CENTER FOR LABOR ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY
WORKING PAPER NO. 35

Do Electoral Cycles in Police Hiring Really Help Us Estimate the Effect of Police on Crime?

Justin McCrary
Department of Economics
University of California, Berkeley
jmccrary@econ.berkeley.edu

September 2001

Abstract

This paper replicates a well-known article by Steven Levitt (*AER*, 1997). Levitt argues that an electoral cycle in police hiring generates exogenous variation in city police force size which may be used to estimate the causal effect of police on crime. I point out that a weighting error in Levitt's computer program led to incorrect inferences for the key results of his paper. Replication estimates using the correct (and intended) weighting scheme are all statistically insignificant at even the 10% level. In addition to correcting the weighting error, I present results based on a new mayoral election year indicator. This new measure suggests that the electoral cycle in police hiring is (moderately) stronger than reported in Levitt. However, two-stage least squares estimates using this new measure are even *less* precise than (correctly weighted) estimates using Levitt's original measure, despite the stronger first-stage relationship.

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. All errors are my own. Data and computer programs are available at http://elsa.berkeley.edu/replications/mccrary/index.html.

Introduction

In an influential paper, Steven Levitt argues that there is an electoral cycle in police hiring, with faster hiring in election years and slower hiring in other years.[1] He then uses elections as an instrument for police hiring to estimate the causal effect of police on crime. This paper 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 no significant effect on property crime. Second, 2SLS estimates of the effect of police on crime are consistently more negative than ordinary least squares (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 the opposite, giving highly variable crimes such as murder 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 just over half the published magnitude and are 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 their OLS counterparts, and, when correctly weighted, OLS and 2SLS estimates are never statistically distinguishable.

In the spirit of replication, I attempted re-collection of all data used in Levitt's paper.[2] The primary correction I report is to Levitt's mayoral election year indicator (see Section 3 and the data appendix). My mayoral election year indicator differs substantially from his, and is (moderately) more predictive of police hiring. Given this stronger first stage relationship, one might expect that use of my measure would lead to greater precision of the 2SLS estimates. Surprisingly, 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 both the state and local electoral cycle. This is an interesting finding in its own right, and warrants further research. However, in these data elections do not induce enough variation in police hiring to generate informative estimates of the effect of police on crime.

---

[1] "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, June 1997.

[2] The data replication effort was mostly successful. For 1975 to 1992, Levitt's data on police and crime, although hand-entered, differ in only minor respects from electronic data available from the Inter-university Consortium for Political and Social Research (ICPSR). A small random sample of earlier years' data on police and crimes were verified against *Crime in the United States*, an annual publication of the Federal Bureau of Investigation summarizing the *Uniform Crime Reports* (UCR), and the source for Levitt's hand-entry. I was able to accurately replicate Levitt's gubernatorial election data using a combination of web search (for 1991 and 1992) and *Candidate and Constituency Statistics of Elections in the United States, 1788-1990*, an electronic file available from ICPSR. 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. I was also unable to replicate Levitt's mayoral electoral year indicator. Throughout, I use Levitt's data with no alterations.

1. Published Estimates

Levitt models year-to-year growth rates in crime per capita as a function of two lags in the growth rate of a city's per capita police force size.[3] The coefficient of interest is the elasticity of crime with respect to police, and is estimated by the sum of the two lag coefficients.[4] Levitt 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.[5]

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, yielding four instruments for two endogenous variables.[6] While the

---

[3] Levitt analyzes data on 59 large U.S. cities with directly elected mayors, 1970 to 1992. Exogenous covariates include six state- and MSA-level variables, city indicators, and year, region, and city-size indicators. The six state- and MSA-level covariates are: (1) MSA percent black and MSA percent female-headed households (from the *Statistical Abstract of the United States* (SAUS) with linear interpolation between census years), (2) MSA percent ages 15 to 24 (from an unknown marketing almanac, also linearly interpolated between census years) (3) state unemployment rates (from Blanchard and Katz, 1992, available annually), and (4) real per capita state and local spending on education and public welfare (from the SAUS, available annually). The region indicators are for the nine Census regions. The city-size indicators are indicators for a city having population of zero to 250,000; 250,000 to 500,000; 500,000 to 1,000,000; and over 1,000,000. The sources for Levitt's state- and MSA-level covariates are taken from Levitt's electronic documentation file, which is posted at the website given in the acknowledgement footnote. Levitt's text (stub to Table 1) cites SAUS as the source for all data not in the UCR.

[4] Levitt's lag structure implies that normalizing crime and police by population does not bias estimates toward one. The lag structure I describe is different from that Levitt describes in the text ("…police are allowed to affect crime both contemporaneously and with a one-year lag," p. 279). The discrepancy in language pertains to the late date of the measurement of the police force size variable. The codebook for ICPSR study number 9028, Police Employee (LEOKA) Data, 1975, states that the "Law Enforcement Employees Report, a form sent to police agencies throughout the country by the FBI once a year, [asks] for a count of employees on the payroll of each agency *as of October 31*. The number of full-time law enforcement employees, officers and civilians, [is requested]. Officers are defined as full-time, sworn personnel with full arrest powers including the chief, sheriff, or other head of the department" (p. ii, my emphasis). Virtually identical language is used for each codebook through 1999. Levitt's police force size measure is the number of officers described in the codebook. The number of officers reported in the 1975 UCR, divided by the population measure reported in the 1975 UCR, is in Levitt's notation P_{1976} (compare with Levitt, p. 275, 279).

[5] This is also the view of many criminologists. See, for example, Nagin (1978, but especially 1998). Other types of bias of the OLS estimator are not discussed in Levitt, and are often downplayed in the criminological literature.

[6] In all specifications, Levitt estimates the 7 crime categories *jointly*. Hence there are, strictly speaking, 14 endogenous variables and 28 instruments (for the specifications in Table 1). The specifications in Table 3 utilize alternately 112 and 252 instruments. The former instrument set interacts the 28 election-year indicators with the four city-size indicators, while the latter instrument set interacts the 28 election-year indicators with the nine region indicators.

growth rate in police per capita is significantly faster in election years than in non-election years, the predictive power of elections is low, as will be discussed further in Section 3, below.

Seven crime categories are considered, ranging in seriousness from murder to larceny. Although Levitt presents separate estimates for each crime category, all specifications impose restrictions on coefficients across the different crime categories.[7] This leads Levitt to use weighted OLS and 2SLS procedures, which are implemented in two steps. In the first step, Levitt 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, and would assign lower weight to crime categories with higher year-to-year variability. 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 code, 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 OLS and 2SLS fitted residuals. Rare crimes such as murder have highly variable growth rates

---

[7] The separate estimates for each crime category are estimated *jointly* and impose restrictions on the coefficients on the city effects, and on the six state- and MSA-level covariates. Specifically, city effects are constrained to be equal across all crime categories considered, and the state- and MSA-level covariates are constrained to have equal effect among the four violent crimes and among the three property crimes. Separate coefficients for each crime category are estimated for the year, region, and city-size indicators. *Pooled* estimates of the effect of police on violent crime and on property crime impose the additional restriction that the effect of police on crime be the same among violent crimes, and among property crimes. This description of Levitt's specification is taken from the stub of Levitt's Table 5 (p. 282), which present the estimates for each crime category: "In all cases, specifications are identical to that [sic] used in Tables 3 and 4 [which report pooled estimates for violent and property crime, respectively], except that the cross-crime restrictions on police elasticities… have been removed in this table." This description is slightly different from Levitt's description (p. 279, 284) of his specification in the text, but I believe it to be the correct one. Finally, note that since covariates and instruments do not vary by crime, without coefficient restrictions across crime

(standard deviation=26%) compared to common crimes such as larceny (standard deviation=10%).[8] Thus, the weighting error is potentially important.

Columns (2) and (3) of Table 1 show Levitt's OLS and 2SLS estimates. The top panel gives estimates for each 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 and the state- and MSA-level covariates 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 estimated 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 two.

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=3.35). Taken seriously, this estimate implies that a 10% increase in police per capita would reduce murders per capita by 30%. The 2SLS estimates for robbery and aggravated assault are also much more negative than the OLS estimates. However, murder is the only crime for which the 2SLS and OLS estimates are statistically distinct (Hausman (1978)-Wu (1973) statistic=7.58, p-

---

categories, the coefficients should be estimated separately by crime category (Zellner 1962). Separate estimation by crime category would have eliminated the need for a weighting correction and (consequently) the weighting error.

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

value<0.01). Murder is also the only crime for which the 2SLS estimate is distinct from zero; the six other t-ratios are all below one.

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.[9] 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.[10] The heavy reliance of the violent crime elasticity on the murder results is potentially troublesome. Perhaps half of all murders are committed in passionate contexts, and one would not necessarily expect such crimes to be affected by marginal changes in police activity.[11] Additionally, any goodness of fit criterion applied to murder alone would clearly fail: the standard deviation of the 2SLS residuals for murder *exceeds* the standard deviation of the data (0.29>0.26, from column (1)).[12]

---

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

[10] Since the correlation between the estimates for the individual crimes is small (between –0.04 and 0.02), one may accurately approximate 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.

[11] In each year from 1995 to 1999, roughly 45% of all murders (for which circumstances were reported) grew out of romantic triangles, inebriated brawling, or arguments. See http://www.fbi.gov/ucr/Cius_99/99crime/99c2_03.pdf, Table 2.14. Accessed 9/3/01. The table summarizes evidence from the Supplemental Homicide files. By comparison, adult and juvenile gang killings account for less than 10%.

[12] In retrospect, the weighting error could have been inferred from the published estimates. Levitt's OLS and 2SLS standard errors exhibit extreme negative correlation (correlation coefficient = –0.975). Since covariates and instruments are the same across crimes, standard errors for the OLS and 2SLS estimates of the police effect should be proportional to standard deviations of the crime category residuals. Such strong negative correlation would be possible only if the rank order of the OLS and 2SLS standard deviations were extremely different. Column (1) shows that they are virtually identical. When the weighting error is corrected, the correlation of the OLS and 2SLS standard errors is 0.997.

2. 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 re-estimated the elasticities from columns (2) and (3) using a data set provided to me by Levitt, but correcting the weighting error described above.[13] 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. The replication estimates are both less than half those published. The property crime estimate is significant at the 5% level, and both the violent and the property crime estimates are significant at the 10% level. 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 estimated police effects.

This tendency is even more pronounced among the unpooled 2SLS estimates, presented in column (5). While the replication unpooled 2SLS point estimates are generally similar to those published, the standard errors are very different. In particular, the rank order of the replication standard errors is virtually the reverse of that of the published, with the result that 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 one.

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). Thus, in using electoral cycles to identify the effect of police on crime, it does not appear possible statistically to reject 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% 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.[14]

---

[13] 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 was able to identify the weighting error.

[14] For his original paper, Levitt 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, reflect the conclusions already drawn. Of the 72 replication estimates shown there, two are (marginally) significant at the 5% level, and none are different from the corresponding OLS estimates.

3. 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 re-collected data on mayoral elections for Levitt's 59 cities.[15] For 23 of the cities, Levitt's measure and my measure are identical. For 33 cities, the measures are in substantial disagreement, and for 3 cities the measures are in moderate disagreement.[16] When the two measures disagree, a common pattern (203 elections, or 14% of all observations) is for my measure to lead Levitt's measure by one year.[17] Another source of discrepancy is changes in mayoral term. For example, for many years Austin, Texas—like most other cities in Texas—conducted mayoral elections every two years. However, in 1985, voters approved amendments to the city charter lengthening the mayor's term to three years.[18] My data on mayoral election dates suggest that alterations to mayoral term occurred in 15% of the cities in Levitt's sample between 1969 and 1992. By comparison, Levitt's measure suggests that *no* city changed mayoral term over this period.

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

---

[15] The data appendix gives a detailed description of the data sources for mayoral election dates, and Table A.1 gives city-by-city comparisons of Levitt's and my measure.

[16] Among the 33 "substantial disagreement" cities, for 21 I was able to obtain complete records on election dates spanning 40 years. For the remaining 12 cities, some interpolation between reporting years was necessary.

[17] Mayoral elections, unlike state or national elections, are not always held in November. Elections in the early spring, especially primary elections, are common and necessitate a judgement regarding whether the mayor would strive to increase the size of a police force in the year leading up to the election, or the year of the election. For elections in the first half of the year, I chose the former; Levitt appears to have opted for the latter. See the data appendix for additional details.

[18] See http://malford.ci.austin.tx.us/election/search.cfm. Accessed 9/3/01.

year indicators. Also included in the specification (but not shown) are the exogenous regressors used in Table 1.[19] Note that the first stage regression does not vary by crime category. The first stage is complicated by the use of two lags of each of the election indicators. Unfortunately, Levitt's choice of lag structure 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.[20] Perhaps surprisingly, mayors appear to have a smaller and less robust effect than governors.[21] The F-statistic on the exclusion of the four election indicators is slightly stronger using my measure, but the point estimates and standard errors are quite similar.[22]

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 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 non-election years, contrasted with 3 to 4 per cent growth in election years.[23] In the context of police officers per capita, this is relatively rapid growth.[24] However, the variation in police hiring induced by elections is small. The F-statistics on the exclusion of

---

[19] Included in the specifications in Table 2 are: year, city, and city-size indicators, and the six state- and MSA-level covariates. Region indicators are excluded. In the overall 2SLS specification, region indicators are separately identifiable from city indicators only because city effects are constrained to be equal across crime categories (in contrast, the region effects are allowed to vary across crime category). Since Levitt defined city-size indicators as time-varying, city-size indicators are separately identifiable from city indicators in the first stage.

[20] The correlation between the two measures is only one-quarter. The cross-sectional intuition that measurement error implies attenuation is somewhat misleading here since observations are serially correlated.

[21] This seems plausible only to the extent that state grants play an important role in the hiring of municipal police officers. Investigation of this issue was deemed beyond the scope of this comment.

[22] Stock and Yogo's (2001) statistic may be viewed as an index of the strength of the first stage relationship. Using Levitt's measure of mayoral elections, this statistic is 6.29, and with my own it is 7.80.

[23] There are four possible estimates: the sum of the mayoral and gubernatorial coefficients based on column (1), at the first and second lag, and the sum based on column (2), at the first and second lag. Using Levitt's mayoral election year indicator, the estimates and standard errors, in the order cited above, are 0.0353 (0.0088), -0.0047 (0.0088), -0.0025 (0.0090), and 0.0408 (0.0090). For comparison, estimates based on my measure are 0.0390 (0.0085), 0.0038 (0.0084), -0.0168 (0.0087), and 0.0307 (0.0085). The first and fourth quantities estimate the growth rate in an election year, and the second and third quantities estimate the growth rate in a non-election year.

[24] In Levitt's data the average growth rate in police per capita is 0.0068, with a standard deviation of 0.0616.

the four election year indicators suggest that only 2% of the growth rate in police per capita may be explained by the electoral cycle.[25]

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. Strikingly, 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 2, none of the estimates using my mayoral election year indicator—either unpooled or pooled—are 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.

4. Conclusion and Discussion

Although Levitt's weighting error led to mistaken inferences, his article makes at least two contributions which 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.[26] His 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 twenty years that such estimates

[25] Although not addressed here, confidence regions should be even larger than the already large Wald regions due to the low predictive power of elections for police hiring. See, for example, Hahn and Hausman (2000) and Staiger and Stock (1997). Anderson and Rubin (1949, 1950) and Moreira (2001) give similar tests.

[26] This conclusion is based on literature reviews in Marvell and Moody (1996) (who themselves utilize Levitt's data) and Eck and Maguire (2000), which together summarize 45 articles.

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 is slightly stronger than originally reported.[27]

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. In the absence of stronger research designs, or perhaps heroic data collection, a precise estimate of the causal effect of police on crime seems likely to remain at large.

[27] That governors exert more influence over city police force size than mayors in these data is, however, a puzzle.

Data Appendix

Mayoral election dates are determined on a city-by-city basis and vary substantially over 1969 to 1992. To my knowledge, there is no electronic source for mayoral election dates, necessitating a good deal of data collection effort on the part of the researcher. In an electronic file provided to me by Levitt, he describes his data source as "various almanacs and documents from city associations." I used a variety of sources, ranging from two annual publications, to web search, to newspaper search, to (a small number of) phone calls to city clerk offices. A few cities have made much of their historical election data available at the city web site.[28]

For most cities, I relied on the *World Almanac (WA)* and the *Municipal Yearbook (MY)*, annual publications available in most libraries. The *WA* and the *MY* contain a listing of cities, the current mayor, and the date of the next election. The coverage in the *WA* is complete, in the sense that the listing appears each year. However, the *WA* sometimes reports information on the city manager in place of mayoral information. In contrast, the *MY* fails to provide the listing for some years, but when the listing is given, it always gives information on the mayor. Using these listings provides two to four observations on each election date since the date of the upcoming election remains unchanged for the length of the mayor's term. This is helpful as there appears to be moderate measurement error in the dates given for upcoming election. The *WA* was used as the primary source, with the *MY* playing a corroborative role for 1979-1983.

Use of these sources allows determination not only of the year of mayoral elections, but also the month of the election. The ability to pin down the month of an election is likely to be important; while national and state elections tend to be held in November, the same is not true of local elections. For example, the current mayor of New Orleans, Marc Morial, was elected to his first term in a runoff election in March, 1994, after placing second in the primary election of February, 1994. Morial then won the primary election outright in February, 1998. Adherence to the calendar year would suggest that the growth rate in police would be higher than average in 1998, but this seems a mistake. If Morial believed that increasing police force size would improve his chances of re-election, he would presumably have striven to hire additional officers in 1997. This seems particularly so, since the police force size measure is a snapshot as of October 31.

I adopted the convention that a mayor would strive to increase the size of the police force in year t if the mayoral election were to occur in year t+1 prior to (weakly) July. If the mayoral election were to occur in year t+1 after July, I judged that the mayor would want to increase the size of the police force in year t+1. July is a good bit later in the year than February/March, the months in the example given, and may be criticized on that ground. I chose July for two reasons. First, as noted in the case of New Orleans, mayoral elections typically involve both a primary and (if necessary) a run-off election. The month given in the two annual publications appears to be the month of the run-off election. For example, for Los Angeles, the month given in both annuals is June. However, Los Angeles has a primary election in April, and a run-off election in June.[29] Note that in many cities, one may win the mayoralty outright by obtaining more than 50% of the vote in the primary. In cities with this rule, the mayor would do well to increase police force size prior to the primary.

Second, and perhaps more practically, both the *WA* and the *MY* occasionally change the heading above the given date from "date of next election" to "date assume office". Since the

[28] Note that the URL for most cities follows the pattern http://www.ci.cityname.st.us, where "cityname" indicates either the city's name or an abbreviated version of it, and "st" denotes the state postal abbreviation.
[29] http://www.lacity.org/clk/election/archiveindex.htm. Accessed 9/10/01.

latter is an upper bound on the former, it seems that when faced with a range of months, the sensible action is to treat the minimum of the months as the month of election.

Table A.1 gives a detailed description of the discrepancies between Levitt's mayoral election year indicator and my own. The election dates published in the *WA* and the *MY,* as well as my mayoral election year indicator based on those dates and all of Levitt's data, are contained in the data set posted at http://elsa.berkeley.edu/replications/mccrary/index.html.

Bibliography

Anderson, T.W. and Herman Rubin, "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, March 1949.

Anderson, T.W. and Herman Rubin, "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, December 1950, pp. 570-582.

Blanchard, Olivier and Lawrence Katz, "Regional Evolutions," *Brookings Papers on Economic Activity*, 1992:1.

Federal Bureau of Investigation, *Crime in the United States*, 1969-1974.

Eck, John E. and Edward R. Maguire, "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.

Hausman, Jerry, "Specification Tests in Econometrics," *Econometrica*, November 1978.

Hahn, Jinyong and Jerry Hausman, "A New Specification Test for the Validity of Instrumental Variables," mimeo, March 2000, forthcoming *Econometrica*.

Inter-university Consortium for Political and Social Research. *Candidate and Constituency Statistics of Elections in the United States, 1788-1990* [Computer file]. 5th ICPSR ed. 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," *NBER Working Paper* # 4991, January 1995.

Levitt, Steven, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, June 1997.

Marvell, Thomas B. and Carlisle Moody, “Specification Problems, Police Levels, and Crime Rates,” *Criminology*, November 1996.

Moreira, Marcelo Jovita, “Tests with Correct Size When Instruments are Arbitrarily Weak,” mimeo, January 2001.

Municipal Yearbook, 1960-98.

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, D.C.: National Academy of Sciences, 1978.

Nagin, Daniel, “Criminal Deterrence Research at the Outset of the Twenty-First Century,” in Michael Tonry, ed., *Crime and Justice: A Review of Research*, Chicago: University of Chicago Press, 1998.

Staiger, Douglas and James H. Stock, “Instrumental Variables Regression with Weak Instruments,” *Econometrica*, May 1997.

U.S. Bureau of the Census, *Statistical Abstract of the United States*, 1969-1992.

Stock, James and Motohiro Yogo, “Testing for Weak Instruments in Linear IV Regression,” mimeo, August 2001.

World Almanac, 1960-98.

Wu, De Min, “Alternative Tests of Independence Between Stochastic Regressors and Disturbances,” *Econometrica*, July 1973.

Zellner, Arnold, “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias,” *Journal of the American Statistical Association*, June 1962.

Table 1. Estimates of the Elasticity of Crime with Respect to Police

	Standard Deviations	**Published**		**Replication**		**New Mayoral Elections Measure**
Estimates for Each 7 Crime Categories	Unconditional (OLS fitted residuals) {2SLS fitted residuals}	OLS	2SLS	OLS	2SLS	2SLS
Violent Crimes	(1)	(2)	(3)	(4)	(5)	(6)
Murder	0.26 (0.25) {0.29}	-0.60 (0.19)	-3.05 (0.91)	-0.56 (0.19)	-3.03 (2.03)	-2.69 (2.07)
Rape	0.17 (0.15) {0.17}	-0.06 (0.13)	0.67 (1.22)	0.00 (0.12)	0.74 (1.19)	0.79 (1.25)
Robbery	0.16 (0.13) {0.14}	-0.31 (0.10)	-1.20 (1.31)	-0.28 (0.11)	-1.39 (1.00)	-0.98 (1.09)
Aggravated Assault	0.17 (0.16) {0.17}	0.11 (0.13)	-0.82 (1.20)	0.17 (0.12)	-0.58 (1.16)	-0.90 (1.32)
Property Crimes						
Burglary	0.12 (0.10) {0.10}	-0.25 (0.08)	-0.58 (1.55)	-0.20 (0.08)	-0.55 (0.67)	-0.47 (0.77)
Larceny--ex MV Theft	0.10 (0.08) {0.08}	-0.10 (0.06)	0.26 (1.66)	-0.05 (0.06)	0.53 (0.58)	0.80 (0.68)
Motor Vehicle Theft	0.15 (0.14) {0.14}	-0.29 (0.10)	-0.61 (1.31)	-0.24 (0.11)	-0.44 (0.98)	-0.77 (1.08)
Type of Weights:		Correct	Incorrect	Correct		Correct
Pooled Estimates						
All Violent Crimes		-0.27 (0.06)	-1.39 (0.55)	-0.12 (0.06)	-0.79 (0.61)	-0.66 (0.65)
All Property Crimes		-0.23 (0.09)	-0.38 (0.83)	-0.13 (0.04)	0.00 (0.34)	0.11 (0.43)
Type of Weights		Incorrect		Correct		Correct
Source of mayoral instrument		Levitt		Levitt		Author
Numbers based on:	Author's calculations	*AER*, June 1997		Author's calculations		Author's calculations

Notes:

Table presents estimates of the elasticity of crime with respect to police. Column (1) gives standard deviations of the growth rates of the 7 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 7 growth rates in crime per capita on growth rates in police per capita. Restrictions across crime categories are imposed on other covariates' coefficients. City indicators are constrained to be equal across all 7 categories,and six state- and MSA-level covariates are constrained to have the same effect among violent crimes, and among property crimes. Also included are crime-specific year, region, and city-size indicators. 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 1129 observations on rape and 1136 observations for each of the other crime categories, for a total of 7945 observations.

Table 2. Estimates of the Electoral Cycle in Police Hiring

	Levitt Measure of Mayoral Elections		New Measure of Mayoral Elections	
	ΔlnPolice $_{t-1}$	ΔlnPolice $_{t-2}$	ΔlnPolice $_{t-1}$	ΔlnPolice $_{t-2}$
Election Year Indicator	(1)	(2)	(3)	(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	5.07	6.09	5.84	6.91
4 election year indicators	p=0.00	p=0.00	p=0.00	p=0.00
F-test on	16.02	0.08	21.09	3.77
$Mayor_{t-1} + Governor_{t-1} = 0$	p=0.00	p=0.78	p=0.00	p=0.05
F-test on	0.28	20.49	0.20	12.92
$Mayor_{t-2} + Governor_{t-2} = 0$	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.

Table 3. Further Estimates of the Elasticity of Crime with Respect to Police

Crime Categories Estimated Separately	**Published**			**Replication, Levitt Mayoral Measure**				**Replication, New Mayoral Measure**			
	2SLS	2SLS	LIML	2SLS	2SLS	LIML	LIML	2SLS	2SLS	LIML	LIML
Violent Crimes	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Murder	-2.09	-1.18	-1.98	-1.99	-1.37	-3.61	-3.16	-2.13	-1.34	-8.01	-5.12
	(0.64)	(0.39)	(0.59)	(1.34)	(0.83)	(2.31)	(1.56)	(1.61)	(0.95)	(6.46)	(2.63)
Rape	0.08	-0.11	-0.27	-0.02	0.36	-0.21	1.22	0.93	0.47	4.89	2.42
	(0.84)	(0.49)	(0.77)	(0.80)	(0.51)	(1.26)	(0.92)	(0.97)	(0.58)	(4.55)	(1.50)
Robbery	-0.38	-0.49	-0.79	-0.40	-0.50	-0.64	-0.95	-0.20	-0.22	0.21	-0.03
	(0.89)	(0.53)	(0.79)	(0.70)	(0.45)	(1.02)	(0.79)	(0.82)	(0.51)	(1.23)	(1.18)
Aggravated Assault	-0.36	-0.41	-1.09	-0.21	-0.53	-0.90	-1.96	-0.39	-0.60	-1.71	-4.23
	(0.81)	(0.50)	(0.73)	(0.86)	(0.53)	(1.48)	(1.01)	(1.01)	(0.61)	(2.18)	(1.94)
Property Crimes											
Burglary	-0.39	-0.11	-0.05	-0.27	-0.12	-0.31	0.07	0.00	0.17	0.68	1.57
	(1.06)	(0.62)	(0.90)	(0.49)	(0.32)	(0.69)	(0.55)	(0.60)	(0.37)	(1.29)	(1.02)
Larceny--ex MV Theft	0.06	-0.21	0.43	0.21	-0.22	0.55	-0.56	0.51	0.05	2.33	0.20
	(1.20)	(0.67)	(1.01)	(0.41)	(0.26)	(0.68)	(0.45)	(0.50)	(0.30)	(2.13)	(0.66)
Motor Vehicle Theft	0.14	-0.34	-0.50	0.22	0.04	0.83	0.65	-0.63	0.40	-2.20	3.02
	(0.89)	(0.53)	(0.80)	(0.69)	(0.45)	(1.11)	(0.79)	(0.83)	(0.52)	(1.73)	(1.61)
Pooled Estimates											
All Violent Crimes	-0.90	-0.65	-1.16	-0.42	-0.36	-1.01	-0.87	-0.14	-0.24	-0.68	-0.94
	(0.40)	(0.25)	(0.38)	(0.44)	(0.27)	(0.74)	(0.47)	(0.51)	(0.30)	(1.16)	(0.63)
All Property Crimes	-0.05	-0.24	-0.34	0.05	-0.14	0.32	-0.14	0.14	0.15	0.78	1.17
	(0.59)	(0.36)	(0.54)	(0.29)	(0.18)	(0.44)	(0.30)	(0.34)	(0.21)	(0.74)	(0.50)
Type of Weights:	Incorrect			Correct				Correct			
Source of mayoral instrument:	Levitt			Levitt				Author			
Instrument:	elections* city-size indicators	elections* region indicators	elections* region indicators	elections* city-size indicators	elections* region indicators	elections* city-size indicators	elections* region indicators	elections* city-size indicators	elections* region indicators	elections* city-size indicators	elections* region indicators
Numbers based on:	*AER* , June 1997			Author's calculations				Author's calculations			

Notes:

Table presents 2SLS and LIML estimates of the effect of police on crime for different crime categories. Specifications are identical to those in Table 1, but use more instruments. The instrument set is described in the second-to-last row, and is alternately the interaction of election year indicators with city-size indicators, and the interaction of election year indicators with region indicators. Columns (1) to (3) present published estimates; Columns (4) to (7) present replication estimates using Levitt's data set; Columns (8) to (11) present replication estimates using my own measure of mayoral election dates. Note that the replications use four columns, while the published estimates comprise three columns. The replication estimates include the LIML estimates using election year indicators interacted with city-size indicators, which were excluded from the published estimates. The quantities in columns (10) and (11) are *not* typographical errors, and underscore the low predictive power of the instrument set.

Table A.1. Discrepancies in Mayoral Election Measures, 1969-92

City	Number Elections: Levitt Measure	Number Elections: New Measure	Number Times (1), (2) Agree	Mean Number Times Election Date Observed in Sources	New Measure Lead by 1 Year?	Comment
	(1)	(2)	(3)	(4)	(5)	(6)
Birmingham, AL	6	6	6	5.00		Election typically in October to November
Mesa, AZ	12	12	0	1.25	Yes	Election typically in June
Phoenix, AZ	6	12	6	1.50		Four-year term beginning in November 1991 (two-year prior); election typically in November
Tucson, AZ	6	6	6	2.50		Election typically in November to December
Anaheim, CA	12	9	4	1.44	Yes (see comment)	Prior to 1986, election typically in April
Fresno, CA	6	6	0	2.00	Yes	Election typically in March to June
Los Angeles, CA	6	6	0	5.00	Yes	Election typically in June
Oakland, CA	6	5	1	2.40		April 1985 was last odd-year election; next election was June 1990 (phone call to city clerk); New measure misclassified by one year for one election (1990=1 instead of 1989=1); Levitt measure incorrect throughout
Sacramento, CA	6	7	0	1.71		Levitt measure off by one year throughout
San Diego, CA	6	6	4	3.00		Change in cycle between December 1975 and December 1980; election in December
San Francisco, CA	6	6	6	4.17		Election typically in November to January
San Jose, CA	6	6	6	2.00		Election typically in November to January
Denver, CO	6	6	0	5.00	Yes	Election typically in April to July
Washington, DC	6	5	5	4.80		First elected mayor November 1974 (home-rule adoption just prior 1974)
Jacksonville, FL	6	6	0	5.00	Yes	Election typically in April to July
Miami, FL	6	10	6	1.40		Four-year term beginning in November 1989 (two-year prior); New measure correct with exception of November 1987 election (newspaper verified); election in November
Saint Petersburg, FL	12	9	0	1.33	Yes	Four-year term beginning April 1981 (two-year prior); election typically in March to April
Tampa, FL	6	6	3	5.00	Yes (see comment)	Starting in 1983, election in March (previously in September to October)
Atlanta, GA	6	6	6	4.00		Election typically in November to January
Honolulu, HI	6	6	6	4.83		Election typically in November to January
Chicago, IL	6	7	0	4.29	Yes	Election in April
Indianapolis, IN	6	6	6	5.00		Election typically in November to December
Louisville, KY	6	6	6	5.00		Election typically in November to December
New Orleans, LA	6	6	0	4.83	Yes	Election typically in February to May
Boston, MA	6	6	0	5.00		Levitt measure off by two years throughout (newspaper verified)
Baltimore, MD	6	6	6	4.83		Election typically in November to December
Detroit, MI	6	7	5	4.00		Sources for new measure in error (newspaper verified)
Minneapolis, MN	6	8	4	3.75		Four-year term beginning in November 1981 (two-year prior); election in November
Saint Paul, MN	6	10	6	2.30		Four-year term beginning in November 1985 (two-year prior); election in November

(see next page)

Table A.1. Discrepancies in Mayoral Election Measures, 1969-92 (continued)

City	Number Elections: Levitt Measure	Number Elections: New Measure	Number Times (1), (2) Agree	Mean Number Times Election Date Observed in Sources	New Measure Lead by 1 Year?	Comment
	(1)	(2)	(3)	(4)	(5)	(6)
Kansas City, MO	6	6	0	3.33	Yes	Election typically in March to April
Saint Louis, MO	6	6	0	5.00	Yes	Election typically in March to April
Charlotte, NC	12	12	12	1.92		Election typically in November to December
Omaha, NE	6	6	0	5.00	Yes	Election typically in May to June
Jersey City, NJ	6	6	0	4.50	Yes	Election typically in May to July
Newark, NJ	6	6	0	5.00	Yes	Election typically in May to July
Albuquerque, NM	6	6	6	4.17		Election typically in October to December
Buffalo, NY	6	6	6	5.00		Election typically in November to December
New York, NY	6	6	6	5.00		Election typically in November to December
Akron, OH	6	6	6	5.00		Election typically in November to January
Cleveland, OH	6	9	6	3.33		Four-year term beginning November 1981 (two-year prior); election in November
Columbus, OH	6	6	6	4.17		Election typically in November to January
Toledo, OH	12	12	12	0.83		Election typically in November to January
Oklahoma City, OK	6	6	6	2.17		Election in April
Tulsa, OK	12	11	0	2.64	Yes	Election typically in April to May
Portland, OR	6	6	6	5.00		Election typically in November to December
Philadelphia, PA	6	6	6	4.67		Election typically in November to January
Pittsburgh, PA	6	6	6	4.83		Election typically in November to January
Memphis, TN	6	6	0	5.00	Yes	Election in November
Nashville, TN	6	6	6	4.83		Election typically in August to September
Arlington, TX	12	12	0	1.08	Yes	Election typically in April to May
Austin, TX	8	10	3	1.20	Yes	Three-year term beginning with the May 1988 election (two-year term prior); election typically in April to May
Corpus Christi, TX	12	12	0	0.75	Yes	Election typically in April to May
Dallas, TX	12	9	0	1.33	Yes	Four-year term beginning with April 1983 election; election typically in April to May
El Paso, TX	12	12	0	2.50	Yes	Election typically in April to May
Fort Worth, TX	12	12	0	1.00	Yes	Election typically in April to May
Houston, TX	12	12	12	2.08		Election typically in November to January
San Antonio, TX	12	12	0	1.67	Yes	Election typically in April
Seattle, WA	6	6	6	5.00		Election typically in November to January
Milwaukee, WI	6	7	0	4.29	Yes	Four-year term beginning in April 1976 (two-year prior); election in April

Notes:

Table compares the Levitt mayoral election year indicator with the new mayoral election year indicator. Column (3) gives the number of times the two measures agree, calculated as the sum (for a city) of the product of the indicators. Column (4) gives the average number of times an election date was observed in either the *Municipal Yearbook* or the *World Almanac*. As there is moderate measurement error in the sources, multiple observations are a great help. Confidence in the new measure is increasing in the number given. Note that I only attempted to collect a full series from the World Almanac; the Municipal Yearbook was used to corroborate the *World Almanac* for 1979 to 1983. Entries in column (5) answer the question: Does my mayoral election measure lead the Levitt measure by one year? Entries in column (6) give information on the month of election, which is relevant for defining the year in which the mayor would want to increase police hiring, as well as general comments on changes in the electoral cycle, term of office, etc.