Steven Levitt's "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime": Replication and Extensions *

Kei Irizawa †   Adam A. Oppenheimer ‡

March 2020

Abstract

Are police effective in maintaining public order and safety? Levitt (1997) approaches this question through the use of instrumental variables. Using election cycles as instruments for changes in police, Levitt finds that increased police presence causally reduces violent crime but does not have an effect on property crime. We replicate Levitt's results with the inclusion of the corrections suggested by McCrary (2002). We also consider the non-parametric bootstrap method to estimate standard errors in addition to 3SLS. Our replication results indicate that elections may be a weak instrument for changes in police, making our estimates biased.

We then investigate the effectiveness of police on crime using an unconventional theory. Most empirical literature seeks to show that police are effective by reducing crime levels in the long-run. However, we argue that short-run effectiveness should instead be measured by increases in the number of crimes reported: increases in the number of police on patrol consequently should increase the number of crimes reported. Using regression discontinuity, event study, and differences-in-differences (synthetic control method) designs, we take police academy graduations as an exogenous shock to police levels and find that this causes a short-term increase in reported crimes. We also verify these results using monthly crime data around mayoral elections in Chicago, finding that reported crime levels increase in the months after mayoral elections.

keywords: Police, Crime, IV, RD, Event Study, Synthetic Control DID

*We would like to thank Dr. Joensen for the guidance and support she provided while writing this paper.

†University of Chicago, Department of Economics. Email: kirizawa@uchicago.edu.

‡University of Chicago, Department of Economics. Email: oppenheimer@uchicago.edu.

1 Introduction

A better understanding of the effect of police on crime could aid policymakers in determining the optimal level of investment in police to maintain public order and safety. However, as argued by Levitt (1997), most historical research trying to estimate this effect falls prey to omitted variable bias caused by simultaneity - because changes in police are highly responsive to changes in crime, simple OLS will not be able to disaggregate which parts of the change in crime are caused by changes in police, and which parts of the change in police are caused by changes in crime. Levitt (1997) attempts to resolve this through an instrumental variables approach by using mayoral and gubernatorial election years as an instrument for changes in police. Levitt argues that election cycles are independent of crime (instrument exogeneity), and because crime levels are important when running for office, incumbents will increase levels of police during campaigns to increase their chance of re-election (instrument relevance). His results indicate that police causally reduce violent crime rates and have unclear effects on property crime rates. However, McCrary (2002) replicates Levitt's results and finds that he was miscalculating the weights for FGLS, causing his standard errors to be estimated incorrectly.

We replicate the main tables of Levitt (1997) and attempt to follow the corrections noted by McCrary (2002). Since we are concerned that the weighting of the errors through FGLS may not be accurately representing the structure of the data, we also consider a non-parametric bootstrap method to estimate standard errors in addition to 3SLS. We argue that Levitt's approach suffers from a weak instrument problem making our estimates biased. Even if the instruments are not weak, our results indicate there is no statistically significant effect of police on crime.

We then introduce three new extensions to attempt to estimate the causal effect of police on crime. Using historical daily-level Chicago crime and police academy graduation data, we compare the causal effect of graduations on crime using regression discontinuity, event study, and synthetic control method differences-in-differences designs. All estimation techniques arrive at the same conclusion: police academy graduations increase reported crime levels in the short-term. This result, while counterintuitive, may partially be a consequence of increased crime reporting because of the additional police on patrol. This explanation, if true, indicates that increasing the size of a police force is effective in stopping crime. Taking these criminals off the street may also explain some of the longer-term reduction in reported crimes found in other research. We also confirm such short-term results using monthly-level Chicago crime and mayoral election data to test Levitt's hypothesis about the effect of elections on crime in Chicago. Consistent with our estimates of the effect of graduations on crime, we find that mayoral elections in Chicago increase reported crime in the short-term.

Our paper is organized as follows: Section 2 replicates Levitt (1997). Section 3 describes our extensions and results. Section 4 concludes.

2 IV Estimation Method Using Levitt (1997)

2.1 Data

In order to replicate the results of Levitt (1997), we use data from McCrary (2001). This is the dataset used in McCrary (2002) to replicate Levitt's results. To assess the similarity of the datasets, we replicate Levitt's Table 1 in Table 1. The summary statistics between the two datasets are similar but not identical in most cases. There are two cases where the variation between the datasets is noticeable: welfare spending and education spending per capita. We therefore should not be surprised if our results differ from Levitt's. Despite these differences, our dataset still has large within-city variation in crime rates. If we are not concerned with measurement error, this indicates the Fixed-Effects estimator should still be efficient.

We then replicate Levitt's Figure 1. 1971 crime levels are indexed at 100. Our replication in Figure 1 shows the same trends pointed out by Levitt - violent and property crime closely follow each other until around the mid-1980s, when the increases in violent crime begin to outpace the increases in property crime. He also notes that levels of sworn police are relatively steady over the sample.

2.2 Instrument Relevance

Levitt begins his discussion of instrument relevance through two simple analyses of the relationship between elections and sworn police. He first looks at the mean change in police during election and nonelection years. Our replication results can be seen in Table 2. Levitt emphasizes the strong positive/weak relationship between election/nonelection years and change in police. This can also be seen in our replication.

Levitt does a similar analysis visually in his Figure 2. We replicate these results in the first panel of Figure 2. This figure seems to indicate a positive relationship between election years and change in police. One potential weakness with this figure is that the sample of cities in election and nonelection years changes over time, causing between-year comparisons of this figure to become difficult to interpret. The effects of this can be seen in the second panel, where we can see how susceptible this figure is to scaling: if we switch from change in ln(police) to percent change in police, there is no longer a clear relationship between election years and change in police.

Levitt then moves to a more formal treatment of instrument relevance. He considers the instrument relevance of election cycles on changes in the police force using the following model specifications:

$$\Delta \ln(P_{it}) = \theta_1 M_{it} + \theta_2 G_{it} + \gamma_t + \nu_{it} \quad \text{(Model 1)}$$

$$\Delta \ln(P_{it}) = \theta_1 M_{it} + \theta_2 G_{it} + \mathbf{X}_{it}\delta + \psi CS_i + \gamma_t + \nu_{it} \quad \text{(Model 2)}$$

$$\Delta \ln(P_{it}) = \theta_1 M_{it} + \theta_2 G_{it} + \mathbf{X}_{it}\delta + \psi CS_i + \gamma_t + \lambda_i + \nu_{it} \quad \text{(Model 3)}$$

$$violent\ \ crime_{it} = \theta_1 M_{it} + \theta_2 G_{it} + \mathbf{X}_{it}\delta + \psi CS_i + \gamma_t + \lambda_i + \nu_{it} \quad \text{(Model 4)}$$

$$property\ \ crime_{it} = \theta_1 M_{it} + \theta_2 G_{it} + \mathbf{X}_{it}\delta + \psi CS_i + \gamma_t + \lambda_i + \nu_{it} \quad \text{(Model 5)}$$

where i and t subscripts indicate city and time, respectively; P gives the number of sworn police; M is an indicator for a mayoral election year; G is an indicator for a gubernatorial election year; CS is an indicator for city size; γ gives a time fixed effect; λ gives a city fixed effect; $\mathbf{X}$ is a matrix of demographic variables; and ν contains all unobservables. Our replication output can be seen in Table 3.

Our first three model specifications indicate a relationship between mayoral and gubernatorial elections and change in police similar to that found in Levitt's paper. The results indicate a statistically significant positive relationship between elections and change in police. We also include F-statistics in our table. The F-statistics for the first two columns are barely above 10, and the F-statistic for the third column is below 10. We should therefore take some caution that election cycles may not be a relevant instrument.

The last two model specifications of our replication differ markedly from Levitt's. This may be related to differences in the datasets being used. Levitt's results indicate a negative relationship between elections and violent and property crime. Our results, however, indicate that there is not a clear relationship between elections and violent crime. While there is no relationship between mayoral elections and property crime, there is a statistically significant negative relationship between gubernatorial elections and property crime.

2.3 Empirical Strategy: IV Estimation

Levitt (1997) uses mayoral and gubernatorial election years as an instrument to estimate the causal effect of police on crime. Levitt estimates his regressions using 2SLS and corrects for heteroskedasticity using FGLS. We also consider a non-parametric bootstrap method to estimate our standard errors and compare these results with 3SLS.

2.3.1 First Stage Regression

We estimate the first stage using the following model specifications:

$$\ln(P_{ijt}) = \beta_0 + \theta_{1i} M_{it} + \theta_{2j} G_{it} + \mathbf{X}_{it}\eta_j + \gamma_{tj} + \lambda_i + \epsilon_{ijt} \quad \text{(Model 1)}$$

$$\ln(P_{ijt}) = \beta_0 + \theta_{1i} M_{it} * CS_i + \theta_{2j} G_{it} * CS_i + \mathbf{X}_{it}\eta_j + \gamma_{tj} + \lambda_i + \epsilon_{ijt} \quad \text{(Model 2)}$$

$$\ln(P_{ijt}) = \beta_0 + \theta_{1i} M_{it} * RG_i + \theta_{2j} G_{it} * RG_i + \mathbf{X}_{it}\eta_j + \gamma_{tj} + \lambda_i + \epsilon_{ijt} \quad \text{(Model 3)}$$

where all variables are as defined in Section 2.2, with the addition of the j subscript, which indicates crime category; RG, which indicates region; and ϵ, which is equivalent to ν. We have the same first stage for the lagged $\ln(P_{ijt-1})$ as well.

F-statistics and p-values for the joint significance of the instruments in the first stage regressions can be seen in Table 4. The F-statistics are all below 10 for the first stage regressions, indicating we may have an issue of weak instruments. These results are similar to those from the replication of Levitt's instrumental relevance table in Section 2.2. This may be caused by allowing the number of instruments to far exceed the number of endogenous regressors (especially for model specifications (3)-(5)).[1] Adding weak instruments through interactions moves the F-statistic toward zero because the instruments lack predictive power. This causes coefficient estimates to become biased. We also know that given weak instruments, estimates will converge towards OLS. Consistent with this, Levitt mentions that his 2SLS estimates get closer to his OLS estimates as he adds more instruments in model specifications (3)-(5).

While we are limited to annual police data using Levitt's dataset, we contacted the City of Chicago Police Department to access monthly police data for Chicago. The data is from July 2017 to January 2020. Figure 3 visualizes this data around the mayoral election in February 2019. We can see that the number of police decreased after the mayoral election. This contrasts with Levitt's theoretical model, which suggests that elections should increase police and consequently decrease crime. The change in trend is likely explained by the fact that the incumbent mayor was not running for re-election. This could partly explain why elections seem to be weak instruments for changes in police. A potential way to correct this concern would be to consider elections only if the incumbent is running for re-election.

2.3.2 Second Stage Regression

The impact of police on crime is estimated using the following model specification:

$$\Delta C_{ijt} = \beta_{1j}\Delta \ln(P_{ijt}) + \beta_{2j}\Delta \ln(P_{ijt-1}) + \mathbf{X}_{it}\eta_j + \gamma_{tj} + \lambda_i + \epsilon_{ijt}$$

where i, t, and j subscripts indicate city, year, and crime category, respectively; C gives the number of crimes per capita; P gives the number of sworn police; and $\mathbf{X}$ is the matrix of covariates including demographic variables, state and local spending controls, city-size indicators, and region and year dummies. We also have city fixed effects λ and crime specific year effects γ. City fixed effects control for between-city variations in crime across crime categories. Crime specific year effects control for national trends in crime categories.

Following Levitt (1997), restrictions are imposed on crime categories such that the sum of the contemporaneous and once-lagged coefficients for the effect of police on crime are equal for all violent crime categories and all property crime categories, but can differ between violent and property crimes. We also impose similar restrictions on the interactions of demographic and state and local spending controls with violent and

[1] Consider the first stage regression for both the contemporaneous and lagged sworn police. For model specification 2, we have number of city size indicators times two (mayor and gubernatorial election indicators) instruments. Similarly, for model specification 3, we have number of region indicators times two instruments.

property crime.[2] Specifically, we have the following constraints:

$$\beta_{11} = \beta_{12} = \beta_{13} = \beta_{14}$$

$$\beta_{15} = \beta_{16} = \beta_{17}$$

$$\beta_{21} = \beta_{22} = \beta_{23} = \beta_{24}$$

$$\beta_{25} = \beta_{26} = \beta_{27}$$

$$\eta_1 = \eta_2 = \eta_3 = \eta_4$$

$$\eta_5 = \eta_6 = \eta_7$$

$$\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4$$

$$\gamma_5 = \gamma_6 = \gamma_7$$

As in Levitt (1997), we report the sum of the contemporaneous and once-lagged coefficients for the effect of police on crime. As Levitt points out, this is an elasticity. Given the restrictions to the model specifications, we provide estimates of the impact of police for violent and property crime separately, but not for each individual crime type in our main replication tables. We also replicate Levitt's Table 5 in Appendix A.1.4, which removes the restrictions for the effect of police on individual crime types.

Levitt chooses to estimate heteroskedastic standard errors using FGLS. In order to verify these results, we also compute standard errors using bootstrap.

2.4 Empirical Results

Our replication results can be seen in Appendix A.1.4. The first two specifications in each table use OLS and not IV. For violent crime, our results for the effect of police on crime are comparable to Levitt's results. There appears to be a statistically significant positive relationship between levels of ln(police) and violent crime (Specification 1), and a statistically significant negative relationship between change in ln(police) and violent crime (Specification 2). However, given the issue of simultaneity bias mentioned by Levitt, these results cannot be interpreted as causal.

The third specification uses mayoral and gubernatorial elections as instruments for change in ln(police). Levitt's 2SLS results indicate a statistically significant negative relationship between change in police and violent crime. Our 2SLS replications indicate a larger negative relationship than Levitt's results, but neither the FGLS nor bootstrap standard errors indicate the results are statistically significant. Moreover, as noted before, there are concerns that elections are weak instruments as the first stage F-statistics are below 10. If this is the case, these results may be biased. The concerns are even larger for specifications four and five, which include more and potentially weaker instruments.

[2] As Stata's 2SLS functions cannot include restrictions, we manually compute our 2SLS estimates. We make sure to correct standard errors for the 2SLS estimates to account for the generated values.

The fourth specification uses election-city size interactions as instruments for change in ln(police). Levitt's 2SLS results once again indicate a statistically significant negative relationship between police and violent crime, although the effect is smaller in this specification than in the third specification. If we believe there is a weak instrument problem, it is not surprising that the results are getting closer to the OLS estimates. Our 2SLS replications differ noticeably depending on the standard errors used. For the FGLS estimates, the coefficient becomes positive but not statistically significant. For the bootstrap estimates, the coefficient remains negative but not statistically significant. Similar to Levitt's estimate, the coefficients becomes smaller in magnitude than the estimate from the third specification. Thus, for both the FGLS and bootstrap methods, we see the coefficient drifts towards the OLS estimate.

The fifth specification uses election-region interactions as instruments for change in ln(police). Levitt's 2SLS results continue to indicate a statistically significant negative relationship between police and violent crime, with an effect that is smaller in magnitude than the previous specification and closer to the OLS estimates. Contrary to Levitt's results, which continue to trend toward the OLS estimates, our estimates have magnitudes very near that of the third specification for both the FGLS and bootstrap methods.

We now turn to property crime. For the first two specifications, our results are comparable to Levitt's. There appears to be a statistically significant positive relationship between levels of ln(police) and property crime (Specification 1), and a statistically significant negative relationship between change in ln(police) and property crime (Specification 2). As with violent crime, the issue of simultaneity bias means these results cannot be interpreted as causal.

Beyond the first two specifications, the magnitudes of Levitt's estimates do not change much and lose their statistical significance. He therefore concludes there may be a weak negative relationship, but it is not statistically significant. Similarly, the estimates from our replication are not statistically significant for both 2SLS models. The estimates of the effect of police on crime for specification 4 are still negative but with a smaller magnitude which is much closer to the OLS estimate. However, the result is not statistically significant. Just as with violent crime, in Specification 5, coefficient estimates for the 2SLS models are negative with much larger magnitudes than in Specification 4. This violates our expectation that the estimates should be moving toward OLS estimates as we include more weak instruments and raises concerns that these models may be biased. We may also be concerned about bias arising from the possibility of FGLS not weighting errors properly given the structure of the data. However, the bootstrap method gives similar results. This indicates that random error may be playing a role in these results. This is supported by the large standard errors of the coefficient estimates, which consequently lead to large 95% confidence intervals.

Since we are concerned that FGLS is not weighting errors properly, we also consider three stage least squares (3SLS) estimation. This weights the errors by using the variance-covariance matrix estimated from 2SLS.[3] The results can be seen in Section A.1.4. Consistent with prior estimates, the relevant coefficients

[3] 3SLS works as follows: we begin by following 2SLS to run first-stage and second-stage regressions. We can then compute a consistent estimator of the variance-covariance matrix. In the third stage, we use the variance-covariance matrix estimated from the 2SLS regression as a weighting matrix to re-estimate the model using GLS.

are not statistically significant and do not trend towards the OLS estimates as we add more instruments.

3 Extensions

Considering the difficulties in identifying the causal effect of police on crime when using IV, our extensions attempt to measure this effect using three other identification strategies: regression discontinuity, event study, and differences-in-differences (synthetic control method). These identification strategies will use police academy graduations as the treatment for increases in the number of police in Chicago. We also look at an event study using Chicago mayoral elections to see if we get similar results using a different treatment.

3.1 Data[4]

Chicago crime data is gathered from Chicago Data Portal (2020). Crime is categorized at the individual-crime level by crime-type, location, and time for all crimes from 2001 to March 2020. We determine Chicago graduation dates from press releases available from the Office of the Mayor (2020). All graduation dates between 2016 and 2019 are considered. We also add the handful of graduations in 2014 and 2015 where we could find exact graduation dates. Figure 4 shows the cyclical nature of aggregate crime, with vertical lines denoting graduations. All our analyses consider the number of days from the nearest graduation. The ideal dataset would include all crimes; however, the drawback that crime data includes only reported crimes likely influences our results.

3.2 Extension One: Regression Discontinuity Design

3.2.1 Empirical Strategy

We first estimate the causal effect of police on crime in Chicago using a regression discontinuity design. Multiple times each year, new classes of potential police begin their six months of training in the police academy. After training, they graduate and begin field training. We consider the date of graduation as the point of discontinuity X_0. The number of police in the field should increase immediately after the graduation date.

Since the number of police increases immediately after this threshold, we use a sharp regression discontinuity design. The main identifying assumption is that in a sufficiently small neighborhood around the discontinuity X_0, crime should be continuous at the daily level. We test this assumption in Section 3.2.2. Moreover, we believe the treatment is as good as randomly assigned: because graduation dates are set far in advance of the window we are considering, changes in crime within this window will be independent of the date of graduation. There may be concerns that the actual number of police graduating could be endogenous,

[4] For further information about our data, please refer to our README.md file.

as the city government could choose to dramatically alter the number of police who can graduate very near the graduation date. For instance, if a city has many recruits but crime levels drop between recruitment and graduation, they may choose to graduate only a small number of the original recruits. However, Chicago is a unique example where this is not a concern. Hinkel (2017) indicates that between 2013 and 2017, Chicago graduated 97% of its recruits, compared to a national average of 86%. Therefore, we can be confident that the number of graduates will be independent of any short-term changes in crime. There may also be concerns that criminals are shifting crime on either side of graduations to reduce their probability of getting caught. This could, for instance, be caused by clustering of senior officer retirements around graduations. This could reduce the quality of police after graduations without increasing the aggregate number of police, potentially explaining the increase in crimes just after graduations. We test this anticipation effect in our event study in Section 3.3.

We construct RD estimates by fitting the following model:

$$Y_t = \alpha + \rho D_t + f_1(X_t) + f_2(X_t * D_t) + f_3(\mathbf{Z}_t) + f_4(\mathbf{Z}_t * D_t) + \epsilon_t$$

where Y_t is $\ln(Crime_t)$, the running variable X_t indicates number of days from the date of discontinuity X_0, $\mathbf{Z}_t$ is a matrix of date covariates, including date, month, year, and day of week, and $D_t = \mathbb{1}\{X_t > X_0\}$. We consider graduations from 2014 to 2020. We approximate $f_1(X_t)$, $f_2(X_t * D_t)$, $f_3(\mathbf{Z}_t)$, and $f_4(\mathbf{Z}_t * D_t)$ using third order polynomials. Including interactions with D_t allows the polynomial coefficient estimates to differ on either side of the discontinuity. The following explains how we constructed this model:

Consider the following model specifications before and after the graduation date:

$$E[Y_{0t}|X_t] = \alpha + \beta_{01}\tilde{X}_t + \beta_{02}\tilde{X}_t^2 + \beta_{03}\tilde{X}_t^3$$
$$E[Y_{1t}|X_t] = \alpha + \rho + \beta_{11}\tilde{X}_t + \beta_{12}\tilde{X}_t^2 + \beta_{13}\tilde{X}_t^3$$

The regression model that can be used to estimate the causal effect of D_t is a deterministic function of X_t:

$$E[Y_t|X_t] = E[Y_{0t}|X_t] + (E[Y_{1t}|X_t] - E[Y_{0t}|X_t])D_t$$

The regression model we estimate becomes the following:

$$Y_t = \alpha + \rho D_t + \beta_1\tilde{X}_t + \beta_2\tilde{X}_t^2 + \beta_3\tilde{X}_t^3 + \beta_1^*\tilde{X}_t D_t + \beta_2^*\tilde{X}_t^2 D_t + \beta_3^*\tilde{X}_t^3 D_t + \epsilon_t$$

where $\beta_j^* = \beta_{1j} - \beta_{0j}$ for all $j = 1, 2, 3$.

Daily crime is highly cyclical. This can be seen in $\ln(crime)$ in Figure 4. In order to control for this cyclicality, we use third order polynomial estimations of the date controls. Including all relevant regressors and controls gives us the following full model specification:

$$
\begin{aligned}
Y_t = \alpha &+ \rho D_t + \beta_1 \tilde{X}_t + \beta_2 \tilde{X}_t^2 + \beta_3 \tilde{X}_t^3 + \beta_1^* \tilde{X}_t D_t + \beta_2^* \tilde{X}_t^2 D_t + \beta_3^* \tilde{X}_t^3 D_t \\
&+ w_1 Day_t + w_2 Day_t^2 + w_3 Day_t^3 + w_4 Day_t * Month_t + w_5 Day_t^2 * Month_t + w_6 Day_t^3 * Month_t \\
&+ w_7 Day_t * Year_t + w_8 Day_t^2 * Year_t + w_9 Day_t^3 * Year_t + w_{10} Day_t * Month_t * Year_t \\
&+ w_{11} Day_t^2 * Month_t * Year_t + w_{12} Day_t^3 * Month_t * Year_t + w_{13} DOW_t + \epsilon_t
\end{aligned}
$$

where we consider the observed dates as a continuous variable Day_t and interact this with $Month_t$ and $Year_t$ indicators. We also control for the day of the week, DOW_t. These controls well approximate the cyclical nature of crime over time. We can see their fit for 30 day windows around graduations in Figure 5.

3.2.2 Testing Identifying Assumption

To test the credibility of our identifying assumption, we run the RD regression on non-graduation dates. To limit the number of tables included, we restrict the sample to regressions using windows of 30 days. Because we are considering windows of 30 days, we include results for start days at ± 15 and ± 30 days from the actual graduation dates to ensure the start days considered are sufficiently far from the actual graduations in their respective windows. Results can be seen in Appendix A.2.2. We are most interested in model specifications 3 and 4, as these control for the cyclicality in the data. We can see that for each of these specifications and for each start day, our coefficient estimate for $Post$ is not statistically significant. We therefore cannot reject the null hypothesis that crime data is continuous over time for non-graduation dates, lending credibility to our identifying assumption.

3.2.3 Empirical Results

From Frisch-Waugh theorem, we know the following:

$$
\begin{aligned}
Y_t = &(\alpha_0 + \rho D_t + \beta_1 \tilde{X}_t + \beta_2 \tilde{X}_t^2 + \beta_3 \tilde{X}_t^3 + \beta_1^* \tilde{X}_t D_t + \beta_2^* \tilde{X}_t^2 D_t + \beta_3^* \tilde{X}_t^3 D_t) \\
&+ (\gamma_1 Day_t + \gamma_2 Day_t^2 + \gamma_3 Day_t^3 + \gamma_4 Day_t * Month_t + \gamma_5 Day_t^2 * Month_t + \gamma_6 Day_t^3 * Month_t \\
&+ \gamma_7 Day_t * Year_t + \gamma_8 Day_t^2 * Year_t + \gamma_9 Day_t^3 * Year_t + \gamma_{10} Day_t * Month_t * Year_t \\
&+ \gamma_{11} Day_t^2 * Month_t * Year_t + \gamma_{12} Day_t^3 * Month_t * Year_t + \gamma_{13} Month_t + \gamma_{14} Year_t + \gamma_{15} DOW_t) + \epsilon_t
\end{aligned}
$$

So, we have the following:

$$\Leftrightarrow \boldsymbol{Y} = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{W}\hat{\gamma} + \hat{\boldsymbol{U}}$$
$$\Leftrightarrow M_{\boldsymbol{W}}\boldsymbol{Y} = M_{\boldsymbol{W}}\boldsymbol{X}\hat{\boldsymbol{\beta}} + M_{\boldsymbol{W}}\boldsymbol{W}\hat{\gamma} + M_{\boldsymbol{W}}\hat{\boldsymbol{U}}$$
$$\Leftrightarrow M_{\boldsymbol{W}}\boldsymbol{Y} = M_{\boldsymbol{W}}\boldsymbol{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{U}}$$
$$\Leftrightarrow \hat{\boldsymbol{\beta}} = ((M_{\boldsymbol{W}}\boldsymbol{X})'(M_{\boldsymbol{W}}\boldsymbol{X}))^{-1}(M_{\boldsymbol{W}}\boldsymbol{X})'M_{\boldsymbol{W}}\boldsymbol{Y}$$

Therefore, we can obtain the estimate of $\hat{\boldsymbol{\beta}}$ by the following procedure:

1) Regress each regressor in $\boldsymbol{X}$ on all regressors in $\boldsymbol{W}$. Denote the residual as $\tilde{\boldsymbol{X}}_{\boldsymbol{W}}$.

2) Regress $\boldsymbol{Y}$ on all regressors in $\boldsymbol{W}$. Denote the residual as $\tilde{\boldsymbol{Y}}_{\boldsymbol{W}}$.

3) Regress $\tilde{\boldsymbol{Y}}_{\boldsymbol{W}}$ on $\tilde{\boldsymbol{X}}_{\boldsymbol{W}}$ and obtain the estimate of $\hat{\boldsymbol{\beta}}$.

Figure 6: Model Specification Fit

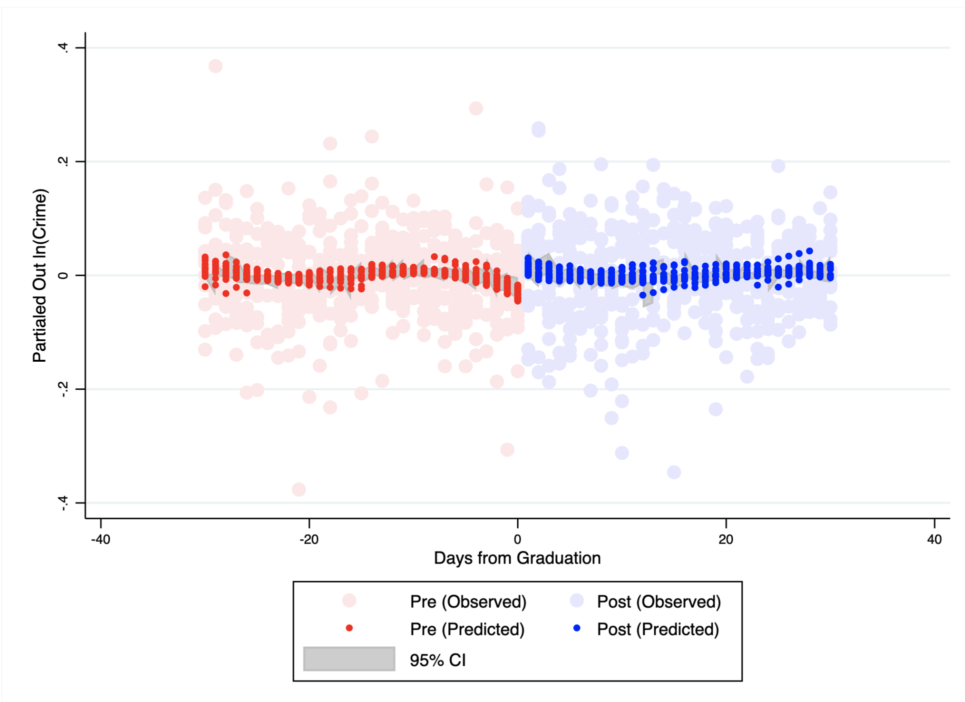


Figure 6 displays the predicted value of $M_{\boldsymbol{W}}\boldsymbol{Y}$ given each observed crime Y_t within a 30 day window of X_0 and controlling for days away from graduation, Day_t, $Month_t$, $Year_t$, DOW_t, and all interactions and polynomial terms. Notice that for each day from graduation $\tilde{X}_t$ there are multiple predicted values of $M_{\boldsymbol{W}}\boldsymbol{Y}$. This can be explained by the three step process, described above, that predicts $M_{\boldsymbol{W}}\boldsymbol{Y}$. This process requires information about $\boldsymbol{W}$. Since we are using data from graduations that occur at different dates, $\boldsymbol{W}$ will be different for graduations that have the same $\tilde{X}_t$, leading to different predictions of $M_{\boldsymbol{W}}\boldsymbol{Y}$. With this considered, our model appears to fit the data well after controlling for cyclicality, and we can now see a clear discontinuity in crimes after graduations.

Our regression estimates can be found in Appendix A.2.3. We consider four model specifications for a range of bandwidths. For the third and fourth specifications, the coefficient of interest, *Post*, is positive and significant for most bandwidths between 10 and 30 days.[5] For specification 4 in particular, the 95% confidence intervals overlap for every bandwidth between 10 and 30 days. The statistically significant coefficient estimates from this sample range from 0.050 to 0.073, indicating a very tight estimate that is robust to different bandwidths.

When interpreting these results, we must consider the tradeoff between bias and variance when choosing bandwidth. Smaller bandwidths lead to less biased but more imprecise estimates: if we consider only crime observations near the graduation date, we should have confidence that the treatment of graduation is as good as randomly assigned. This will give us unbiased results. However, smaller intervals will include less observations, leading to imprecise estimates. On the other hand, when considering bandwidths larger than 30 days, we lose precision on the *Post* coefficient. This could for instance be caused by systematic changes in the environment relevant to criminals over long periods of time.

We finally test for the percent of the variation in crimes that is explained by graduations in specification 4 after partialling out all variables except *Post*. Our estimates can be seen in Table 19. These results look at the R^2 specifically for data after graduations. We can see that for bandwidths between 10 and 30 days, *Post* is explaining between 1 and 1.5% of the variation in crimes.

These results seem somewhat surprising. Common insight into the causal effect of police on crime would indicate there should be a negative effect. How can we resolve the positive effect we are finding? We must consider how crime data is collected: we can only have crime data if the crime is reported. As police academy graduations will increase the number of police on patrol, it seems plausible that this will increase the number of crimes observed by police, subsequently increasing the number of crimes reported in the data. While this would make it appear as if crime rates are increasing, it would actually be a reflection of an effective (or overzealous) police force. However, we cannot discount the possibility that crimes are actually increasing after graduations or that some other factor is at play without further investigation of the data.

[5] We are interested in the third and fourth specifications as controlling for cyclicality is necessary with high frequency crime data.

3.3 Extension Two: Event Study Design

3.3.1 Empirical Strategy

We consider the following model for daily Chicago crime data around graduations:

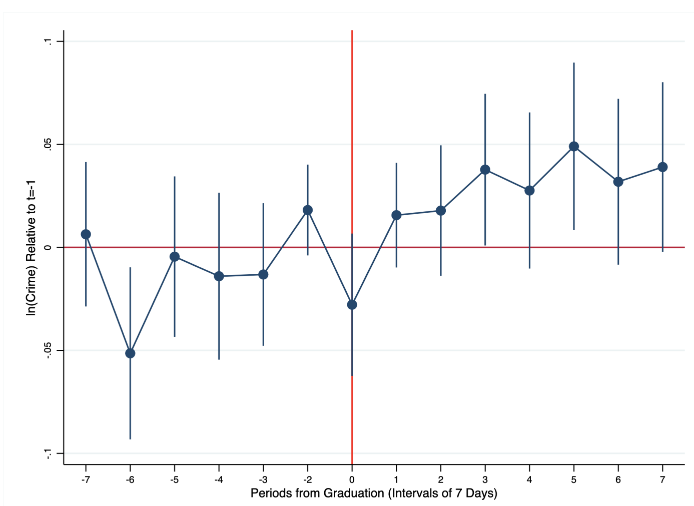
$$
\begin{aligned}
Y_t = \alpha + \sum_{\tau} \sigma_\tau D_{\tau,t} + \gamma_t \\
&+ \gamma_1 Day_t + \gamma_2 Day_t^2 + \gamma_3 Day_t^3 + \gamma_4 Day_t * Month_t + \gamma_5 Day_t^2 * Month_t + \gamma_6 Day_t^3 * Month_t \\
&+ \gamma_7 Day_t * Year_t + \gamma_8 Day_t^2 * Year_t + \gamma_9 Day_t^3 * Year_t + \gamma_{10} Day_t * Month_t * Year_t \\
&+ \gamma_{11} Day_t^2 * Month_t * Year_t + \gamma_{12} Day_t^3 * Month_t * Year_t + \gamma_{13} Month_t + \gamma_{14} Year_t + \gamma_{15} DOW_t + \epsilon_t
\end{aligned}
$$

where Y_t is $\ln(Crime_t)$. The vector $D_{\tau,t}$ is composed of separate indicator variables for each bin before and after the graduation date, excluding the bin just prior to graduation. Each bin is an indicator for a range of seven days. The bins range from -7 (for 42-49 days before graduation) to 7 (for 42-49 days after graduation). Bin 0 is composed of only the graduation date. The σ_τ's measure the average crime level in their corresponding bin after controlling for the cyclicality of $\ln(Crime_t)$.

3.3.2 Empirical Results

We construct a function that estimates the event study coefficients given a bin size and the number of bins before and after the graduation date to be considered. The bin just prior to graduation is excluded from the regression in order to use it as a reference. Event study results can be seen in Figure 7 for a threshold of 7 bins, each including 7 days except for bin 0, which includes only the day of graduation.

Figure 7: Event Study: Graduation Effect on Crime



Regression results can be found in Appendix A.2.4. Similar to the RD design estimates, we see a positive effect of graduation on crimes reported. Moreover, these results indicate that the positive effect may be slowly increasing over time and eventually leveling off in the long-run.

At the end of Section 3.2.1, we raised concerns that there may be an anticipation effect where criminals shift crime to just after graduations in anticipation of a decrease in the quality of police after graduations. From our event study results, we can see that there is no anticipation effect in the data.

3.4 Extension Three: Differences-in-Differences (Synthetic Control Method) Design

3.4.1 Empirical Strategy

We consider a differences-in-differences model for daily Chicago crime data around the graduation on September 23, 2018. Based on the Office of the Mayor (2020), we know that nearly 100 police officers were deployed to the 9 districts in Chicago as listed in Table 21. We consider these districts to be in the treatment group and measure the average aggregate crime before and after graduation. We use a synthetic control method to determine which other districts in Chicago should be included in the control group. We follow the algorithm proposed by Abadie et al. (2010) to estimate weights for each non-treatment group district such that we have similar trends for pre-graduation crime data. Instead of selecting weights such that the vector of controls have similar values between the treatment and control group in the pre-treatment period, we choose weights to match crime levels. The estimated weights for each district can be seen in Table 22.

3.4.2 Empirical Results

Partialled out $\ln(crime)$ around the graduation date of interest can be seen in Figure 8. We use the controls described in Section 3.2.1 to control for cyclicality in the data. We can see that the high variance in daily data makes it somewhat difficult to find a good synthetic control group. Daily differences in $\ln(crime)$ between the treatment and control groups can be seen in Figure 9. From this figure, we can see that the fit for the synthetic control group is not excellent prior to the graduation. However, more information can be extracted by looking at the cumulative differences in crime between the treatment and control groups: this can be seen in Figure 10. In the period prior to graduation, the cumulative difference between the treatment and control groups tends to stay around zero. Therefore, while the synthetic control group may not fit the treatment group at a granular level, it does do a good job at fitting the general trend of reported crime. We then see a divergence after the graduation: similar to our previous results, we see a positive effect of graduation on reported crimes.

3.5 Extension Four: Event Study Design (Mayoral Election)

3.5.1 Empirical Strategy

We consider the following model for monthly Chicago crime data around mayoral elections:

$$\begin{aligned} Y_t = \alpha + \sum_{\tau} \sigma_\tau D_{\tau,t} + \gamma_t \\ + \gamma_1 Month\_Year_t + \gamma_2 Month\_Year_t^2 + \gamma_3 Month\_Year_t^3 + w_4 Month\_Year_t * Year_t \\ + \gamma_5 Month\_Year_t^2 * Year_t + \gamma_6 Month\_Year_t^3 * Year_t + \gamma_7 Month_t + \gamma_8 Year_t + \epsilon_t \end{aligned}$$

where Y_t is $\ln(Crime_t)$ and $Month\_Year_t$ is continuous in time. The vector $D_{\tau,t}$ is composed of separate indicator variables for each month before and after the election month. The indicators range from -12 (for 12 months before the election) to 12 (12 months after the election). The σ_τ's measure the average crime level in the months before and after the election after controlling for the cyclicality of $\ln(Crime_t)$. We can see the monthly cyclicality of crime in the first panel of Figure 11. The fit of our controls around elections can be seen in the second panel.

3.5.2 Empirical Results

Event study results can be found in Appendix A.4.2 for the 6 months before and after mayoral elections. Results include a figure and regression results. Similar to our graduation results, although no longer statistically significant, we see a positive effect of mayoral elections on crime that slowly increases over time and eventually levels off in the long-run.

4 Conclusion

This paper seeks to advance our understanding of the effect of police on crime. We begin by replicating the results of Levitt (1997) including the changes noted by McCrary (2002). Following Levitt, we use election cycles as an instrument for changes in the number of police to overcome simultaneity between police and crime. Our first stage results indicate that elections may be a weak instrument for changes in police. However, even if we overlook this potential source of bias, our results differ from Levitt's: we find that police do not have a statistically significant effect on either violent or property crime. To verify that the weighting of errors in FGLS properly represents the structure of the data, we also consider the non-parametric bootstrap method to compute standard errors in addition to 3SLS. All three methods give similar results. Regardless of statistical significance, any results using this method must be taken with caution because of the concern of bias caused by weak instruments.

We then introduce an unconventional approach to argue for the effectiveness of police: short-term increases in reported crimes after increases in levels of police. We argue that because crime data is all reported, if crime reports increase, it may be an indication that police are more effective at stopping crimes. We propose three new strategies to estimate this causal effect: regression discontinuity, event study, and synthetic control differences-in-differences designs around police academy graduations. All estimation techniques indicate that increases in the number of police increase reported crime rates in the short term. Moreover, we respond to potential concerns that criminals may alter their behavior around graduations through our event study. Our event study estimates indicate there is not an anticipation effect in the data. Finally, we test a reduced form event study of Levitt's hypothesis about mayoral elections using Chicago data. We find that similar to our graduation results, there is a short-term increase in crimes after mayoral elections.

While these results seem surprising, and may even seem somewhat reasonable, they inevitably lead to questions about the credibility of crime data. If the increases in crime we observe in the data are caused by increases in reporting but not actual crime levels, this introduces new concerns that reported crime levels do not accurately reflect changes in absolute crime over short periods of time. Moreover, we should keep in mind that our conclusions specifically reflect short-term effects for overall crime levels. It may be that there is a feedback effect from the short-term increases in crime reports that causes long-term reductions in crimes. This would be consistent with past research. However, if this is the case, what portion of these long-term effects comes from taking criminals off the street and what portion comes from deterrence? Further, do these effects differ by particular crime categories? In consideration of these concerns, we must look into new empirical methods to more rigorously investigate the credibility of crime data and the long-term effects and the sources of these effects of police on crime. It is our hope that future research will attempt to resolve these issues.

References

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

Chicago Data Portal, "Crimes - 2001 to Present," https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2/data 2020. Accessed: 2020-03-04.

Hinkel, Dan, "Chicago Police Recruits Rarely Flunk Out, Raising Concerns about Training," *Chicago Tribune*, 2017. Accessed: 2020-03-15.

Levitt, Steven D., "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, 1997, *87* (3).

McCrary, Justin, "Replication of Steven Levitt, AER, 1997, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime"," https://eml.berkeley.edu/replications/mccrary/index.html 2001. Accessed: 2020-02-20.

_ , "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment," *American Economic Review*, 2002, *92* (4).

Office of the Mayor, "Mayor's Press Releases," https://www.chicago.gov/city/en/depts/mayor/press_room/press_releases.html 2020. Accessed: 2020-03-08.

A Appendix

A.1 Replication

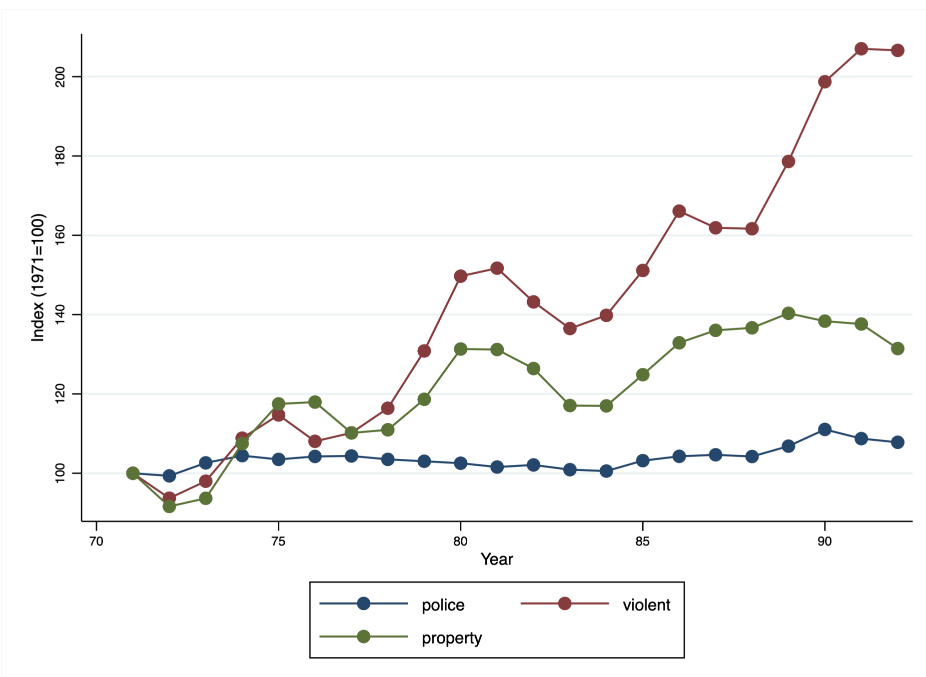
A.1.1 Data

Table 1: Summary Statistics (All Values Per 100,000 Residents Except Population)

		Standard deviation			
Variable	Mean	Within-city	Across cities	Minimum	Maximum
Population	718,043	59,190	1,039,385	85,000	7,896,000
Violent crime	1,159	203	685	103	4,353
Murder	19	3	12	0.6	81
Rape	68	9	31	4	199
Assault	522	128	353	42	2,386
Robbery	557	98	378	26	2,338
Property crime	7,748	615	2,112	2,707	16,739
Burglary	2,308	149	728	600	4,994
Larceny	4,365	400	1,398	550	10,003
Motor vehicle theft	1,074	221	695	165	5,369
Sworn officers	237	10	99	70	781
Percent black	23.1	2.0	18.0	0.1	78.2
Percent female-headed households	15.0	0.7	4.3	5.8	31.9
Percent ages 15-24	17.1	0.4	2.1	11.5	25.3
Public welfare spending per capita (1992 dollars)	381.6	10.4	74.3	148.3	556.2
Education spending per capita (1992 dollars)	501.7	5.8	70.6	230.1	684.9
State unemployment rate	6.7	0.4	2.0	2.0	15.5

Notes: All variables except population are per 100,000 residents. The sample used is a set of 59 large US cities with directly elected mayors over the period 1970-1992.

Figure 1: Trends in Crime and Police



A.1.2 Instrumental Relevance

Table 2: Change in Police over Election and Nonelection Years

	Gubernational	Mayoral	No election
	(1)	(2)	(3)
Δ ln sworn police officers per capita	0.018***	0.014***	0.006***
	(0.004)	(0.003)	(0.002)
N	1,278	1,278	1,278
R^2	0.021	0.016	0.010

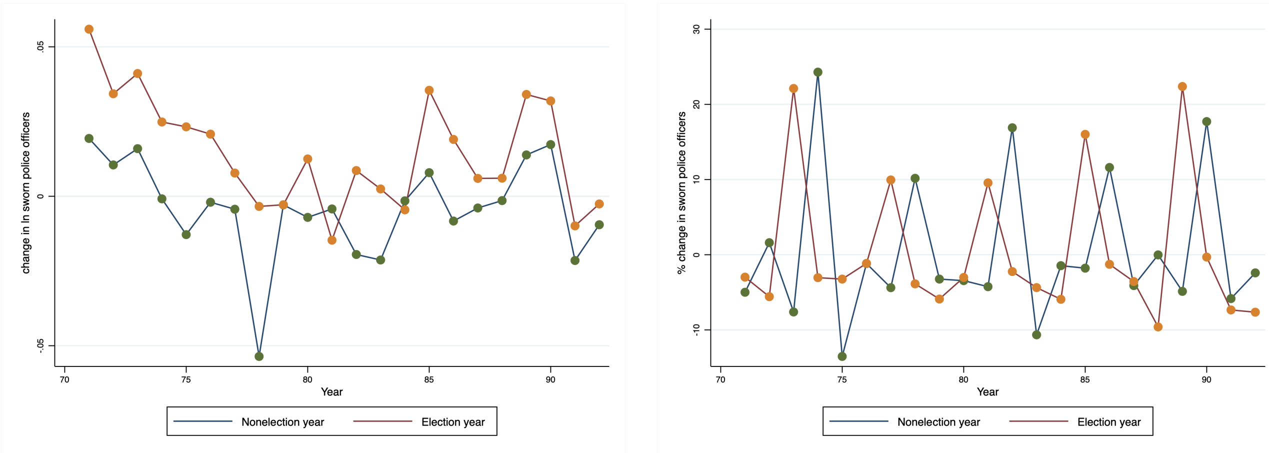
Robust standard errors in parentheses.

Table 3: Levitt Table 2 Replication

	Δln Sworn officers			Violent crime	Property crime
	(1)	(2)	(3)	(4)	(5)
Mayoral election year	0.013***	0.012***	0.010***	0.007	0.001
	(0.004)	(0.004)	(0.004)	(0.008)	(0.005)
Gubernational election year	0.025***	0.021***	0.021***	–0.012	–0.021***
	(0.007)	(0.007)	(0.007)	(0.011)	(0.008)
Δln Public welfare spending per capita		0.013	0.014	0.027	0.047
		(0.010)	(0.011)	(0.038)	(0.030)
Δln Education spending per capita		0.318***	0.359***	0.315***	0.307***
		(0.042)	(0.045)	(0.078)	(0.055)
Δ State unemployment rate		–0.317	–0.360	0.107	1.121***
		(0.252)	(0.256)	(0.428)	(0.277)
Δ(Percent ages 15-24 in SMSA)		0.418	4.979**	–1.475	5.384
		(2.175)	(2.498)	(5.433)	(3.520)
Δ(Percent black)		–0.009	–0.009	–0.015	–0.007
		(0.006)	(0.010)	(0.017)	(0.011)
Δ(Percent female-headed households)		0.011	0.020	–0.021	0.025
		(0.012)	(0.018)	(0.036)	(0.024)
Constant	0.029***	0.025**	0.042***	0.062	–0.008
	(0.008)	(0.012)	(0.015)	(0.041)	(0.025)
N	1,278	1,276	1,276	1,264	1,271
R^2	0.075	0.153	0.195	0.245	0.413
Adj. R^2	0.058	0.131	0.134	0.187	0.368
Year indicators?	Yes	Yes	Yes	Yes	Yes
City-size indicators?	No	Yes	Yes	Yes	Yes
City-fixed effects?	No	No	Yes	Yes	Yes
F-stat: Joint significance of election years?	13.81	11.23	9.00	1.15	3.84
P-value: Joint significance of election years?	<0.0001	<0.0001	0.0001	0.3171	0.0218

Robust standard errors in parentheses.

Figure 2: Yearly Changes in Police over Election and Nonelection Years



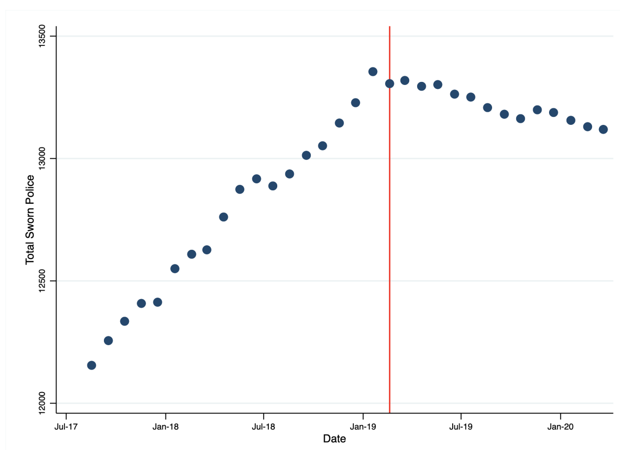
A.1.3 Empirical Strategy: IV Estimation

Table 4: First Stage Results

	Concurrent			Once-Lagged		
	(3)	(4)	(5)	(3)	(4)	(5)
F-stat	5.52	2.28	1.69	5.48	1.96	1.68
p-val	0.0002	0.0025	0.0066	0.0002	0.0121	0.0071

Calculated using robust standard errors.

Figure 3: Effect of Mayoral Election on Police (Chicago)



A.1.4 Empirical Results

Table 5: Violent Crime 2SLS (FGLS)

	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)
ln Sworn officers per capita	0.263***	-0.099*	-2.282	0.370	-2.914
	(0.036)	(0.057)	(3.496)	(2.575)	(2.094)
State unemployment rate	−0.207	−0.226	−0.222	−0.911	0.220
	(0.290)	(0.285)	(0.885)	(0.730)	(0.623)
ln Public welfare spending per capita	−0.116***	−0.020	−0.030	−0.019	−0.029
	(0.018)	(0.017)	(0.057)	(0.047)	(0.039)
ln Education spending per capita	0.147***	0.281***	0.100	0.414*	0.004
	(0.031)	(0.043)	(0.321)	(0.238)	(0.196)
Percent ages 15-24 in SMSA	0.645	−2.391	−1.228	−3.571	−0.698
	(0.742)	(3.070)	(5.041)	(4.608)	(4.916)
Percent black	0.008***	−0.009	−0.018	−0.008	−0.019
	(0.002)	(0.009)	(0.019)	(0.016)	(0.016)
Percent female-headed households	0.005	−0.013	0.017	−0.015	0.027
	(0.004)	(0.020)	(0.052)	(0.042)	(0.040)
N	8,456	7,945	8,883	8,883	8,883
Data differenced?	No	Yes	Yes	Yes	Yes
Instrument	None	None	Election	Election* CS	Election* RG

Standard errors in parentheses.

Table 6: Violent Crime 2SLS (Bootstrap)

	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)
ln Sworn officers per capita	0.308***	-0.177*	-2.717	-1.757	-2.263
	(0.057)	(0.091)	(4.175)	(2.116)	(2.107)
State unemployment rate	−0.028	−0.173	0.759	0.068	0.663
	(0.302)	(0.291)	(0.726)	(0.784)	(0.561)
ln Public welfare spending per capita	−0.061**	−0.004	0.005	−0.006	0.008
	(0.026)	(0.028)	(0.048)	(0.039)	(0.035)
ln Education spending per capita	0.049	0.265***	−0.102	0.098	−0.052
	(0.041)	(0.062)	(0.296)	(0.217)	(0.171)
Percent ages 15-24 in SMSA	1.645*	−1.071	3.947	1.610	3.511
	(0.983)	(3.844)	(4.620)	(4.625)	(4.129)
Percent black	0.009***	−0.013	−0.019	−0.017	−0.017
	(0.002)	(0.012)	(0.016)	(0.014)	(0.013)
Percent female-headed households	0.008	0.007	0.032	0.024	0.026
	(0.006)	(0.021)	(0.043)	(0.036)	(0.028)
N	8,456	7,945	8,883	8,883	8,883
Data differenced?	No	Yes	Yes	Yes	Yes
Instrument	None	None	Election	Election* CS	Election* RG

Standard errors in parentheses.

Table 7: Property Crime 2SLS (FGLS)

	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)
ln Sworn officers per capita	0.099***	-0.130***	-2.551	-0.333	-3.018
	(0.035)	(0.043)	(3.076)	(2.350)	(1.863)
State unemployment rate	1.857***	1.423***	1.786***	1.436**	1.789***
	(0.271)	(0.226)	(0.681)	(0.577)	(0.501)
ln Public welfare spending per capita	−0.028*	0.026	0.039	0.059	0.029
	(0.016)	(0.019)	(0.051)	(0.043)	(0.037)
ln Education spending per capita	0.245***	0.305***	0.025	0.240	0.008
	(0.030)	(0.036)	(0.272)	(0.209)	(0.170)
Percent ages 15-24 in SMSA	5.268***	−1.254	1.911	0.540	1.444
	(0.727)	(2.568)	(4.113)	(3.772)	(4.038)
Percent black	−0.001	−0.020***	−0.025	−0.016	−0.028**
	(0.002)	(0.008)	(0.016)	(0.014)	(0.013)
Percent female-headed households	−0.007*	0.019	0.039	0.013	0.052
	(0.004)	(0.017)	(0.046)	(0.038)	(0.035)
N	8,456	7,945	8,883	8,883	8,883
Data differenced?	No	Yes	Yes	Yes	Yes
Instrument	None	None	Election	Election* CS	Election* RG

Standard errors in parentheses.

Table 8: Property Crime 2SLS (Bootstrap)

	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)
ln Sworn officers per capita	0.279***	-0.172***	-1.871	-0.331	-1.230
	(0.052)	(0.053)	(2.999)	(1.789)	(1.948)
State unemployment rate	1.360***	0.997***	1.630**	1.164***	1.363***
	(0.423)	(0.297)	(0.511)	(0.380)	(0.290)
ln Public welfare spending per capita	−0.020	0.014	0.040	0.047	0.040
	(0.023)	(0.032)	(0.042)	(0.036)	(0.034)
ln Education spending per capita	0.217***	0.276***	0.026	0.216	0.121
	(0.035)	(0.043)	(0.193)	(0.163)	(0.143)
Percent ages 15-24 in SMSA	6.461***	0.377	4.500	2.943	3.592
	(0.968)	(3.339)	(3.639)	(3.187)	(2.925)
Percent black	−0.000	−0.018	−0.025	−0.021*	−0.024
	(0.002)	(0.011)	(0.012)	(0.011)	(0.012)
Percent female-headed households	−0.001	0.015	0.039	0.023	0.033
	(0.006)	(0.019)	(0.034)	(0.026)	(0.030)
N	8,456	7,945	8,883	8,883	8,883
Data differenced?	No	Yes	Yes	Yes	Yes
Instrument	None	None	Election	Election* CS	Election* RG

Standard errors in parentheses.

Table 9: Violent Crime 3SLS

	IV	IV	IV
	(3)	(4)	(5)
ln Sworn officers per capita	-0.165	-0.187	-0.230
	(0.529)	(0.401)	(0.267)
State unemployment rate	–0.039	–0.092	0.062
	(0.397)	(0.386)	(0.371)
ln Public welfare spending per capita	–0.008	–0.009	–0.007
	(0.028)	(0.027)	(0.026)
ln Education spending per capita	0.289***	0.297***	0.267***
	(0.076)	(0.067)	(0.060)
Percent ages 15-24 in SMSA	0.658	0.404	0.972
	(4.474)	(4.365)	(4.255)
Percent black	–0.001	–0.001	–0.002
	(0.016)	(0.014)	(0.013)
Percent female-headed households	0.005	0.005	0.005
	(0.031)	(0.029)	(0.027)
N	1,129	1,129	1,129
Data differenced?	Yes	Yes	Yes
Instrument	Election	Election* CS	Election* RG

Standard errors in parentheses.

Table 10: Property Crime 3SLS

	IV	IV	IV
	(3)	(4)	(5)
ln Sworn officers per capita	-0.160	-0.149	-0.212
	(0.489)	(0.351)	(0.219)
State unemployment rate	0.952***	0.951***	1.031***
	(0.307)	(0.287)	(0.278)
ln Public welfare spending per capita	0.018	0.019	0.018
	(0.022)	(0.020)	(0.020)
ln Education spending per capita	0.304***	0.305***	0.287***
	(0.063)	(0.053)	(0.046)
Percent ages 15-24 in SMSA	0.061	–0.050	0.349
	(3.740)	(3.549)	(3.470)
Percent black	–0.009	–0.009	–0.010
	(0.014)	(0.013)	(0.012)
Percent female-headed households	0.016	0.016	0.016
	(0.028)	(0.025)	(0.023)
N	1,129	1,129	1,129
Data differenced?	Yes	Yes	Yes
Instrument	Election	Election* CS	Election* RG

Standard errors in parentheses.

Table 11: Levitt Table 5 Replication 2SLS (FGLS)

	OLS	OLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Murder	0.534***	-0.544***	-2.333	1.057	-3.304
	(0.045)	(0.142)	(4.116)	(3.232)	(2.748)
Rape	-0.231***	0.029	-2.390	0.676	-1.828
	(0.045)	(0.115)	(4.002)	(3.095)	(2.548)
Assault	0.056	0.159*	-1.900	0.932	-2.928
	(0.043)	(0.094)	(3.887)	(2.942)	(2.312)
Robbery	0.640***	-0.254**	-3.545	-0.474	-3.698
	(0.044)	(0.111)	(3.995)	(3.082)	(2.513)
Burglary	0.091**	-0.216***	-3.115	-0.721	-2.977
	(0.039)	(0.073)	(3.451)	(2.681)	(2.038)
Larceny	-0.069*	-0.050	-2.987	-0.737	-3.617*
	(0.039)	(0.062)	(3.360)	(2.594)	(1.980)
Motor vehicle	0.714***	-0.230**	-0.582	1.717	-1.215
theft	(0.048)	(0.102)	(3.552)	(2.810)	(2.265)
Data differenced?	No	Yes	Yes	Yes	Yes
Instrument	None	None	Election	Election* CS	Election* RG

Standard errors in parentheses.

Table 12: Levitt Table 5 Replication 2SLS (Bootstrap)

	OLS	OLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Murder	0.627***	-0.567**	-2.393	-1.480	-2.330
	(0.069)	(0.238)	(4.290)	(2.401)	(2.772)
Rape	-0.199***	-0.011	-1.438	-0.418	-0.694
	(0.064)	(0.139)	(4.140)	(2.482)	(2.779)
Assault	0.156**	0.158	-3.088	-2.209	-2.847
	(0.077)	(0.119)	(4.709)	(2.299)	(2.542)
Robbery	0.649***	-0.286***	-3.938	-2.887	-3.164
	(0.058)	(0.099)	(4.411)	(2.573)	(2.595)
Burglary	0.136***	-0.210**	-2.618	-1.036	-1.873
	(0.053)	(0.082)	(3.974)	(2.179)	(1.801)
Larceny	-0.033	-0.060	-2.462	-0.893	-1.872
	(0.050)	(0.069)	(3.860)	(2.125)	(1.619)
Motor vehicle	0.739***	-0.247**	-0.527	0.949	0.059
theft	(0.055)	(0.105)	(3.704)	(2.346)	(2.027)
Data differenced?	No	Yes	Yes	Yes	Yes
Instrument	None	None	Election	Election* CS	Election* RG

Standard errors in parentheses.

A.2 Extension One: Regression Discontinuity Design

A.2.1 Empirical Strategy

Figure 4: Cyclical Nature of Crime (Chicago)

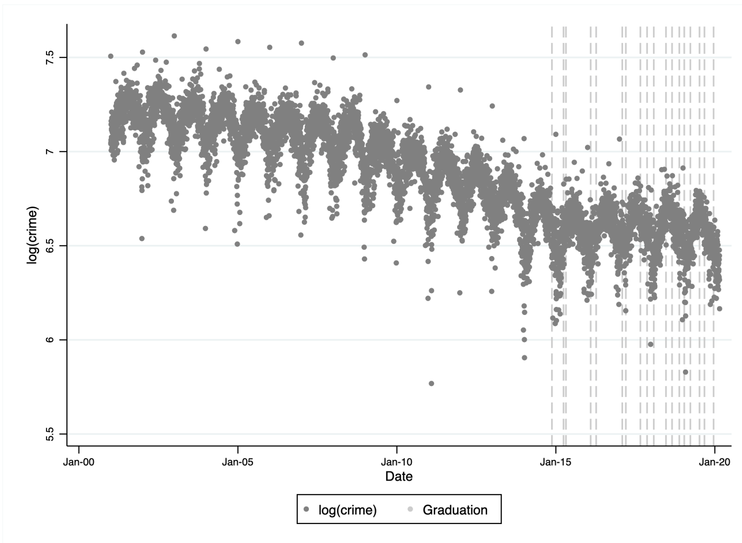
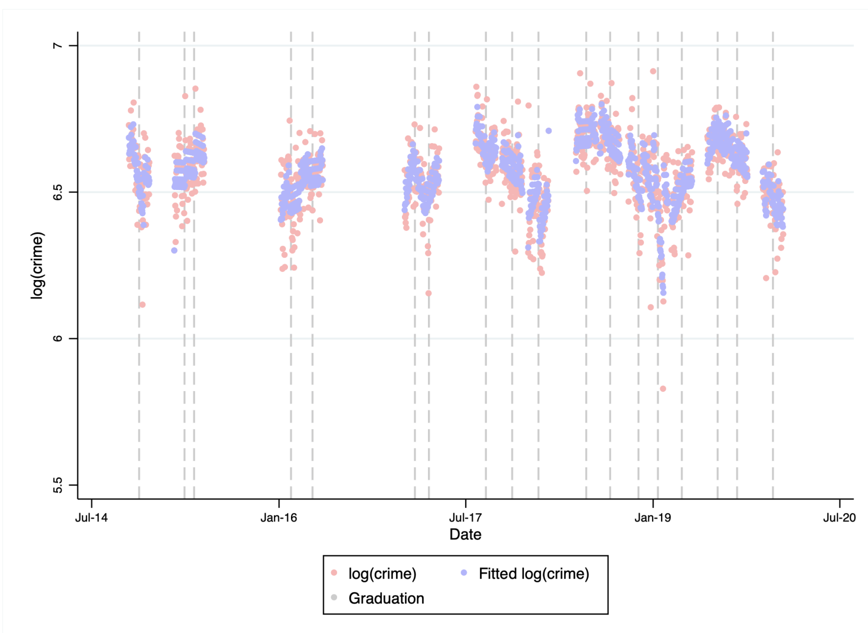


Figure 5: Cyclicality Fit



A.2.2 Testing Identifying Assumptions

Table 13: Regression Discontinuity Identifying Assumptions - 15 day window

	Start at 15 days before graduation				Start at 15 days after graduation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	−0.004	−0.055***	−0.102	−0.004	0.014*	0.024	0.115	0.251
	(0.009)	(0.018)	(0.065)	(0.246)	(0.007)	(0.019)	(0.071)	(0.208)
Days from Start / 100		0.258***	1.017*	−0.166		−0.223***	−0.039	−0.183
		(0.063)	(0.596)	(2.935)		(0.062)	(0.097)	(0.171)
Post * (Days from Start / 100)		−0.520***	−1.014*	0.073		0.182*	−0.601	−2.064
		(0.097)	(0.598)	(2.936)		(0.094)	(0.584)	(2.377)
(Days from Start / 100)^2			3.596***	−0.798			1.760**	1.711*
			(1.349)	(10.571)			(0.896)	(0.896)
Post * (Days from Start / 100)^2			−0.701	3.679			0.173	6.064
			(1.719)	(10.679)			(1.440)	(8.593)
(Days from Start / 100)^3				−5.030				8.047
				(11.341)				(7.456)
Post * (Days from Start / 100)^3				11.381				−14.686
				(14.329)				(12.191)
Constant	6.563***	35.380***	2,854.279***	2,796.525***	6.559***	37.549***	2,099.727***	2,111.768***
	(0.007)	(8.213)	(706.847)	(725.077)	(0.005)	(8.634)	(473.924)	(477.549)
N	954	954	954	954	952	952	952	952
R^2	0.000	0.480	0.627	0.627	0.004	0.457	0.620	0.621
Adj. R^2	−0.001	0.465	0.585	0.584	0.003	0.441	0.577	0.577

Notes: Robust standard errors in parentheses.

Table 14: Regression Discontinuity Identifying Assumptions - 30 day window

	Start at 30 days before graduation				Start at 30 days after graduation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	0.024**	0.082**	0.091	−1.704	−0.001	0.066**	0.153	−1.792
	(0.010)	(0.038)	(0.225)	(1.082)	(0.009)	(0.028)	(0.254)	(1.245)
Days from Start / 100		−0.092	0.254	13.215*		0.125***	−1.047	−1.063
		(0.076)	(1.275)	(7.616)		(0.046)	(0.882)	(0.977)
Post * (Days from Start / 100)		0.097	0.129	−13.498*		−0.259***	−0.872	12.921
		(0.097)	(1.211)	(7.597)		(0.077)	(1.315)	(8.891)
(Days from Start / 100)^2			1.738	32.473*			0.792	0.902
			(1.292)	(17.199)			(1.020)	(4.056)
Post * (Days from Start / 100)^2			−0.660	−37.144**			1.162	−30.576
			(1.712)	(17.326)			(1.837)	(21.040)
(Days from Start / 100)^3				22.775*				0.011
				(12.689)				(8.099)
Post * (Days from Start / 100)^3				−36.790**				23.582
				(14.914)				(17.697)
Constant	6.553***	28.135***	2,757.997***	3,299.593***	6.558***	29.699***	2,010.603**	2,052.837**
	(0.009)	(8.975)	(906.136)	(893.767)	(0.005)	(8.898)	(910.310)	(915.055)
N	790	790	790	790	787	787	787	787
R^2	0.008	0.509	0.643	0.653	0.000	0.500	0.606	0.608
Adj. R^2	0.006	0.493	0.597	0.606	−0.001	0.482	0.554	0.554

Notes: Robust standard errors in parentheses.

A.2.3 Empirical Result

Table 15: Regression Discontinuity

	5 day bandwidth				10 day bandwidth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	0.006	0.045	0.056	0.047	−0.029**	0.048***	0.033	0.072*
	(0.017)	(0.029)	(0.049)	(0.110)	(0.012)	(0.018)	(0.027)	(0.043)
Days from Graduation / 100		−0.787	−2.452	1.282		−0.767***	−0.026	−2.332
		(0.580)	(1.825)	(4.099)		(0.186)	(0.699)	(1.600)
Post * (Days from Graduation / 100)		−0.245	1.657	−2.179		−0.333	−1.298	−1.218
		(0.876)	(4.505)	(15.229)		(0.300)	(1.228)	(3.708)
(Days from Graduation / 100)^2			−26.811	174.034			8.753	−52.008
			(33.480)	(195.083)			(6.130)	(37.072)
Post * (Days from Graduation / 100)^2			19.208	−186.096			−4.014	104.274
			(64.825)	(563.340)			(11.322)	(69.011)
(Days from Graduation / 100)^3				2,649.318				−407.878*
				(2,482.906)				(241.139)
Post * (Days from Graduation / 100)^3				−2491.430				122.144
				(6,331.585)				(524.941)
Constant	6.555***	−7.698	−1758.974	−2176.286	6.567***	−22.946	246.123	300.572
	(0.010)	(33.030)	(2,144.199)	(2,096.163)	(0.007)	(16.716)	(933.313)	(940.759)
N	198	198	198	198	378	378	378	378
R^2	0.001	0.502	0.675	0.685	0.015	0.533	0.667	0.671
Adj. R^2	−0.004	0.430	0.555	0.557	0.012	0.500	0.599	0.600

Notes: Robust standard errors in parentheses.

Table 16: Regression Discontinuity

	15 day bandwidth				20 day bandwidth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	−0.032***	0.007	0.063***	0.073**	−0.023**	−0.027**	0.045**	0.069**
	(0.010)	(0.016)	(0.023)	(0.032)	(0.009)	(0.013)	(0.019)	(0.028)
Days from Graduation / 100		−0.447***	−0.510	−0.060		−0.146*	−0.494	−0.261
		(0.124)	(0.422)	(0.848)		(0.086)	(0.300)	(0.613)
Post * (Days from Graduation / 100)		0.273	−1.332*	−2.720		0.320**	−0.192	−1.822
		(0.200)	(0.786)	(1.700)		(0.128)	(0.487)	(1.164)
(Days from Graduation / 100)^2			0.105	8.020			−0.969	2.389
			(2.642)	(12.469)			(1.472)	(6.977)
Post * (Days from Graduation / 100)^2			12.726**	18.897			6.549**	18.958
			(5.343)	(25.972)			(2.563)	(12.993)
(Days from Graduation / 100)^3				35.860				9.964
				(53.827)				(22.566)
Post * (Days from Graduation / 100)^3				−94.694				−56.446
				(107.260)				(38.819)
Constant	6.575***	20.057	−2488.641***	−2456.538***	6.576***	37.097***	−1771.268**	5,593.040***
	(0.006)	(14.490)	(836.975)	(823.840)	(0.006)	(9.929)	(720.706)	(1,603.349)
N	555	555	555	555	725	725	725	725
R^2	0.018	0.454	0.645	0.646	0.009	0.468	0.645	0.661
Adj. R^2	0.016	0.426	0.591	0.590	0.008	0.447	0.598	0.613

Notes: Robust standard errors in parentheses.

Table 17: Regression Discontinuity

	25 day bandwidth				30 day bandwidth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	–0.021**	–0.034***	0.033*	0.064***	–0.017**	–0.037***	0.020	0.050**
	(0.008)	(0.012)	(0.017)	(0.024)	(0.007)	(0.012)	(0.016)	(0.022)
Days from Graduation / 100		–0.011	–0.545**	–0.600		0.030	–0.267	–1.073***
		(0.067)	(0.237)	(0.493)		(0.049)	(0.230)	(0.395)
Post * (Days from Graduation / 100)		0.126	0.452	–0.855		0.089	0.343	0.579
		(0.091)	(0.393)	(0.844)		(0.068)	(0.362)	(0.660)
(Days from Graduation / 100)^2			–1.417	–1.881			0.221	–6.855**
			(0.910)	(4.648)			(0.797)	(3.299)
Post * (Days from Graduation / 100)^2			3.648**	17.572**			0.618	12.148**
			(1.602)	(7.924)			(1.151)	(5.612)
(Days from Graduation / 100)^3				–1.133				–16.353**
				(12.365)				(7.815)
Post * (Days from Graduation / 100)^3				–33.797*				7.342
				(19.408)				(11.521)
Constant	6.577***	34.100***	2,002.758**	1,812.342**	6.576***	32.376***	288.742	536.734
	(0.005)	(8.818)	(864.172)	(902.550)	(0.005)	(7.454)	(621.238)	(619.533)
N	885	885	885	885	1,037	1,037	1,037	1,037
R^2	0.007	0.477	0.649	0.652	0.005	0.475	0.626	0.631
Adj. R^2	0.006	0.461	0.608	0.611	0.004	0.461	0.587	0.592

Notes: Robust standard errors in parentheses.

Table 18: Regression Discontinuity

	35 day bandwidth				40 day bandwidth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	–0.013*	–0.038***	0.017	0.032	–0.008	–0.039***	0.017	0.026
	(0.007)	(0.012)	(0.016)	(0.020)	(0.007)	(0.011)	(0.015)	(0.019)
Days from Graduation / 100		0.076*	–0.181	–0.673*		0.157***	–0.229	–0.364
		(0.046)	(0.255)	(0.367)		(0.041)	(0.230)	(0.309)
Post * (Days from Graduation / 100)		0.027	0.256	0.550		–0.133**	0.342	0.248
		(0.062)	(0.361)	(0.582)		(0.057)	(0.349)	(0.509)
(Days from Graduation / 100)^2			0.209	–3.770			–0.012	–0.901
			(0.808)	(3.004)			(0.681)	(1.963)
Post * (Days from Graduation / 100)^2			0.701	5.859			0.778	3.176
			(1.000)	(4.351)			(0.824)	(3.272)
(Days from Graduation / 100)^3				–7.860				–1.529
				(6.637)				(3.743)
Post * (Days from Graduation / 100)^3				5.768				–1.075
				(8.937)				(6.003)
Constant	6.574***	38.925***	8,222.434***	8,144.857***	6.568***	37.214***	8,291.282***	8,290.293***
	(0.005)	(7.074)	(2,044.055)	(2,057.701)	(0.005)	(6.834)	(2,051.770)	(2,057.720)
N	1,167	1,167	1,167	1,167	1,269	1,269	1,269	1,269
R^2	0.003	0.469	0.607	0.608	0.001	0.458	0.601	0.601
Adj. R^2	0.002	0.457	0.570	0.570	0.000	0.446	0.564	0.563

Notes: Robust standard errors in parentheses.

Table 19: Regression Discontinuity - Percent Variation of Crime Explained by Graduation

	5 day bandwidth	10 day bandwidth	15 day bandwidth	20 day bandwidth	25 day bandwidth	30 day bandwidth	35 day bandwidth	40 day bandwidth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
N	90	180	268	353	433	509	573	624
R^2	0.001	0.011	0.016	0.015	0.015	0.009	0.005	0.004
Adj. R^2	–0.010	0.005	0.012	0.013	0.012	0.007	0.003	0.002

Notes: Robust standard errors in parentheses.

A.2.4 Event Study: Graduation

Table 20: Event Study: Graduation Model Specification Estimates

	lcrime
gradMinus7	0.00636
	(0.0179)
gradMinus6	-0.0514**
	(0.0213)
gradMinus5	-0.00450
	(0.0199)
gradMinus4	-0.0140
	(0.0207)
gradMinus3	-0.0132
	(0.0177)
gradMinus2	0.0181
	(0.0112)
grad0	-0.0278
	(0.0176)
gradPlus1	0.0156
	(0.0130)
gradPlus2	0.0178
	(0.0162)
gradPlus3	0.0377**
	(0.0188)
gradPlus4	0.0276
	(0.0193)
gradPlus5	0.0490**
	(0.0207)
gradPlus6	0.0319
	(0.0205)
gradPlus7	0.0390*
	(0.0210)
N	6995
R^2	0.916
Adj. R^2	0.909

Robust standard errors in parentheses.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.

A.3 Diff-in-Diff (Synthetic Control Method) Design: Graduation

A.3.1 Empirical Strategy

Table 21: Districts of Police Deployed

District	Deployed Officers
District 1 (Central)	8
District 2 (Wentworth)	8
District 4 (South Chicago)	10
District 6 (Gresham)	10
District 7 (Englewood)	10
District 9 (Deering)	10
District 11 (Harrison)	8
District 18 (Near North)	8
District 24 (Rogers Park)	5

Table 22: Weights for Synthetic Control

District	Weights
District 3	0.08
District 5	0.081
District 8	0.532
District 9	0
District 10	0
District 12	0.071
District 14	0
District 15	0
District 16	0
District 17	0
District 19	0.156
District 20	0
District 22	0.039
District 25	0.043

A.3.2 Empirical Results

Figure 8: Diff-in-Diff: Around Graduation on September 23, 2018

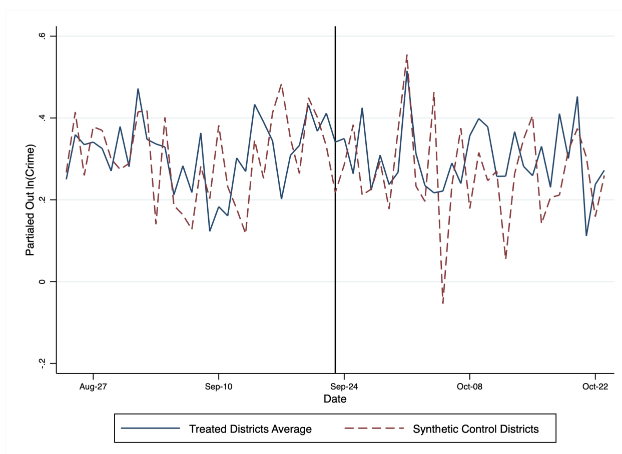


Figure 9: Diff-in-Diff: Around Graduation on September 23, 2018 (Difference)

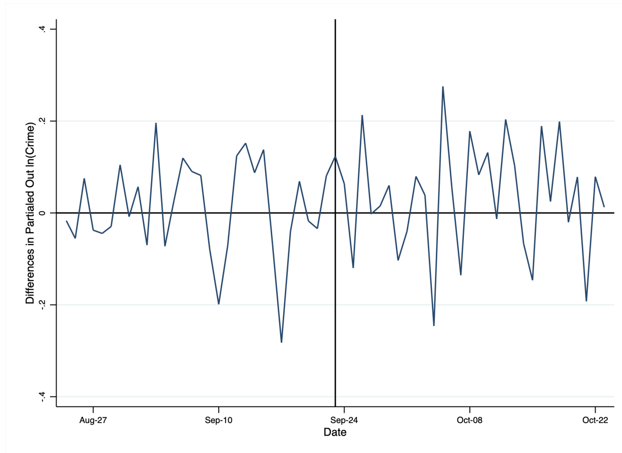
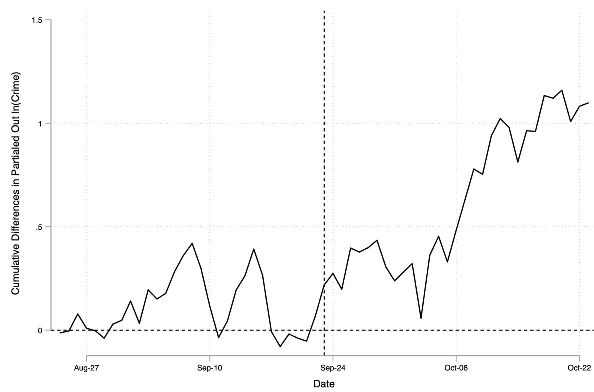


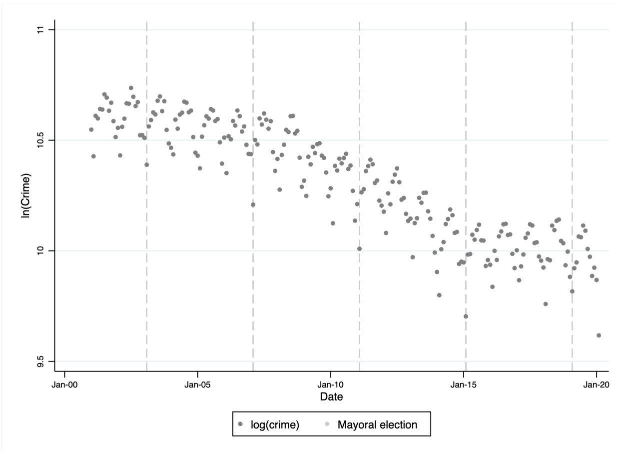
Figure 10: Diff-in-Diff: Around Graduation on September 23, 2018 (Cumulative Difference)

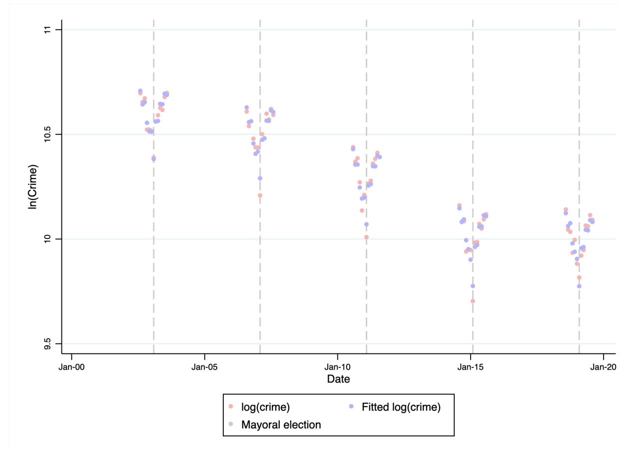


A.4 Event Study Design: Mayoral Election

A.4.1 Empirical Strategy

Figure 11: Monthly Crime Cyclicality and Fit





A.4.2 Empirical Results

Figure 12: Event Study: Mayoral Election Effect on Crime

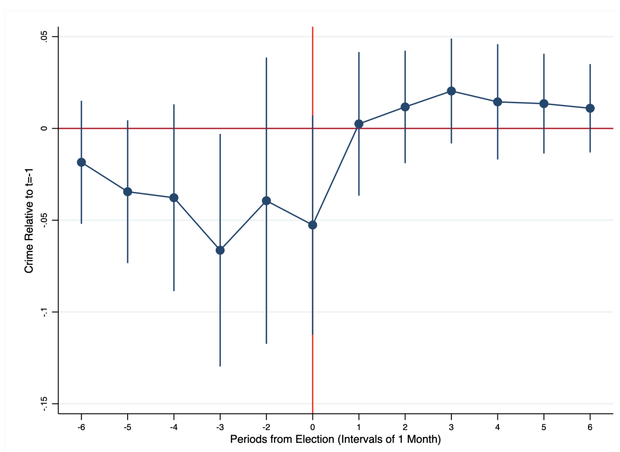


Table 23: Event Study: Mayoral Election Model Specification Estimates

	lcrime
electionMinus6	-0.0184
	(0.0170)
electionMinus5	-0.0345*
	(0.0197)
electionMinus4	-0.0377
	(0.0258)
electionMinus3	-0.0664**
	(0.0321)
electionMinus2	-0.0394
	(0.0395)
election0	-0.0526*
	(0.0302)
electionPlus1	0.00247
	(0.0198)
electionPlus2	0.0117
	(0.0155)
electionPlus3	0.0204
	(0.0145)
electionPlus4	0.0145
	(0.0159)
electionPlus5	0.0135
	(0.0138)
electionPlus6	0.0110
	(0.0122)
N	230
R^2	0.991
Adj. R^2	0.988

Robust standard errors in parentheses.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$.