spanning more than two years, and the first to use such data to examine the effect of police on crime.[16] Replication OLS estimates of the effect of police on violent and property crime are both roughly -0.12 and are estimated with some precision. Given that criminologists have argued for over 20 years that such estimates exhibit positive bias, these might be taken as evidence in favor of the hypothesis that police reduce crime.

Second, Levitt provides reasonably convincing evidence of an electoral cycle in police hiring. This, too, is an important contribution. An electoral cycle in police hiring represents a failure of the political process to allocate resources efficiently. Although often asserted, evidence of such failures is somewhat rare. The results presented here suggest that the electoral cycle in police hiring may be somewhat stronger than originally reported.

However, it does not appear possible to use these data to learn about the causal effect of police on crime. Although elections significantly predict growth rates in city police force size, they do not significantly predict crime growth rates. As a result, 2SLS estimates of the effect of police on crime using election-year indicators as instruments are indistinct from zero, and indistinct from OLS estimates. Consequently, this identification strategy provides little evidence that police reduce crime, and even less evidence that OLS estimates of the effect of police on crime exhibit positive bias. In the absence of stronger research designs, or perhaps heroic data collection, a precise estimate of the causal effect of police on crime will remain at large.

[15] In these data, the average growth rate in police per capita is 0.0068, with a standard deviation of 0.0616.

[16] This conclusion is based on literature reviews in Thomas B. Marvell and Carlisle Moody (1996) and John E. Eck and Edward R. Maguire (2000), which together summarize 45 articles.

REFERENCES

Eck, John E. and Maguire, Edward R. "Have Changes in Policing Reduced Violent Crime? An Assessment of the Evidence," in Alfred Blumstein and Joel Wallman, eds. *The crime drop in America*. New York: Cambridge University Press, 2000, pp. 207–65.

Federal Bureau of Investigation. *Crime in the United States*. Washington, DC: U.S. Government Printing Office, 1969–1974.

Hausman, Jerry. "Specification Tests in Econometrics." *Econometrica*, November 1978, *46*(6), pp. 1251–71.

International City Managers' Association. *Municipal yearbook*. Washington, DC: International City Managers' Association, 1960–1998.

Inter-university Consortium for Political and Social Research. *Candidate and constituency statistics of elections in the United States,* *1788–1990*. [Computer File] Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 1994.

Levitt, Steven. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 4991, January 1995.

______. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review*, June 1997, *87*(3), pp. 270–90.

Marvell, Thomas B. and Moody, Carlisle. "Specification Problems, Police Levels, and Crime Rates." *Criminology*, November 1996, *34*(4), pp. 609–46.

McCrary, Justin. "Do Electoral Cycles in Police Hiring Really Help Us Estimate the Effect of Police on Crime?" Center for Labor Economics (Berkeley, CA) Working Paper No. 35, September 2001.

Moreira, Marcelo Jovita. "Tests with Correct Size When Instruments are Arbitrarily Weak." Mimeo, University of California–Berkeley, January 2001.

Moulton, Brent J. "Random Group Effects and the Precision of Regression Estimates." *Journal of Econometrics*, August 1986, *32*(3), pp. 385–97.

Nagin, Daniel. "General Deterrence: A Review of the Empirical Evidence," in Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin, eds., *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates*. Washington, DC: National Academy of Sciences, 1978, pp. 95–139.

______. "Criminal Deterrence Research at the Outset of the Twenty-First Century," in Michael Tonry, ed., *Crime and justice: A review of research*, Vol. 23. Chicago: University of Chicago Press, 1998, pp. 1–42.

Newspaper Enterprise Association. *World almanac*. New York: Newspaper Enterprise Association, 1960–1998.

Zellner, Arnold. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." *Journal of the American Statistical Association*, June 1962, *57*(298), pp. 348–68.