More COPS, Less Crime*

Steven Mello
Princeton University
Industrial Relations Section
Firestone Library A-16-H-2
Princeton, NJ 08544
smello@princeton.edu

January 3, 2017

Abstract

I exploit a unique natural experiment to estimate the causal effect of police on crime. The American Recovery and Reinvestment Act increased funding for the COPS hiring grant program from $20 million in 2008 to $1 billion in 2009 and over $150 million annually in 2010-2013. During this period, grant applications were scored and funding was allocated according to a fuzzy cutoff rule. I leverage quasi-random variation in grant receipt by comparing the change over time in police and crimes for cities above and below the score threshold. Relative to low-scoring applicants, cities above the cutoff experience increases in police levels of about 3.6% and decreases in violent and property crimes of about 4.8% and 3%, respectively. The effects are driven by large and statistically significant effects of police on robbery, larceny, and auto theft. I also find evidence that police reduce murders, with the point estimate implying that one life can be saved by hiring eleven officers. Arrest rates do not increase with police force expansions, suggesting a deterrence mechanism underlying the crime reductions. The program passes a cost-benefit test under some assumptions but not others. The results highlight that police hiring grants may offer higher benefit-cost ratios than other stimulus spending.

JEL Classification: K42, H76.
Keywords: Police, crime, deterrence.

*I am grateful to Ilyana Kuziemko and Alex Mas, who provided considerable advice and encouragement on this project. I thank Jessica Brown, John Donohue, and Felipe Goncalves, who read earlier drafts and offered valuable insights and criticisms. Mingyu Chen, David Cho, Janet Currie, Will Dobbie, Hank Farber, Andrew Langan, Chris Neilson, and participants of the Princeton Public Finance Working Group provided helpful comments. An Online Data Appendix, which describes the processing of the data in detail, is available at www.princeton.edu/∼smello/papers/CopsDataAppendix.pdf. I acknowledge financial support from a Princeton University Graduate Fellowship. Any errors are my own.

1 Introduction

In February 2009, President Obama signed into law the American Recovery and Reinvestment Act (ARRA), which provided for over $490 billion in stimulus spending between 2009 and 2011. The Recovery Act allocated about $2 billion to the Department of Justice (DOJ), the majority of which was used to finance a reinvigoration of the DOJ's police hiring grant program. The Community Oriented Policing Services (COPS) hiring program, which covers the salary cost of new police hires for local law enforcement agencies, was a cornerstone of President Clinton's Violent Crime Control and Law Enforcement Act of 1994. Between 1995 and 2005, the COPS hiring program spent almost $5 billion to help local police departments hire about 64,000 officers (Evans and Owens 2007). Allocations for the program fell from over $1 billion per year in the late 1990's to almost zero in the years 2005–2008. The injection of Recovery Act funding restored the COPS hiring program budget to $1 billion in FY 2009, and allocations for the program remained above $150 million annually through 2013.

I rely on variation in police levels generated by the program's rebirth, termed COPS 2.0, to estimate the effect of police on crime.[1] Crime is estimated to cost Americans over $200 billion per year, and local government expenditures on police protection exceed $87 billion annually (Chalfin 2016). Given that provision of public safety is a key responsibility of local governments, and that hiring additional police is the main policy instrument for crime prevention, the causal effect of expanding police forces on crime rates is a parameter of substantial interest. In practice, estimating this effect is made difficult by the fact that police hiring decisions are endogenous to local crime conditions, which introduces simultaneity bias in OLS estimates.[2]

Beginning with Levitt (1997), researchers have tried to overcome endogeneity issues by relying on quasi-experimental research designs. Instruments used in the literature include mayoral election years (Levitt 1997, McCrary 2002), firefighter hiring (Levitt 2002), federal policing grants (Evans and Owens 2007, Worrall and Kovandzic 2010), terror alert levels (Klick and Tabarrok 2005), and state sales tax rates (Lin 2009). Quasi-experimental studies have consistently documented that police reduce crime, although estimated magnitudes vary widely across papers.[3] Further, these instruments are not without potential flaws. Binary instruments, such as election years, discard most of the variation in police rates and are often weak by conventional standards. Federal grant instruments suffer from the the possibility that these grants are targeted where they

[1] To the best of my knowledge, the term was coined by David Muhlhausen in a report for the Heritage Foundation titled *Why Would COPS 2.0 Succeed when COPS 1.0 Failed?*

[2] See, e.g., Klick and Tabarrok (2010) for further discussion.

[3] For a more complete summary of the existing literature on the police-crime relationship, see reviews by Levitt and Miles (2006), Klick and Tabarrok (2010), and Chalfin and McCrary (2016b).

are most needed or most likely to succeed, either of which would violate the exclusion restriction. Papers using sharp micro-time series variation in police presence generated by terror alert systems or terrorist attacks, including Klick and Tabarrok (2005), Draca, Machin and Witt (2011), and Di Tella and Schargrodsky (2004), provide convincing evidence that police deter property crimes. However, these studies estimate treatment effects specific to single jurisdictions, raising questions of external validity (Klick and Tabarrok 2010).

In this paper, I exploit a unique natural experiment generated by the scale up the COPS grant program. Beginning in 2009, hiring grants were awarded based on an open solicitation application process. Local law enforcement agencies applied for funding, and the COPS office scored the applications and distributed funds. A fuzzy cutoff rule was used in selecting winners, with the probability of grant receipt jumping discontinuously at a state by year-specific application score threshold. I leverage this feature of the program for identification. My primary empirical strategy, which is similar to the dynamic regression discontinuity approach implemented in Cellini, Ferreira and Rothstein (2010), is to compare the change over time in police and crimes for cities with scores exceeding the threshold with those below. The approach exploits the discontinuous allocation rule while still allowing the inclusion of police agency fixed effects to account for level differences across cities. I control independently for the effect of the application score on the time path of police and crimes by including interactions between the score and a set of event time (i.e. years since application) indicators in the regressions. Further, I construct year fixed effects that vary by city size and pre-program crime trends, so that estimates are identified by comparing cities above the threshold with cities below that are of similar size and followed similar trends in the years prior to grant application.

I show that high and low scoring cities follow similar trends in police and crime prior to the application year. Compared with cities just below, however, police rates increase by about 3.6% in the years following the program application for cities above the threshold, while violent crimes and property crimes fall by about 4.8% and 3% respectively. To estimate a single police-crime elasticity for each crime category, I instrument the police rate with an interaction between an indicator for the post-application period and an indicator for whether the score exceeded the threshold in 2SLS regressions where the crime rate is the dependent variable. The estimates imply crime-police elasticities of -1.36 for violent crime and -0.84 for property crime. Estimates obtained with a conventional Regression Discontinuity Design approach are almost identical. An analysis of individual crime types reveals that police reduce murders, robberies, larcenies, and auto thefts. The estimated elasticities for murder and robbery are particularly large relative to existing studies. The point estimate in my primary murder specification implies that one life can be saved by hiring eleven police officers. I find little evidence, however, that arrests increased following the program-induced police force expansions, which suggests a deterrence mechanism underlying the estimated crime effects.

While mine is not the first paper to study the effect of police on crime, this study contributes to the existing literature in several important ways. The discontinuous allocation rule used in distributing grant funding allows for a cleaner identification strategy than has been used in past studies. While the majority of existing literature has focused on large cities or state-level data because of data quality concerns, I rigorously clean the FBI crime data and examine all cities with populations above 1,000 that applied for COPS funding between 2009–2013. My results, therefore, may be relevant to a larger share of local governments than prior estimates. Finally, several of the most-cited papers on the topic have studied the high crime periods of the 1980's and 1990's. I study a period with low and falling crime rates and show that additional police still have a meaningful impact in this very different environment.

The rest of the paper is organized as follows. Section 2 provides background on the COPS hiring program. I describe the data in Section 3 and my empirical strategy in Section 4. Section 5 presents the results. I conduct a cost-benefit analysis in Section 6 and conclude in Section 7.

2 Background on the COPS Hiring Program

2.1 Program History

In September 1994, President Bill Clinton signed into law the Violent Crime Control and Law Enforcement Act, the largest federal crime bill to date. The bill authorized $8.8B in spending on grants for state and local law enforcement agencies between 1994 and 2000 and established the office of Community Oriented Policing Services (COPS) to administer the new grant programs. A key tenet of the crime bill was the creation of the COPS Universal Hiring Program (CHP), which covered 75% of the cost of new police hires for grant recipients. The stated goal of the hiring grant program was to put 100,000 new police officers on the street.[4]

CHP funding exceeded $1B in fiscal years 1995–1999, but appropriations fell considerably in the early 2000's. Less than $200M was allocated for the hiring program in 2003–2004, and less $20M was appropriated in each year 2005–2008 (James 2013). The program was defunded due both to the retreat of crime as a central policy issue and to questions over the program's effectiveness (Evans and Owens 2007). Reports

[4]See http://www.justice.gov/archive/opa/pr/Pre_96/October94/590.txt.html.

produced by the Heritage Foundation in 2001 and 2006, for example, argued that hiring grants did not reduce crime because grants were used to supplant other expenditures rather than to expand police forces.[5]

Funding for the hiring program saw a dramatic resurgence in 2009 with President Obama's signing of the American Recovery and Reinvestment Act. The Recovery Act provided $2B in new funds to the Department of Justice, with $1B earmarked specifically for the COPS hiring program. The funding was seen both as a precautionary measure for keeping crime rates low in the face of a worsening economy and as a means to create or preserve as many as 5,000 police officer jobs across the country. Following the injection of ARRA funds in FY2009, congressional appropriations exceeded $140M annually between 2010 and 2013, a large increase from the 2004–2008 funding levels.[6] Hiring grants awarded in FY's 2009–2011 were also more generous than in previous years, covering 100%, rather than 75%, of entry-level salary and fringe benefits for hires or rehires for three years.[7]

2.2 Details of COPS 2.0

Hiring grants were distributed based on an open solicitation application process – any state, local, or tribal agency with primary law enforcement responsibility was eligible to apply for funding. As part of the application, agencies were required to submit an array of statistical information, including indicators of fiscal health, local unemployment rates, local poverty rates, and local crime rates. Indicators of municipality fiscal health included police agency operating budget, local government operating budget, and locally generated revenue for the current and prior two fiscal years, as well as details about local government employee layoffs. Applicants were also required to submit an essay on their community policing strategy and request a specific number of police officers for which they required funding.[8]

Using the submitted information, the COPS office assigned each applicant a *fiscal need* score and a *crime* score. COPS office documentation indicates these scores were generated by ranking applicants in the same state and population group (smaller or larger than 150,000) against each other on each application question, then weighting each question to obtain an aggregate ranking for each agency. I was unable to replicate the reported application scores by following this approach, most likely due to my inability to observe a large share of the application material. Municipal level employment and financial data, for example, are publicly available on an annual basis for only a small fraction of cities. As such, I proceed treating the function

[5] See, e.g., http://www.heritage.org/research/reports/2008/04/why-would-cops-20-succeed-when-cops-10-failed.
[6] See (James 2013) for a detailed history of COPS funding.
[7] The program reverted to covering 75% of salary and benefits beginning in 2012.
[8] See http://www.cops.usdoj.gov/pdf/CHP/e05105273-CHP.pdf.

mapping application data to scores as unknown. The two component scores were added together to create an aggregate application score, and funding was allocated according to the within-state score ranking. Data on hiring grant awards strongly suggest that a de facto cutoff strategy was used, which can be seen in Figure 1.

Applicants were eligible to receive a grant of up to 5% of current force size, capped at a maximum. For example, a department employing 100 sworn officers at the time of application was eligible to receive a grant of 5 officers from COPS. The maximum grant size was 50 in 2009, 25 in 2010–2012, and 15 in 2013. These maximums were binding for only a small share of applicants. In my sample of 4,374 cities, only 9 departments employed over 1,000 officers and only 55 employed over 300. Two final rules governed grant allocations. First, the COPS office was required to distribute at least 1.5% of total hiring program funding to each state. Second, they were required to distribute at least 50% of all funding to jurisdictions with populations exceeding 150,000.

2.3 Research on the COPS Program

Although this paper is, to my knowledge, the first to examine COPS 2.0, several papers have studied the first iteration of the COPS hiring program. The most noteworthy paper on the topic is the careful and well-regarded study by Evans and Owens (2007). Papers by the GAO (2005) and Worrall and Kovandzic (2010) also study the original COPS program and employ similar research designs.

In the first part of the paper, Evans and Owens (2007) examine whether COPS grants increased police forces. Using a twelve-year (1990-2001) panel of 2074 cities, they regress sworn officers per 10,000 residents on the lagged number of officers granted by the COPS office per 10,000 residents in panel data models, finding that local police forces increased by 0.7 sworn officers for each granted officer. In the second part of the paper, the authors instrument the police rate with the lagged grant rate in 2SLS regressions where the crime rate is the outcome of interest, finding that increases in police are associated with statistically significant declines in robberies, assaults, burglaries, and auto thefts.

Relative to Evans and Owens (2007), my contribution is twofold. First, I improve on their identification strategy. I observe data on grant applications, which they do not, and am able to infer a discontinuity-based allocation rule from these data. This allows the use of cities who applied for but were not offered hiring grants as a control group for grant winners. I argue that the set of applicants denied funding is a superior control group to the broader set of cities who report crimes to the FBI. Applicants may differ from non-applicants in their beliefs about future crime, for example. Further, the use of the discontinuous allocation rule helps circumvent the possible endogeneity of grant take-up. In my setting, some cities with scores above the threshold are not observed as winning grants, which may reflect cities denying grant offers or the COPS office specifically rejecting applicants on the basis of private information. My identification strategy considers these cities as treated, however, alleviating concerns about unobservable differences across cities that do and do not receive hiring grants.

Second, I study a different era of the program. Evans and Owens (2007) examine the introduction of the COPS program in the mid 1990's, when crime rates were high and crime in general was a central policy issue. The stated goal of the program was to induce large increases in police forces across the country. My focus is the reinvigoration of the program following the injection of ARRA funding. The goal of COPS 2.0 was to preserve law enforcement jobs and prevent a rise in crime due to worsening economic conditions. The poor fiscal health of many cities during this period, combined with a lower program budget than during the original COPS period, generated a highly competitive application process. The different context, various program changes, and the availability of a cleaner identification strategy warrant a new evaluation. Further, this paper contributes to a broader literature on the effectiveness of the Recovery Act and offers insights on the relative benefits of including law enforcement funding in stimulus packages.

3 Data

3.1 COPS Program Data

I obtained data on applications for COPS hiring grants for the years 2009–2013 from the COPS office website.[9] These data provide FBI ORI codes and application scores for the universe of applicants in each year.[10] The application score scale varies across years, and I standardize the scores to have mean zero and variance one for each year individually so that score "distance" has the same interpretation across program years. The COPS office also provides information on grant recipients for each year beginning in 2008. A small number of hiring grants were awarded in 2008, and since I do not have application information for 2008, I discard these data. The grants data include name of agency, number of officers granted, and dollar value associated with the grant, for all CHP awards in each year. I collected data on award winners from 2009–2013 and merged them with the application information using a name-matching algorithm, with a match rate of 97.7%.

The application score cutoffs were computed as follows. The COPS office documentation indicates that

[9] For example, 2009 application scores are at http://www.cops.usdoj.gov/pdf/Applicant_Rankings2.pdf.

[10] An ORI code is the unique identifier given to each agency that reports crimes to the FBI through the Uniform Crime Reporting Data System.

applicants from the same state and population group (greater or smaller than 150,000) were ranked against each other. I divided the applicants into state $\times$ size $\times$ program year groups g accordingly. Note that while I use only municipal agencies with populations above 1,000 in the regression analysis, I use all the applicants to compute the score thresholds. Within each group, I follow the strategy in Hoekstra (2009) to infer the cutoffs. That is, at each score s in group g, I estimate the regression

$$\mathbf{1}[\text{Win Grant}_i] = \alpha + \beta \mathbf{1}[s_i \geq s] + \epsilon_i$$

using only group g observations and identify the value of s_g that maximizes the R-squared of this regression. In several cases, the selected s_g value is a non-negligible margin above the next highest score. In these cases, setting the cutoff at s_g may overstate the true "closeness" to the cutoff of the city at s_g. For this reason, I compute the threshold in group g as the average of the regression-selected s_g and the next highest score.

In the analysis, I discard applications from groups without a competitive application process. Group g is deemed noncompetitive if it meets any of three conditions: (1) group g has no winners; (2) group g has no losers; (3) the coefficient from a regression of a *grant win* dummy on the application score, using only observations in group g, is negative.

The results of the cutoff computations can be seen in Figure 1. The figure pools all program years together and plots the probability of winning a hiring grant as a function of the application relative to the cutoff. There is a clear discontinuity in the probability of grant receipt at the threshold. In a regression of a grant win indicator on the relative score and an indicator for whether the score exceeds the threshold, the coefficient (standard error) on the high score indicator is 0.89 (.0088). A more formal RD estimate (see Table 3) yields a coefficient of 0.69 (0.03).

3.2 FBI Data

Data on police employees and crimes reported at the agency level are from the FBI's Uniform Crime Reporting Data System (UCR), which are compiled by and available for download from the NACJD. The UCR provides monthly counts of index I crimes for all reporting agencies in the *Offenses Known* file. Index I crimes include the core violent (murder, rape, robbery, aggravated assault) and property (burglary, larceny, motor vehicle theft) crimes. The number of sworn officers employed by each agency in each year is reported in the UCR *Law Enforcement Officers Killed in Action* (*LEOKA*) file. Because police officer counts are reported only once yearly, and many agencies report their full-year crime counts once rather than report monthly, I aggregate these counts to the agency-year level. I collected UCR data for 2005–2014, the most recent year currently available. For city population, I use a smoothed version of the measure reported in the UCR files.[11]

The UCR data requires thorough cleaning before use. I implement a regression-based approach similar to that of Evans and Owens (2007) to identify record errors and extreme outliers. Using data from 1990–2014, I fit the time series of police and crime rates to a quartic time trend for each agency individually. Observations were then identified as errors if the percent difference between the observed and predicted values differed by more than a pre-determined city size group × crime type threshold. Observations flagged as errors are recoded as missing in the dataset. The data-cleaning procedure is described in more detail in the Data Appendix.

3.3 Other Data Sources

The COPS and UCR data are supplemented with basic demographic variables measured at the county level. I use county-level data because most demographic measures are not available at the city level on an annual basis. I computed percent black, percent hispanic, and percent aged 15-24 from county population estimates obtained from the SEER program at the NIH for the period 2005–2014. County-level per-capita income was obtained from the Bureau of Economic Analysis (BEA), and county-level unemployment rates were obtained from the BLS Local Area Unemployment Statistics data files. I use percent black, percent Hispanic, percent aged 15-24, log per capita income, and unemployment rate as controls in the crime regressions.

3.4 Sample Construction

Using the 2000, 2005, and 2012 *Law Enforcement Agency Identifiers Crosswalks*, which are compiled by the NACJD, I identified the universe of municipal police agencies that report crimes to the FBI. 9,949 police agencies meet this requirement. I drop cities with population below 1,000, which reduces the number of agencies to 9,162, because per-capita measures are much noisier, and often orders of magnitude higher, below this threshold.[12] I also require cities to have five (out of ten possible) years of valid population, police, violent crime, and property crime observations. 8,917 cities survive this restriction. From this sample, I drop cities with any observations in the top and bottom 1% of the distribution of sworn officers per capita, which leaves 8,034 police agencies. Of these, 4,374 applied for COPS program funding between 2009 and 2013 and were in a competitive application group. This list of 4,374 cities comprises the main sample.

[11]Chalfin and McCrary (2016a) note that the UCR population measure tends to jump discontinuously around census years. For this reason, I follow their procedure and smooth the population measure using local linear regression. For more detail, see the Online Data Appendix.

[12]See Figure 3 in the Online Data Appendix.

3.5 Sample Characteristics

As described in more detail below, the unit of analysis in this study is an application, indexed by a city and program year. Table 1 illustrates the distribution of applications across city size groups and program years. The sample is heavily weighted towards smaller cities, with 49% of cities having populations between 5,000 and 25,000. Only 10% of cities are above the 50,000 resident threshold. The average city applies twice. About 34% apply only once, while about 56% apply either twice or three times. The rest apply four or more times. The applications under study are distributed across program years as follows: 43% in 2009, 25% in 2010, 13% in 2011, 7.5% in 2012, and 11% in 2013.

Table 2 presents summary statistics for each of the 8,804 city × applications in the year prior to application. The average city has about 25,000 residents, 21 sworn officers per 10,000, 35 violent crimes per 10,000, and 314 property crimes per 10,000. Breaking down the applications by high (above the threshold) and low scores reveals some disparities in the types of cities. High-scoring cities are about twice as large on average and have about 2.5 more police per 10,000 residents. Violent (property) crime rates are about 80% (50%) higher in cities with applications above the threshold.

Figure 2 illustrates the relationship between the application score and characteristics of the city × application observations. The first panel plots the frequency of the relative application score. The lack of excess mass just above the cutoff suggests no systematic manipulation by cities into eligibility. The estimated discontinuity (standard error) from the McCrary (2008) test is -.072 (.076), implying that smoothness in the density function cannot be rejected. The second panel plots the average covariate index by score bins of width 0.25 in the year prior to application. The covariate index is the predicted crime rate from a regression of total crimes per 10,000 residents on controls (see above) and year fixed effects (to account for the fact that applications were submitted in different calendar years). Cities above the cutoff have higher covariate indices on average, but the difference does not appear to be discontinuous. Similarly, it appears that in the year prior to application, high scoring cities have higher police and crime rates, but cities exceeding the threshold do not differ discontinuously on these measures.[13]

[13] This is confirmed statistically in Table A-1, which presents RD estimates using local linear regression and Imbens and Kalyanaraman (2012) bandwidths where the pre-application covariate index, police rate, and crime rate are the outcomes of interest. None of the estimated discontinuities is statistically significant.

4 Empirical Strategy

The goal of the empirical analysis is to leverage quasi-random variation in police rates induced by COPS program allocation rules to estimate the causal effect of police on crime. As discussed above and pictured in Figure 1, a fuzzy cutoff rule in the application score was used in allocating grants among applicants. Less than 10% of applicants with scores just below the threshold received grants, while over 80% of applicants just above were funded. Such an allocation rule lends itself naturally to a regression discontinuity approach, which would compare cities just above the threshold with cities just below with the underlying assumption that exceeding the cutoff is random.

To begin, I consider a standard regression discontinuity design. Let i index cities and denote s_i for city i's application score. City i faces a cutoff score s_i^* and I let

$$h_i = \mathbf{1}[s_i \geq s_i^*]$$

be a high score indicator. A standard RD specification in this context is

$$\Delta y_i = \theta h_i + f(\tilde{s}_i) + \epsilon_i$$

where $\tilde{s} = s - s^*$, $f(\cdot)$, and Δy_i is the change in some outcome y. In the main specification, I consider the change between one year prior to one year after the program application as the change of interest. A complication is the fact that cities need not apply only once. In the RD setup, I treat each city $\times$ application as its own observation – a single city appears in the dataset once for each submitted application. In the estimation, I approximate f with local linear regression and select the application score bandwidth as the optimal bandwidth from Imbens and Kalyanaraman (2012). The coefficient of interest, θ, measures the extent to which the change in y differed for cities above and below the threshold.

In practice, the standard RD estimates are imprecisely estimated. To address this, I instead implement a dynamic, difference-in-differences version of the regression discontinuity design. My strategy is based closely on that used by Cellini et al. (2010), who study the effect of passing bonds for school facility investments on housing prices. The setup is as follows. Suppose that city i applies for funding in year $\tilde{t}$ and receives the application score s_i. The cutoff score faced by city i in year $\tilde{t}$ is s_i^* and let h_t be an indicator for exceeding the threshold. Let $\tau = t - \tilde{t}$ be the year relative to the application, or the *event time* year. Assuming cities vary in their application years, one could examine the effect over time of crossing the threshold at $\tau = 0$ by estimating

$$y_{it\tau} = \alpha_\tau + \theta_\tau h_i \times \alpha_\tau + \phi_i + \kappa_t + \epsilon_t \tag{1}$$

where α_τ is an event time fixed effect, ϕ is a city fixed effect, and κ is a year fixed effect. Equation (1) is an event study regression where the event is a grant application and the event time coefficients are allowed to vary by whether the application score is above or below the threshold.

Again, an important complication is that cities need not apply only once. Indeed, in my sample of 4,374 applicant cities, 2,867 apply multiple times. Considering a single application for each city would require taking a stand on which is the focal application. Further, it would reduce the number of events to be analyzed from 8,804 to 4,374. Instead, I follow the *stacking* approach for dealing with multiple events detailed in Lafortune, Rothstein and Schanzenbach (2016) and Cellini et al. (2010). The time series for each city is copied once for each application that city submits. That is, one observation in the dataset is a city × application year × calendar year.

The analogue to (1) with stacked data is

$$y_{iat\tau} = \alpha_\tau + \theta_\tau h_{ia} \times \alpha_\tau + \phi_{ia} + \kappa_t + \epsilon_t \tag{2}$$

where h_{ia} is an indicator for whether the application score in year a is above the cutoff and ϕ_{ia} is city × application fixed effect. While the city fixed effects are replaced with city × application fixed effects, I cluster the standard errors at the city level. Because of the stacked data, cities with more applications receive more weight in estimates of (2). To account for the relative overrepresentation of cities that apply more often, I weight cities by one over the number of applications when estimating regressions of this form.

Underlying equation (2) is an assumption that whether s_{ia} exceeds the threshold is random. This assumption is plausible for application scores very close to the cutoff but less believable in cases where the score is far from the threshold. Rather than discard city × applications outside an arbitrary bandwidth of the threshold, I make two adjustments to the estimating equation to improve the plausibility of this assumption. First, I add interactions between the score in year a and the event time dummies. These interactions capture the effect of the score on the time path of y independent of whether the score exceeds the threshold. Second, instead of controlling only for year fixed effects, I group cities according to their size and pre-program trends and construct year effects that vary by these groups. Using this technique, cities with scores above the threshold are not compared with all cities below, but instead with cities below the threshold that are of similar size and followed similar pre-program trends.

The pre-program cell by year effects, also used in Evans and Owens (2007), are constructed as follows. First, cities are grouped into six size categories according to their population: 1,000-2,500; 2,500-5,000; 5,000-10,000; 10,000-25,000; 25,000-50,000; 50,000+.[14] Cities that fall in multiple categories during the sample period are placed in the group they appear most often. Then, for each city, I estimate a linear time trend in the police rate and crime rate for the period 2005-2008. Cities are then grouped according to the size category by (within size category) quartile of police pre-trend by (within size category) quartile of crime pre-trend. That is, a city falls into one of $6 \times 4 \times 4 = 96$ groups, and regressions include year effects that vary by these groups. Note that this approach is done separately for each crime type, so that in regressions where, for example, robbery is the dependent variable, cities are grouped according to their pre-program robbery trends.[15]

The main estimating equation is then

$$y_{iat\tau} = \alpha_\tau + \theta_\tau h_{ia} \times \alpha_\tau + \lambda_\tau s_{ia} \times \alpha_\tau + \gamma X_{it} + \phi_{ia} + \kappa_t + \epsilon_{iat\tau} \tag{3}$$

where λ_τ is the coefficient on an interaction between the event time fixed effect and the application score associated with applicant ia and κ_t is a pre-program cell by year fixed effect. The coefficients of interest are the θ_τ's, which trace out the effect of crossing the threshold in event time. The pre-period ($\tau < 0$) θ_τ's provide a specification check for whether the pre-program trends differ among high and low-scoring cities. The post-period θ_τ's are intent-to-treat estimates of the effect of hiring grants on y.

Estimates of (3) where y is police officers per 10,000 residents and crimes per 10,000 residents can be thought of as the first stage and reduced form components of an instrumental variables estimate of the causal effect of police on crime, generating one estimate for each event time year. To estimate a single average effect, I estimate the equation

$$Crime_{iat\tau} = \theta Police_{iat\tau} + \beta Post_{iat\tau} + \alpha_\tau + \lambda_\tau s_{ia} \times \alpha_\tau + \gamma X_{it} + \phi_{ia} + \kappa_t + \epsilon_{iat\tau} \tag{4}$$

[14] 50,000 residents is a lower than desirable cutoff for the largest city size group. However, there are only 123 cities with populations $> 100{,}000$. This results in a very small number of cities in each pre-program cell bin. Hence, I pool the 50,000-100,000 and 100,000+ groups together when creating the groups.

[15] I present a test of whether grouping cities accordingly, as well as controlling for event time $\times$ application score interactions, improves the plausibility of the randomness assumption in Table A-4. Using data from the year prior to application, I regress an indicator for whether the score exceeded the threshold on a set of city characteristics. When the full sample is used and cities are ungrouped (Column 1), the F-statistic from a joint significance test of the observable characteristics is 129.25. Adding pre-program cell fixed effects reduces the F-stat to 94.25 (column 2), and controlling linearly for the score further reduces it to 22.4 (Column 3). It is important to note that level differences in observables across cities above and below the threshold are, in practice, irrelevant because such differences are absorbed by the city $\times$ application fixed effects in (2). This exercise simply demonstrates that above-threshold cities are compared with more similar below-threshold ones after making these adjustments.

where $Post = \mathbf{1}[\tau \geq 0]$ and $Post_{iat\tau} \times h_{ia}$ instruments for police. This specification is a standard difference-in-differences IV estimating equation except that the event time indicators and event time $\times$ application score interactions are included.

5 Results

5.1 Regression Discontinuity Evidence

Figure 3 plots binned averages of the two-year change in police rates by the application score. The unit of observation is a city $\times$ application and I focus on city $\times$ applications within one point of the cutoff. Bin widths are selected to have equal mass. The squares plot this change as of one year prior to application (i.e. the change from three years prior to one year prior), which is a placebo test. The circles plot this change as of one year after the application (i.e. the change from one year prior to one year after), which captures the short term effects of the program. Although the binned averages are noisy, the figure suggests that changes prior to the program are relatively continuous through the score distribution, while changes after the program increase discontinuously at the score threshold. Police rates increase slightly for cities above the threshold, who become eligible for funding, while police rates fall by between 0.5 and 1 sworn officer per 10,000 for cities below the cutoff. The differences at the threshold suggest that COPS grants were associated with relative increases in police levels.

Figure 4 presents identical plots where the two-year change in crimes per 10,000 residents is the outcome of interest. Again, the average changes are relatively noisy, but the plots suggest that cities above and below the threshold experienced similar changes in violent and property crimes prior to program application. Cities above the cutoff, however, appear to experience larger declines in crime than cities below following program application. The disparity in crime changes at the threshold suggests that (relative) police increases induced by the program caused a (relative) drop in crime.

Table 3 presents the RD estimates. All regressions use the optimal IK bandwidth and control for the score with local linear regression. Robust standard errors from Calonico, Cattaneo and Titiunik (2014) in parentheses. Columns 1-2 show that crossing the threshold is associated with an increase in the probability of grant receipt of 0.69 percentage points and an increase in active grant funding of about \$61,000.[16] Column 3 indicates that police forces increase by 0.63 sworn officers at the threshold. Columns 4 and 5 suggest that crossing that threshold was associated with relative declines of 1.68 violent crimes and 10.46 property

[16] *Active Funding* is my estimate of grant dollars spent on a given city in a given year. This is computed by summing total funding over the prior three years (because grants cover three years of salary) then dividing by three (to annualize the amount).

crimes. Evaluating at pre-period means, these results imply crime-police elasticities of -1.24 for violent crime and -0.86 for property crime. The reduced form crime estimates are not statistically signifiant, however.[17]

5.2 Event Time Evidence

Figure 5 plots the θ_τ's from estimates of (3) where sworn officers per 10,000 residents is the dependent variable. One year prior to the application is the excluded year, so the differences between high and low scoring cities are normalized to that year. Cities with scores above and below the threshold follow similar trends in the years prior to application. Beginning with the application year, however, the cities diverge, with police increasing among cities above the threshold relative to those below. Note that given the timing of program events and police measurement, it is reasonable to expect hiring grants to have an impact in the application year. Grant funding is distributed in the summer, while police levels are measured in October. Further, anecdotal evidence suggests that many COPS 2.0 awards were used to avoid scheduled layoffs, and such an effect would be seen more quickly than actually hiring and training new officers.

The corresponding estimates for violent crimes are presented in Figure 6. As with police, violent crime rates follow similar trends in high and low scoring cities during the pre-period, but diverge beginning with the application year. The coefficients in years 0 through 2 are all statistically distinguishable from zero. The high score coefficients are noisy for murder, but do turn more negative in the post period, with the coefficient at $\tau = 1$ statistically significant. For robbery, above cutoff cities appear to follow a slight upward trend relative to those below in the pre-period. Compared with low-scorers, however, high-scoring cities experience statistically significant declines following the application year. The assault coefficients follow a pattern similar to those for aggregate violent crime, turning from zero in the pre-period to negative in the post-period before returning to zero. None of the post-period coefficients are significant, however.

An analogous pattern emerges for property crimes, as depicted in Figure 7. For aggregate property crime, larceny, and auto theft, exceeding the threshold at time zero is unrelated with changes in the crime rate in years prior, but is associated with statistically signifiant declines in the post period. Coefficients also turn negative following the application year for burglary, albeit more slowly.

The point estimates underlying the event study figures for police and aggregate violent and property crimes are presented in Table 4. The estimates indicate that on a base of 22.54, exceeding the threshold is associated with about a 3.6% increase in the police rate. Similarly, cities with high scores experience a relative decline

[17] Regression discontinuity plots and estimates for individual crime types are presented in Figure A-2 and Table A-2.

in violent crime of about 4.8%, which suggests an elasticity of violent crime with respect to police of about -1.33. Property crimes decline by about 3% following the application year for grant eligible cities, suggesting an elasticity of about -0.8. The analogous estimates for the individual crime types are shown in Table A-3.

5.3 IV Results

The differences-in-differences IV strategy outlined in Section 4 suggests a straightforward way to estimate the effect of grants on police. Column 1 of Table 5 presents coefficients on *Post* and $Post \times High$ from a regression where active grant funding per 10,000 residents is the dependent variable.[18] The coefficient on $Post \times High$ suggests that relative to cities below the threshold, grant funding increases by about \$51,000 dollars per 10,000 residents following the application. Column 2 shows that police forces increases by about 0.75. A simple Wald estimate of the effect of grant funding on police is then 0.75/\$51,000, which implies that one officer-year is added for every \$68,000 in grant funding awarded.

Columns 3 and 4 present IV estimates of the effect of police on violent and property crime based on equation 4, that is using $Post \times High$ as an instrument for the police rate. The F-statistic corresponding to the instrument is in the range of 30, confirming that this instrument satisfies the relevance condition. The estimates suggest that an additional officer is associated with 2.9 fewer violent crimes and 16.23 fewer property crimes. To convert the coefficients to elasticities, I multiply the point estimates by the ratio of the mean police and crime rate.[19] The results imply crime-police elasticities of -1.36 for violent crime and -0.84 for property crime. Worth noting is the fact that these elasticities are nearly identical to those obtained via the standard RD specification, which lends credence to the dynamic RD identification strategy.

Table 6 presents the corresponding estimates for individual crime types. The results suggest that police force increases generate statistically significant declines in murders, robberies, larcenies, and auto thefts. Estimated coefficients imply that an additional officer leads to 0.09 fewer murders, 1.4 fewer robberies, 8.6 fewer larcenies, and 3.5 fewer auto thefts. The implied elasticities are -4.8, -2.9, -0.62, and -2.18 respectively. Worth noting is the fact the event time estimates also suggested an effect on assault and that the p-values for the assault and burglary coefficients in Table 6 are 0.275 and 0.156, respectively.

The magnitude of the estimated murder effect warrants further discussion. The point estimate of -0.0896

[18] Active funding is computed as the sum of grants received over the previous four years divided by four. Grants cover three years of salary but the event study estimates suggest that year zero and year four are partially treated.

[19] The means are for 2005 and for cities within 1 point of the score threshold. These means were selected because the instrumental variables estimates are local to cities near the cutoff.

implies that one life can be saved by hiring eleven additional police officers. At a value of a statistical life (VSL) of $7M and an annual cost per officer of $130,000 (Chalfin and McCrary 2016a), the benefits of this proposition outweigh the costs by a wide margin of $5.6M. The lower bound of the 90% confidence interval around the point estimate (-0.019) implies one life saved for every 53 officers hired, which still passes a cost-benefit test by over $100,000.

5.4 Comparison with Existing Studies

A common finding in the existing literature is that crime-police elasticities are larger for violent than property crime. My results confirm this finding, as I estimate a violent crime elasticity of -1.36 and a property crime elasticity of -0.84. Evans and Owens (2007) find a similar violent crime elasticity of -1.34. Their estimated property crime elasticity of -0.26, however, is about one third the size of my estimate. Lin (2009), who studies state-level data, also finds a similar violent crime elasticity of -1.13, but a considerably larger property crime elasticity of -2.18. Chalfin and McCrary (2016a) find smaller elasticities of -0.34 and -0.17, but these disparities could be due to differing samples – their study examines only large cities.

Estimates for individual crime types vary widely across existing papers. Murder-police elasticities range from -0.24 (Marvell and Moody 1996) to -3.03 (Levitt 1997). My point estimate implies a quite large elasticity of -4.8, but the 95% confidence interval includes elasticities as small as -0.35. Consistent with existing work, I find little statistical evidence that police reduce rapes but strong evidence of an effect on robbery. My estimate of the robbery-police elasticity of -2.93, however, is larger than most existing estimates. For example, Lin (2009) and Evans and Owens (2007), whose estimates are at the larger end of the current literature, find statistically significant elasticities of -1.86 and -1.34.

Past studies have diverged on whether police reduce aggravated assaults. Evans and Owens (2007) find a large and statistically significant elasticity of -0.96, while the corresponding figure in Chalfin and McCrary (2016a) is a small and insignificant -0.1. The implied elasticity in Table 6, -0.77, is on the larger end of existing estimates. Although the coefficient is not statistically significant, the event time coefficients in Figure 6 appear to suggest an effect. Several studies have found that police reduce burglaries, with most estimated elasticities between -0.3 (Klick and Tabarrok 2005) and -0.59 (Evans and Owens 2007). My estimate, -0.64, is similar in magnitude but not significant.

Existing estimates for larceny also vary widely. Many studies have found essentially no effect, while my results suggest a statistically significant larceny-police elasticity of about -0.62. On the other hand, prior studies have commonly found large and robust effects of police on auto thefts. Estimates range from -0.44 (Worrall and Kovandzic 2010) to -0.85 (Klick and Tabarrok 2005) to as large as -4.14 (Lin 2009). My estimate of -2.18 is on the larger end of estimates in the literature.

5.5 Robustness

In this section, I probe the robustness of the crime-specific estimates discussed above. Figure 8 illustrates the first stage and reduced form estimates underlying the crime regressions when varying bandwidths are used – that is, when only city × applications within a certain distance of the application score threshold are used in the estimation. A bandwidth of 4 includes 99% of the data and 4.5 includes all city × applications used in the primary regressions.

The police panel indicates that the first stage effects are stable regardless of the bandwidth used. The reduced form murder effect is similar in magnitude for bandwidths larger than 0.1 and statistically significant for bandwidths larger than 1.5. The point estimate shrinks slightly, however, when only applications within 0.5 points are used. The rape estimates hover around zero and are never distinguishable from zero. For robbery, the reduced form effects are statistically significant at the 95% level regardless of the bandwidth, although the point estimate is slightly smaller when using only closer applications. Coefficients for assault and burglary are never significant and reveal no clear pattern in terms of the relationship between the bandwidth and effect size. The magnitude of the estimated larceny effect is consistent across bandwidth sizes and most estimates are significant at the 90% level, as was the case in Table 6. The results for auto theft reveal that although estimated effect size for bandwidths greater than one are similar and always significant, the effect is smaller and significant only at the 90% level when only the closest applications are used in the analysis.

I present additional robustness checks in Table 7, which repeats the estimates in Table 6 using alternative specifications. In the second row, I repeat the estimation using only cities with ten years of valid data for that crime type. The effects are similar, and if anything, larger than in the main specification. In the third row, I conduct the estimation using "balanced panels" – that is, for each city × application, only years between four years prior and two years after the application year are used. I also drop 2013 applications, so that all cities used in the estimation are observed over the four years prior to two years after range. These regressions test whether the estimated effects are driven by changes in years well after treatment. The murder and robbery estimates shrink, although the robbery estimate is still highly significant. The estimated assault effect is notably larger and marginally significant, implying an elasticity of about -1.15. The larceny and auto theft estimates are similar in magnitude and statistical significance to the main specification.

In the fourth row, I restrict the analysis to cities with population greater than 10,000.[20] This sample criterion is the same as that used in Evans and Owens (2007) and is meant to ensure that the estimated effects are not merely a result of noisiness in the crime data for very small cities. Relative to the main specification, the effects for robbery, assault, larceny, and auto theft grow, with the robbery and auto theft estimates remaining significant. The estimated murder and assault remain large but not statistically distinguishable from zero.

In the fifth row, I augment the main specification by replacing the event time $\times$ application score interactions with event time $\times$ relative score (score minus the threshold) interactions that vary by whether the score exceeds the threshold, similar to a standard regression discontinuity design specification that controls linearly for the running variable and allows the slope of the line to vary on either side of the threshold. Relative to the first row, this approach strengthens the murder effect but weakens the robbery, larceny, and auto theft effects. In the sixth row, I repeat the specification from the fifth row using only city $\times$ applications within one point of the cutoff. This specification, which is the most demanding in Table 7, yields results very similar in magnitude and significance to those in the main specification.

As an additional robustness check, I explore the sensitivity of the results to varying the definition of the treatment. COPS hiring grants cover a police officer salary for three years (or 75% of that salary for grants in 2012 and 2013). In the main specification, however, assumes that cities above the cutoff at year t are treated through the end of the sample. Cities above the cutoff in 2009 are coded as treated 4-5 years after the application, for example. To test the extent to which this coding scheme is relevant for the estimated effects, I repeat the first stage and reduced form evidence under varying treatment length specifications. That is, I create $Post^k = \mathbf{1}[0 <= \tau <= k]$ and $Post^k \times High$ and estimate

$$Crime_{iat\tau} = \beta Post^k_{iat\tau} + \theta Post^k_{iat\tau} \times High_{ia} + \alpha_\tau + \lambda_\tau s_{ia} \times \alpha_\tau + \gamma X_{it} + \phi_{ia} + \kappa_t + \epsilon_{iat\tau}$$

and examine how θ varies across values of k.

The results of this exercise are shown in Figure 9. Note that $k=5$ includes all post-application years in the data and therefore corresponds to the main specification. Setting $k=2$ or $k=3$ would capture the

[20]To be precise, cities whose population exceeds 10,000 in at least five years.

program "letter of the law" and revert cities to "untreated" at the conclusion of the grant period. An immediate takeaway is that the first stage is increasing in the specification of the treatment length. The coefficient is roughly twice as large when all post period years are indicated as treated as when only the application year and the year after are included. The estimates for robbery, burglary, larceny, and auto theft follow an inverse pattern, with effects smallest when the shortest period is used and largest when using the longest period. The inverted relationship is comforting in the sense that implied Wald estimates (the reduced form divided by the first stage) appear to be relatively stable across the treatment length specifications. For murder, the reduced form is quite small and insignificant when $k < 4$, but grows when k is 4-5 and is only significant when $k = 5$, which may shed doubt on the robustness of the estimated murder effect. This result seems to be at odds with the event time estimates, however, as we observed in Figure 6 that high scoring cities experienced large drops in murder at $\tau = 1$ relative to those below the threshold. Further, the estimated murder effect survived the majority of specification tests in Table 7.

5.6 Mechanisms

As with other crime control policies, police hiring may reduce crime through two channels – deterrence or incapacitation. Standard economic models of crime, such as Becker (1968), predict that police increases *deter* crime by raising the expected cost associated with criminal behavior. Cost increases elicit a behavioral response, with fewer potential offenders choosing to engage in crime. However, police may also increase the number of individuals detained or incarcerated, which reduces crime by *incapacitating* potential offenders. By which mechanism police reduce crime is of considerable interest because incapacitation is associated with incarceration costs in addition to the police wage bill.

In practice, my estimates almost surely identify a combination of deterrence and incapacitation effects (Chalfin and McCrary 2016b). To get a sense of the relative importance of the two mechanisms, I examine whether COPS-induced police force expansions were associated with increases in arrest rates. As highlighted in Owens (2012), the intuition behind this test is that if police deter crime, increased police presence will reduce crime without necessarily increasing the number of offenders actually apprehended. On the other hand, for police to have an incapacitation effect, hiring police must increase the number of arrested, and therefore incapacitated, potential offenders.

For this exercise, I rely on data from the UCR *Arrests* file, which reports yearly arrest counts by offense category at the agency level. Not all agencies that report crimes to the FBI also submit arrest data. The sample for the arrest rate analysis includes 3,519 cities (out of 4,374) and 7,138 city $\times$ applications out of (8,804). The arrest data were cleaned and processed using the same method as the crimes data which is described in the Data Appendix. Figure 10 plots the event time $\times$ high score interaction coefficients from regressions where arrests per 10,000 residents is the outcome of interest.[21] The figures suggest that exceeding the threshold is associated with a decline in murder and robbery arrests. There is little evidence, however, that arrest rates changed for any other crime types. In Table 8, I repeat the IV estimates from Table 6 for arrest rates. For reference, I show the coefficient from an identical-sample regression of crimes (instead of arrests) on police in the table. The estimates suggest that police increases were associated with a decline in robbery arrests of similar magnitude to the actual decline in robberies (elasticity of -2.6), which suggests that the estimated robbery reductions were accomplished through deterrence. Past work has also shown that robbery may be a particularly deterrable crime type. Abrams (2012), for example, finds that sentence enhancements for gun crimes enacted in the 1970's and 1980's were associated with large and statistically significant declines in gun robberies. The arrest rates coefficients for no other crime type are statistically significant, and the coefficients are negative for all crime types except assault and burglary. In line with the argument in Owens (2012), I take these results as evidence that police deter murders, robberies, larcenies, and auto thefts.

6 Cost-Benefit Analysis

Given that police added by the program reduced crime, a natural question is whether the COPS hiring program passes a cost-benefit test. The first-stage estimates in Table 5 imply that police forces increased by one for each \$68,000 in grant funding. Just over \$900M was allocated to cities in my sample, implying that grants awarded between 2009-2013 added 13,393 officer-years for these cities.[22] After accounting for deadweight loss associated with raising government funds, the federal cost is in the range of \$1.2B. Most estimates in the literature suggest that the annual cost of a fully-equipped police officer is around \$130,000, which implies that local governments spent an additional \$830M on the estimated police increases. Hence, a reasonable estimate of the program's total cost is about \$2B.

I use victimization cost estimates common in the literature to convert crime reduction into social value dollars. The estimates I use are taken from Cohen and Piquero (2009) and Chalfin and McCrary (2016a)

[21] For comparability, the police and crime event study plots specifically for the arrests sample are shown in Figure A-5.
[22] \$910,764,901/68,000 = 13,393.

and are shown in Table 9. The social benefit associated with an additional police officer is the sum of the police coefficients for the individual crime types, with each coefficient weighted by the social cost associated with that particular crime type. Obviously this computation is sensitive to the choice of coefficients. To be conservative, I use the smallest (in absolute value) coefficient for each crime type in Table 6, which yields an estimated social value per officer of \$417,456. Under this assumption, the total benefit generated by the program is 13,393 $\times$ \$417,456 $\approx$ \$5.6B, which suggests the program easily passes a cost-benefit test.

As noted in existing studies, whether additional police are cost-effective depends particularly on whether they reduce murder. The sensitivity of cost-benefit analysis to the estimated murder effect is, of course, due to the fact that murder is the most costly crime type by a factor of 50.[23] Using the same set of coefficients as above but setting the murder effect to zero gives a total benefit estimate of about \$900M, approximately equal to the federal spending on grants but well below the program's total costs. On the other hand, if an additional police officer prevents at least 0.0214 murders annually, the program is cost effective even if police have no effect on all other crime types.[24] A police-murder effect of -0.0214 is within the 90% confidence interval of the estimate in my preferred specification (-0.0896 ± 0.07) and smaller (in absolute value) than the point estimate in all specifications in Table 7.

As a component of the American Recovery and Reinvestment Act, the increase in COPS program funding was intended, at least in part, to create or save police officer jobs. Hence, it is useful to compare the costs and benefits associated with the COPS program to those associated with other stimulus spending under the heading of job creation. The degree to which ARRA spending increased employment has been the subject of much debate. The academic literature has focused on estimating the cost per job created (or saved) by the Recovery Act, relying on cross-state variation in the generosity of transfers received from the federal government. Despite apparently similar methodologies, existing estimates vary widely. Chodorow-Reich, Feiveson, Liscow and Woolston (2012) estimate a cost per job-year of \$26,000, with most job-creation in the private sector. Conley and Dupor (2013) find that most jobs created were in government and estimate a cost per job-year of \$200,000. My analysis implies a cost per police officer job-year of \$68,000, which is squarely in the range of the existing estimates. Depending on the crime coefficients used, my estimate of the social benefit per officer-year is between \$25,000 and \$418,000. Given the relatively modest cost per job-year, and potentially large positive

[23] Although murder is rare, its costliness outweighs its rarity. For example, while robberies are about 275 times more common than murders, one murder is 550 times more costly when using a VSL of \$7M.

[24] 13,393 officer years $\times$ 0.0213 fewer murders per officer-year $\times$ \$7M (social cost per murder) = \$2B (total program cost).

crime reduction externalities, the benefit-cost ratio associated with police hiring grants may compare favorably with other forms of stimulus spending. Such programs may be more politically feasible, as well, since spending under the heading of crime reduction is more likely to gain bipartisan support than many federal programs.[25]

7 Conclusion

In this paper, I exploit quasi-random variation in police levels induced by COPS program allocation rules to circumvent the endogeneity of police hiring and estimate the casual effect of police on crime. My identification strategy relies on the fact that grant funding was awarded as a discontinuous function of cities' application scores. I compare the change over time in police and crime for cities above and below the application score threshold with the underlying premise that cities below are a valid control group for cities above. Studying the dynamics non-parametrically, I show that police and crimes follow similar trends in high and low scoring cities prior to the application year, but the trends diverge as the high scoring cities become eligible for funding. The corresponding instrumental variables estimates demonstrate that an additional officer prevents 2.9 violent crimes and 16.23 property crimes, with implied elasticities of -1.36 and -0.84.

An examination of individual crime types reveals that the results are driven by large effects of police on robbery, larceny and auto theft. My estimates suggests that an additional officer is associated with 1.39 fewer robberies, 9.6 fewer larcenies, and 3.5 fewer auto thefts. I also find evidence of a sizable effect of police on murder – the coefficient in the main specification is statistically significant and implies that one murder per year can be prevented by hiring eleven officers. The magnitudes of the murder, robbery, larceny, and auto theft estimates are generally robust to specification checks and the effects remain when examining only cities close to the application score thresholds. Further, I show that program-induced police increases were not coupled with increases in arrest rates, which suggests that crime reductions were achieved through deterrence rather than incapacitation.

The conclusion of a cost-benefit analysis of the program hinges on whether police reduce murder. Assuming no effect of police on murder, I estimate a total social benefit attributable to the program approximately equal to federal spending on hiring grants but well below my best estimate of the program's total cost. On the other hand, if an additional police officer prevents at least .0214 murders annually, the social value associated with murder reduction alone is enough to justify the program's cost. A police-murder effect of

[25]See, e.g. *Bipartisan House group seeks to bolster nation's police forces with COPS bill*, Mike Lillis for the thehill.com, 5/14/2011.

-.0214 is within the 90% confidence interval of the point estimate in my preferred specification (-.089), and I estimate effects of at least this magnitude in all specifications. Regardless of the cost-benefit conclusion, however, the results highlight that programs to increase police officer employment may offer higher returns than other stimulus spending because of the associated crime reduction externality. I estimate that one officer-year was added for every \$68,000 spent by the federal government, and that the social benefit of the ensuing crime reduction is at the very least \$25,000 but quite possibly as large as \$400,000.

My analysis raises several questions for future work. The fact that police increases were, on average, associated with large crime reductions suggests that police levels were not set optimally ex ante. This conclusion is also reached in a recent paper by Chalfin and McCrary (2016a), who argue that the average U.S. city is underpoliced. Understanding what frictions prevent the efficient allocation of local government resources for crime prevention might prove an interesting avenue for new research. Additionally, the treatment effect of police on crime is almost certainly heterogeneous. A careful examination of this heterogeneity, which is beyond the scope of this paper, may reveal the optimal allocation of police across cities and assist the COPS program in targeting future grants.

References

Abrams, David S, "Estimating the Deterrent Effect of Incarceration Using Sentencing Enhancements," *American Economic Journal: Applied Economics*, October 2012, *4* (4), 32–56.

Angrist, Joshua and Jorn-Steffen Pischke, *Mostly Harmless Econometrics*, Princeton University Press, 2009.

Ater, Itai, Yehonatan Givati, and Oren Rigbi, "Organizational Structure, Police Activity, and Crime," *Journal of Public Economics*, July 2014, *115* (1), 62–71.

Baicker, Katherine and Mireille Jacobson, "Finders Keepers: Forfeiture Laws, Policing Incentives, and Local Budgets," *Journal of Public Economics*, December 2007, *91* (11), 2113–2136.

Becker, Gary S, "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76* (2), 129–217.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs," *Econometrica*, December 2014, *82* (6), 2295–2326.

Cellini, Stephanie, Fernando Ferreira, and Jesse Rothstein, "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *Quarterly Journal of Economics*, February 2010, *125* (1), 215–261.

Chalfin, Aaron, "The Economic Cost of Crime," in Wesley Jennings, ed., *The Encyclopedia of Crime and Punishment*, January 2016, pp. 1–12.

____ **and Justin McCrary**, "Are U.S. Cities Underpoliced?: Theory and Evidence," *Review of Economics and Statistics*, 2016, pp. 1–55.

____ **and** ____ , "Criminal Deterrence: A Review of the Literature," *Journal of Economic Literature*, 2016, pp. 1–60.

Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Gui Woolston, "Does State Fiscal Relief During Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act," *American Economic Journal: Economic Policy*, August 2012, *4* (3), 118–145.

Cohen, Mark A and Alex R Piquero, "New Evidence on the Monetary Value of Saving a High Risk Youth," *Journal of Quantitative Criminology*, January 2009, *25* (1), 25–49.

Conley, Timothy G and Bill Dupor, "The American Recovery and Reinvestment Act: Solely a Government Jobs Program?," *Journal of Monetary Economics*, July 2013, *60* (5), 535–549.

Conover, Christopher, "Congress Should Account for the Excess Burden of Taxation," *Cato Institute Policy Analysis*, October 2010, *669*, 1–12.

Corman, Hope and Naci Mocan, "Carrots, Sticks, and Broken Windows," *The Journal of Law and Economics*, April 2005, *48* (1), 235–266.

DeAngelo, Gregory and Benjamin Hansen, "Life and Death in the Fast Lane: Police Enforcement and Traffic Fatalities," May 2014, *6* (2), 231–257.

Donohue, John, "Assessing the Relative Benefits of Incarceration: Overall Changes and the Benefits on the Margin," in Steven Raphael and Michael Stoll, eds., *Do Prisons Make Us Safer*, 2009, pp. 269–341.

____ **and Jens Ludwig**, "More COPS," *Brookings Institution Policy Brief*, March 2007, pp. 1–7.

Draca, Mirko, Stephen Machin, and Robert Witt, "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks," *American Economic Review*, August 2011, *101* (5), 2157–2181.

Durlauf, Steven N and Daniel S Nagin, "Imprisonment and Crime: Can Both be Reduced?," *Criminology and Public Policy*, January 2011, *10* (1), 13–54.

Evans, William N and Emily G Owens, "COPS and Crime," *Journal of Public Economics*, February 2007, *91* (1), 181–201.

Garrett, Thomas and Gary Wagner, "Red Ink in the Rearview Mirror: Local Fiscal Conditions and the Issuance of Traffic Tickets," *The Journal of Law and Economics*, February 2009, *52* (1), 71–90.

Hines, James and Richard Thaler, "Anomalies: The Flypaper Effect," *Journal of Economic Perspectives*, October 1995, *9* (4), 217–226.

Hoekstra, Mark, "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*, November 2009, *91* (1), 717–724.

Imbens, G and K Kalyanaraman, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies*, July 2012, *79* (3), 933–959.

Jackson, C Kirabo, Rucker C Johnson, and Claudia Persico, "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms," *The Quarterly Journal of Economics*, February 2016, *131* (1), 157–218.

James, Nathan, "Community Oriented Policing Services (COPS): Background and Funding," *Congressional Research Service*, May 2013, pp. 1–14.

Klick, Jonathan and Alexander Tabarrok, "Using Terror Alert Levels to Estimate the Effect of Police on Crime," *The Journal of Law and Economics*, April 2005, *48* (1), 267–279.

____ **and** ____ , "Police, Prisons, and Punishment: Empirical Evidence on Crime Deterrence," in Bruce Benson and Paul Zimmerman, eds., *Handbook on the Economics of Crime*, Edward Elgar, 2010, pp. 127–144.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach, "Shool Finance Reform and the Distribution of Student Acheivement," *NBER Working Paper*, July 2016, pp. 1–86.

Lee, David S and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, June 2010, *48* (2), 281–355.

Levitt, Steven, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, June 1997, *87*, 270–290.

____ , "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?," *Economic Inquiry*, 1998, *36* (3), 353–372.

____ , "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply," *American Economic Review*, September 2002, *92* (4), 1244–1250.

____ **and Thomas Miles**, "Economic Contributions to the Understanding of Crime," *Annual Review of Law and Social Science*, December 2006, *2* (1), 147–164.

____ **and** ____ , "Empirical Study of Criminal Punishment," in A Mitchell Polinsky and Steven Shavell, eds., *Hanbook of Law and Economics*, Elsevier, 2007, pp. 455–495.

Lin, Ming-Jen, "More Police, Less Crime: Evidence from US State Data," *International Review of Law and Economics*, June 2009, *29* (2), 73–80.

MacDonald, John, Jeffrey Fagan, and Amanda Geller, "The Effects of Local Police Surges on Crime and Arrests in New York City," *Columbia Public Law Research Paper No. -*, October 2015, pp. 1–43.

____ **, Jonathan Klick, and Ben Grunwald**, "The Effect of Privately Provided Police Services on Crime," *Institute of Law and Economics Research Paper*, November 2012, *12-36*, 1–26.

Machin, Stephen and Olivier Marie, "Crime and Police Resources: The Street Crime Initiative," *Journal of the European Economic Association*, March 2011, *9* (4), 678–701.

Marvell, Thomas and Carlisle Moody, "Specification Problems, Police Levels, and Crime Rates," *Criminology*, November 1996, *34* (4), 609–646.

Mas, Alexandre, "Pay, Reference Points, and Police Performance," *Quarterly Journal of Economics*, August 2006, *121* (3), 783–821.

McCrary, Justin, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment," *American Economic Review*, November 2002, *92*, 1236–1243.

____ , "Manipulation of the Running variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, February 2008, *142* (2), 698–714.

____ , "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police," *American Economic Review*, April 2009, *97* (1), 318–353.

Owens, Emily G, "More Time, Less Crime? Estimating the Incapacitative Effect of Sentence Enhancements," *The Journal of Law and Economics*, August 2009, *52* (3), 551–579.

____ , "COPS and Cuffs," in Phillip Cook, Stephen Machin, Olivier Marie, and Giovanni Mastrobouni, eds., *Lessons from the Economics of Crime: What Works in Reducing Offending*, 2012.

Tella, Rafael Di and Ernesto Schargrodsky, "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack," *American Economic Review*, March 2004, *94* (1), 115–133.

U.S. Government Accountability Office, "COPS Grants Were a Modest Contributor to Declines in Crime in the 1990s," *GAO Report*, October 2005, *06* (104), 1–124.

Vollaard, Ben and Joseph Hamed, "Why the Police Have an Effect on Violent Crime After All: Evidence from the British Crime Survey," *The Journal of Law and Economics*, November 2012, *55* (4), 901–924.

Weisburd, Sarit, "Police Presence, Rapid Response Rates, and Crime Prevention," *Unpublished Manuscript*, March 2016, pp. 1–59.

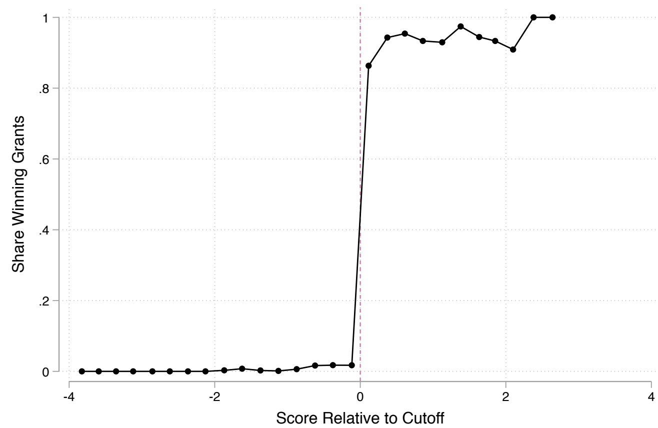
Whalen, Charles and Felix Reichling, "The Fiscal Multiplier and Economic Policy Analysis in the United States," *Congressional Budget Office Woring Paper Series*, February 2015, pp. 1–20.

Worrall, John, "The Effects of Local Law Enforcement Block Grants on Serious Crime," *Criminology and Public Policy*, August 2008, *7* (3), 325–350.

—— **and Tomislav Kovandzic**, "Police Levels and Crime Rates: An Instrumental Variables Approach," *Social Science Research*, May 2010, *39* (3), 506–516.

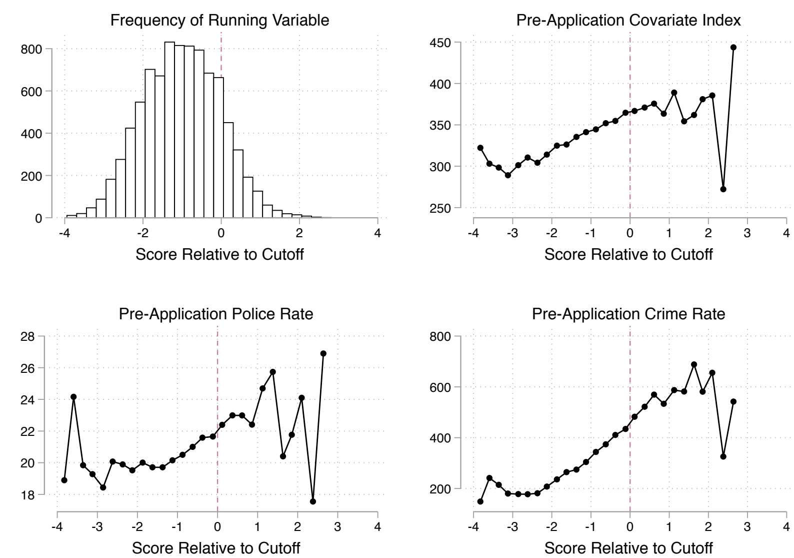
Zhao, Jihong, Matthew Scheider, and Quint Thurman, "Funding Community Policing to Reduce Crime: Have COPS Grants Made a Difference?," *Criminology and Public Policy*, January 2002, *2* (1), 7–32.

Figure 1: Discontinuity in Probability of Grant Receipt



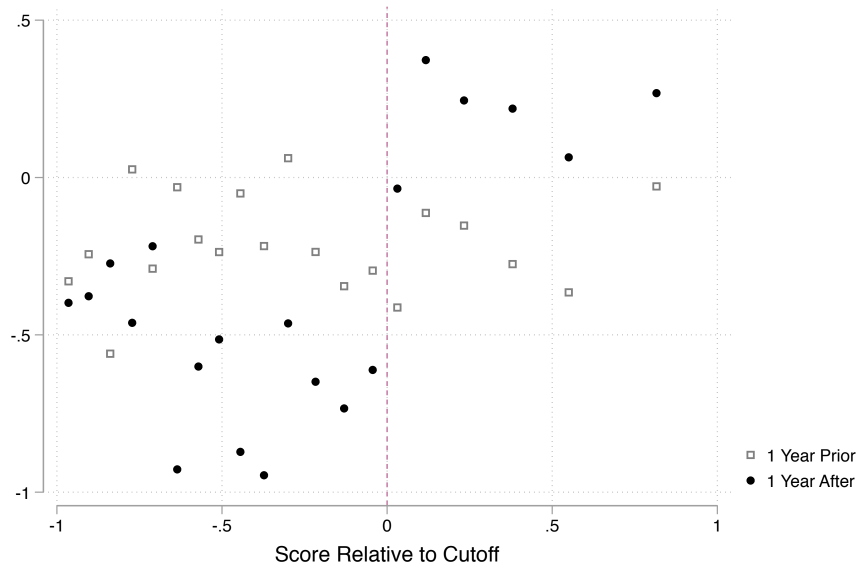
Notes: A unit is a city × application. Bin width equals 0.25. Figure shows share of applications in each bin that resulted in grants.

Figure 2: Characteristics of Applicants by Application Score



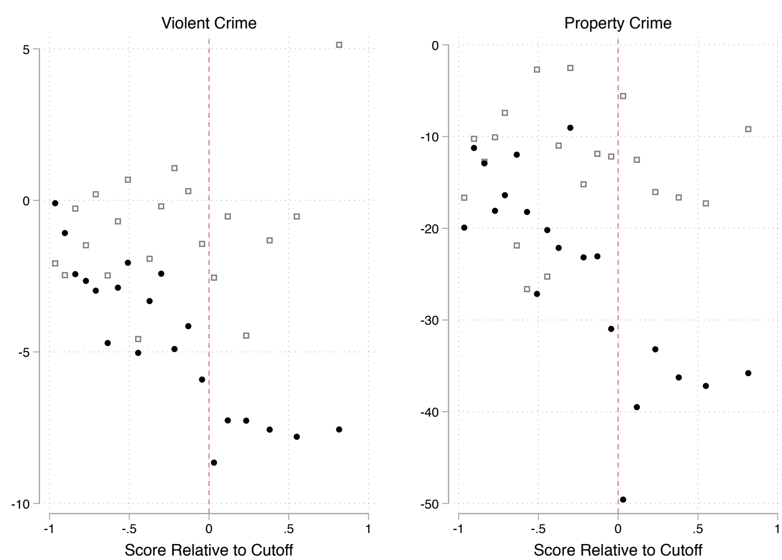
Notes: An observation is a city × application. Bin width equals 0.25. Covariate index is the predicted crime rate from a regression of crimes per 10,000 residents on controls and year fixed effects. Covariate index, police rate, and crime rate are plotted for one year prior to the application.

Figure 3: Two-Year Change in Police by Application Score



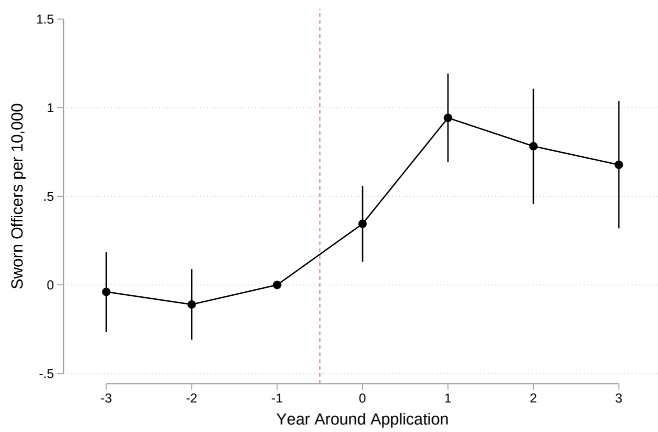
Notes: An observation is a city × application. Bin width selected to have equal mass. Plot shows the mean two-year change (i.e. change from $t-2$ to t) in police per 10,000 residents for one year prior and one year after the application year. Cities are weighted by one over the number of applications.

Figure 4: Two-Year Change in Crime by Application Score



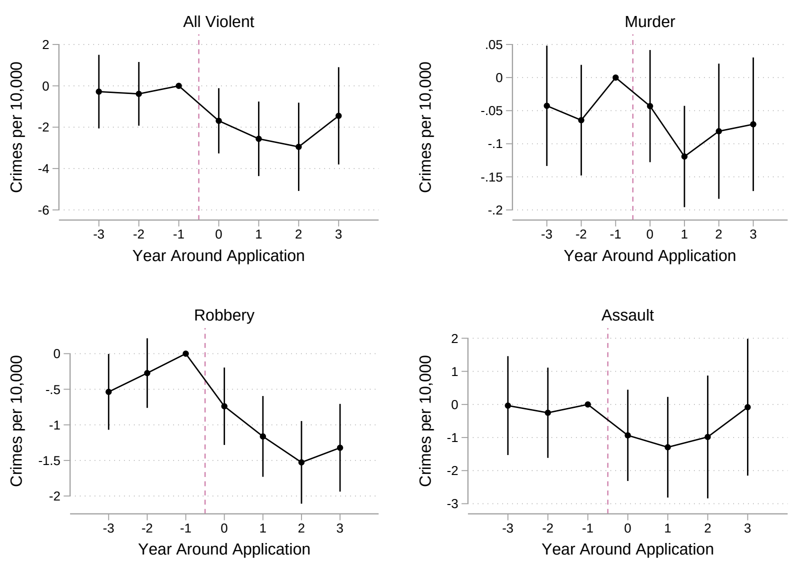
Notes: See Figure 3 for legend. An observation is a city × application. Bin width selected to have equal mass. Plot shows the mean two-year change (i.e. change from $t-2$ to t) in crimes per 10,000 residents for one year prior and one year after the application year. Cities are weighted by one over the number of applications.

Figure 5: Effect of Exceeding the Threshold on Police



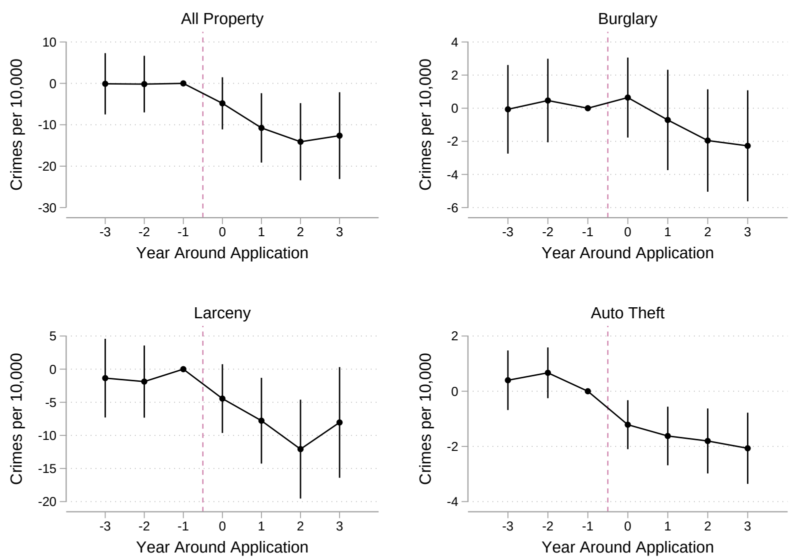
Notes: Dependent variable is sworn officers per 10,000 residents. 95% confidence bands from standard errors clustered at the city level. Figure plots coefficients on interaction between an event time indicator and a high score indicator. Regression includes controls, event time fixed effects, event time fixed effects interacted with the application score, city × application fixed effects and pre-program cell × year fixed effects. Cities are weighted by one over the number of applications. Regressions also include event time coefficients (and high score interactions) for 4+ years prior and 4+ years after. Corresponding coefficients shown in Table 4.

Figure 6: Effect of Exceeding the Threshold on Violent Crime



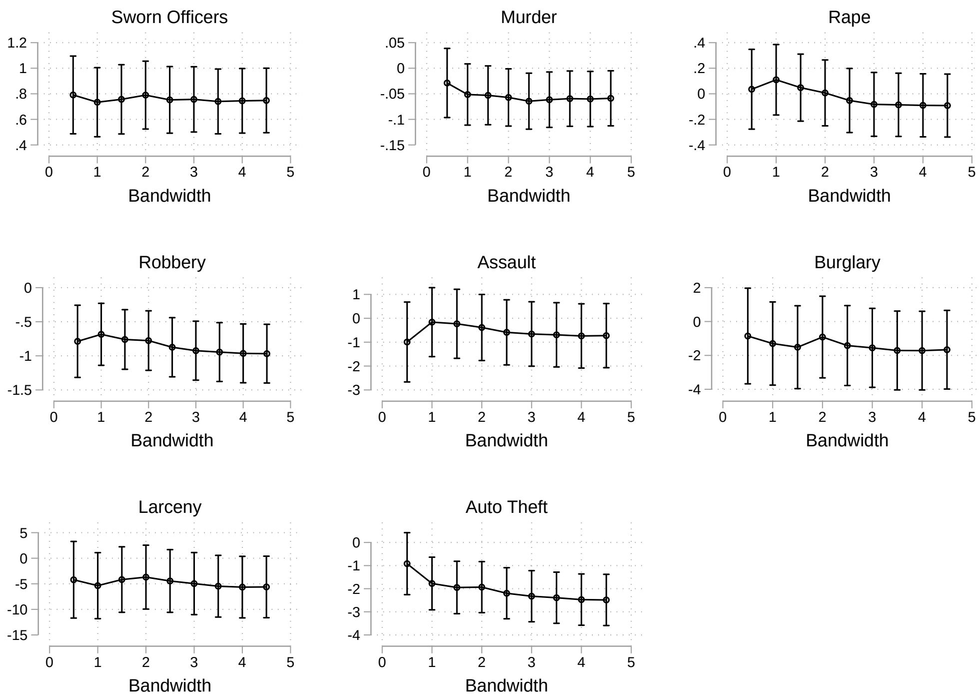
Notes: Dependent variable is crimes per 10,000 residents. 95% confidence bands from standard errors clustered at the city level. Figures plots coefficients on interactions between an event time indicator and a high score indicator. Regression includes controls, event time fixed effects, event time fixed effects interacted with the application score, city × application fixed effects and pre-program cell × year fixed effects. Cities are weighted by one over the number of applications. Regressions also include event time coefficients (and high score interactions) for 4+ years prior and 4+ years after. Corresponding coefficients shown in Table 4 and A-3.

Figure 7: Effect of Exceeding the Threshold on Property Crime



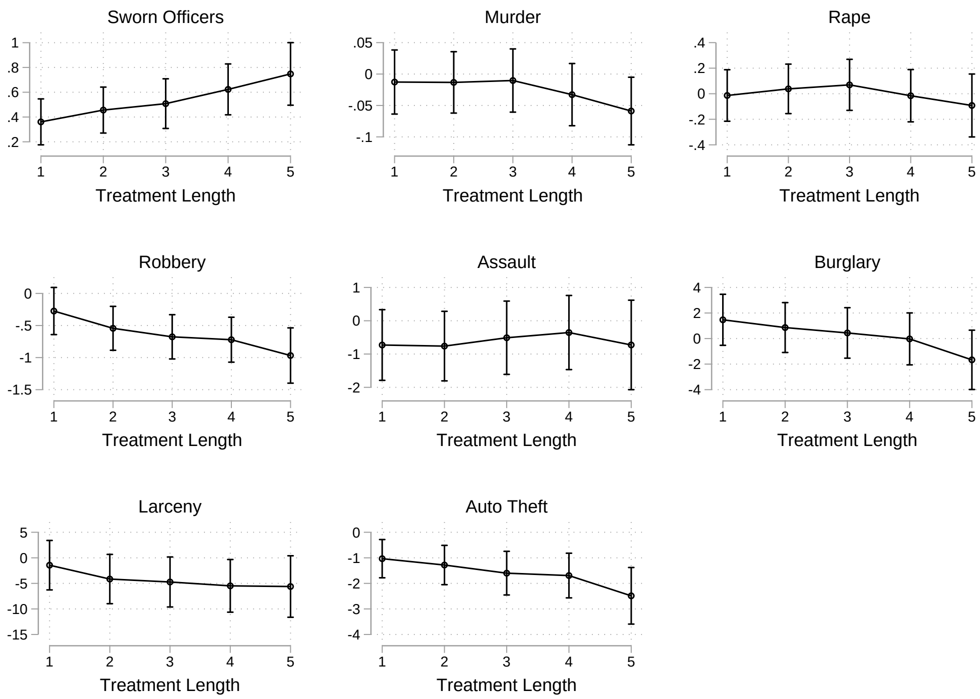
Notes: Dependent variable is crimes per 10,000 residents. 95% confidence bands from standard errors clustered at the city level. Figures plots coefficients on interactions between an event time indicator and a high score indicator. Regression includes controls, event time fixed effects, event time fixed effects interacted with the application score, city × application fixed effects and pre-program cell × year fixed effects. Cities are weighted by one over the number of applications. Regressions also include event time coefficients (and high score interactions) for 4+ years prior and 4+ years after. Corresponding coefficients shown in Table 4 and A-3.

Figure 8: Sensitivity of First Stage and Reduced Form Estimates to Varying the Bandwidth



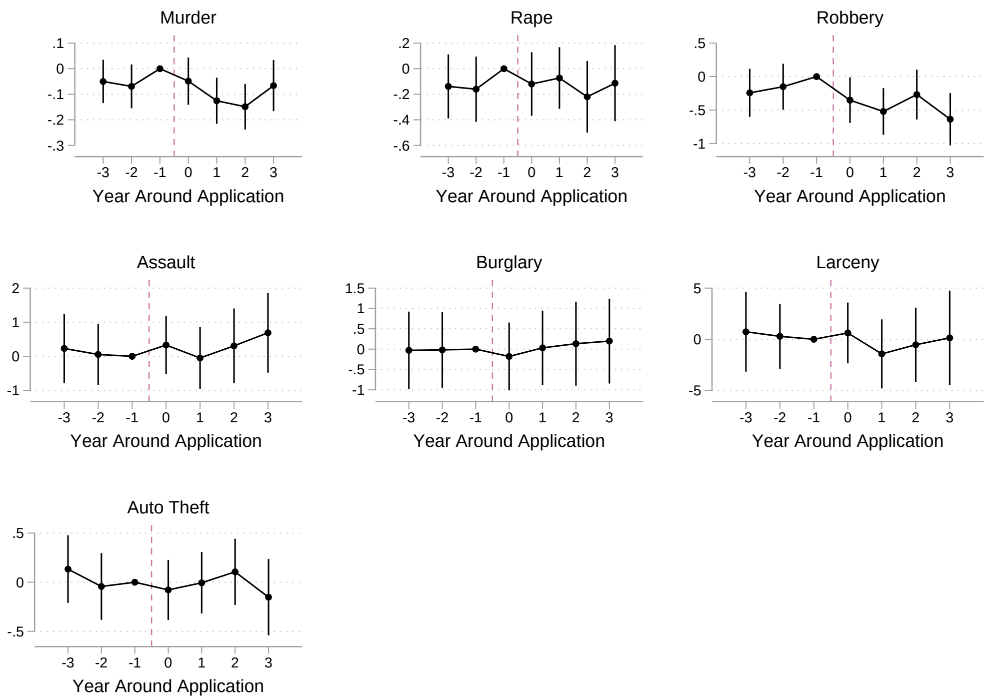
Notes: Each figure plots the coefficients on *Post* $\times$ *High* from regressions where police (crimes) per 10,000 residents is the dependent variable and only applications with scores within the indicated bandwidth of the cutoff are used. 95% confidence bands are from standard errors clustered at the city level. All regressions includes controls, city $\times$ application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year $\times$ pre-program cell fixed effects. Cities are weighted by one over the number of applications.

Figure 9: First Stage and Reduced Form Estimates by Varying Definitions of *Post*



Notes: Each figure plots the coefficients on *Post* $\times$ *High* from regressions where police (crimes) per 10,000 residents is the dependent variable and only applications with scores within the indicated bandwidth of the cutoff are used. 95% confidence bands are from standard errors clustered at the city level. All regressions includes controls, city $\times$ application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year $\times$ pre-program cell fixed effects. Cities are weighted by one over the number of applications.

Figure 10: Effect of Exceeding the Threshold on Arrests



Notes: Dependent variable is arrests per 10,000 residents. 95% confidence bands from standard errors clustered at the city level. Figures plots coefficients on interactions between an event time indicator and a high score indicator. Regression includes controls, event time fixed effects, event time fixed effects interacted with the application score, city × application fixed effects and pre-program cell × year fixed effects. Cities are weighted by one over the number of applications.

Table 1: Breakdown of Sample by City Size and Application Year

		Applications				
Size Group	Unique Cities	2009	2010	2011	2012	2013
1,000-2,500	475	421	241	79	45	53
2,500-5,000	687	587	340	145	77	104
5,000-10,000	941	804	477	221	130	195
10,000-25,000	1204	1032	652	331	199	309
25,000-50,000	594	528	304	173	111	167
50,000-100,000	350	312	169	153	76	105
100,000+	123	99	48	45	23	49
Total	4374	3783	2231	1147	661	982

Notes: Cities observed in multiple categories are placed in the group they appear in most often. The groups indicated are those used to construct the pre-program cell $\times$ year fixed effects except that the 50,000-100,000 and 100,000+ groups are pooled together.

Table 2: Summary Statistics for Applicant Cities

	High Score	Low Score	Total
Population	3.992	2.058	2.438
	(10.76)	(4.503)	(6.297)
Per Capita Income	3.704	4.027	3.963
	(1.041)	(1.171)	(1.154)
Unemployment Rate	7.558	7.504	7.515
	(2.977)	(2.689)	(2.748)
Percent Black	0.133	0.101	0.108
	(0.151)	(0.111)	(0.120)
Percent Hispanic	0.121	0.108	0.111
	(0.153)	(0.129)	(0.134)
Percent Age 15-24	0.140	0.138	0.139
	(0.0288)	(0.0316)	(0.0311)
Sworn Officer Rate	23.71	20.66	21.26
	(9.160)	(8.684)	(8.862)
Violent Crime Rate	66.36	27.37	35.04
	(48.77)	(28.41)	(36.82)
Property Crime Rate	475.0	274.9	314.2
	(222.0)	(167.4)	(196.3)

Notes: Standard deviations in parentheses. An observation is a city $\times$ application. Number of observations by group: 1,371 (High Score); 7,433 (Low Score); 8,804 (Total). There are 4,374 unique cities and cities are weighted by one over the number of applications. Characteristics reported as of the year prior to application.

Table 3: Regression Discontinuity Estimates

	Pr(Win)	Grant Funding	Police	Violent	Property
Above Threshold	0.69***	50998***	0.633***	-1.678	-10.456
	(0.03)	(4977)	(0.201)	(1.64)	(7.3)
Mean	.15	16703	21.85	46.32	425.48
IK Bandwidth	.39	.78	.93	.72	.79
Cities	1343	2223	2510	2098	2245
Observations	1766	3371	3947	2887	3198

Notes: Robust standard errors from Calonico et al. (2014) in parentheses. Bandwidth choice is the optimal bandwidth from Imbens and Kalyanaraman (2012). Dependent variable in Column 1 is an indicator for grant receipt. Dependent variable in Columns 2-5 is the change in y per 10,000 residents between one year prior and one year after program application. Application score is controlled for via local linear regression with a triangular kernel.

Table 4: Effect of Exceeding the Threshold on Police and Crime

	Police	Violent Crime	Property Crime
4+ Years Prior	0.0328	-0.280	-2.321
	(0.117)	(0.971)	(4.130)
3 Years Prior	-0.0391	-0.173	-0.0973
	(0.115)	(0.900)	(3.774)
2 Years Prior	-0.110	-0.330	-0.180
	(0.101)	(0.785)	(3.488)
Application Year	0.345***	-1.711**	-4.823
	(0.109)	(0.805)	(3.211)
1 Year After	0.943***	-2.518***	-10.77**
	(0.128)	(0.919)	(4.273)
2 Years After	0.783***	-2.896***	-14.13***
	(0.165)	(1.087)	(4.756)
3 Years After	0.678***	-1.382	-12.65**
	(0.183)	(1.197)	(5.349)
4+ Years After	0.882***	-2.497*	-16.62**
	(0.201)	(1.331)	(7.105)
Mean	22.54	48.05	433.58
Cities	4201	4201	4201
City x Applications	8459	8459	8459
Observations	84590	78993	80587

Notes: Standard errors clustered at the city level in parentheses. Dependent variable is police (crimes) per 10,000 residents. Table displays coefficient on event time indicators interacted with an indicator for whether the application score at time zero exceeds the threshold. Each regression includes controls, city $\times$ application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year $\times$ pre-program cell fixed effects. Cities are weighted by one over the number of applications.

Table 5: Effects of Grants on Police and Crime

	Grant Funding	Police	Violent Crime	Property Crime
Post	-2298.4***	0.226	0.525	3.509
	(659.1)	(0.166)	(1.114)	(6.039)
Post x High	50926.6***	0.748***		
	(1552.9)	(0.129)		
Police			-2.904**	-16.23**
			(1.244)	(6.412)
Mean	-	22.54	48.05	433.58
Elasticitiy	-	-	-1.36	-.84
F-Stat	-	-	31.19	29.76
Cities	4201	4201	4201	4201
City x Applications	8459	8459	8459	8459
Observations	84590	84590	78993	80587

Notes: Standard errors clustered at the city level in parentheses. Table presents IV estimates corresponding to equation (4). Each regression includes controls, city $\times$ application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year $\times$ pre-program cell fixed effects. Cities are weighted by one over the number of applications.

Table 6: IV Estimates of the Effect of Police on Individual Crime Types

	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto Theft
Police	-0.0896**	-0.138	-1.390***	-1.127	-2.464	-8.579*	-3.536***
	(0.0420)	(0.179)	(0.383)	(1.032)	(1.738)	(4.778)	(0.979)
Mean	.42	4.08	10.69	33.16	86.8	311.12	36.52
Elasticitiy	-4.8	-.76	-2.93	-.77	-.64	-.62	-2.18
F-Stat	28.52	28.32	31.77	26.77	29.4	28.74	31.66
Cities	4205	4203	4203	4090	4109	4109	4110
City x Applications	8473	8465	8468	8258	8297	8290	8297
Observations	79850	78462	78981	77141	78190	79286	77404

Notes: Standard errors clustered at the city level in parentheses. Table presents IV estimates corresponding to equation (4). Each regression includes controls, city $\times$ application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year $\times$ pre-program cell fixed effects. Cities are weighted by one over the number of applications.

Table 7: Robustness Checks

	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto Theft
Same as Table 6	-.09**	-.14	-1.39***	-1.13	-2.46	-8.58*	-3.54***
	(.04)	(.18)	(.38)	(1.03)	(1.74)	(4.78)	(.98)
	[28.52]	[28.32]	[31.77]	[26.77]	[29.4]	[28.74]	[31.66]
Balanced Samples	-.09**	-.05	-2.25**	-.9	-2.51	-14.5**	-4.46***
	(.04)	(.21)	(.89)	(1.35)	(1.92)	(6.95)	(1.65)
	[11.98]	[11.8]	[9.550]	[12.53]	[23.46]	[16.55]	[14.05]
Balanced Panels	-.05	-.13	-1.38***	-1.72*	-1.4	-8.70*	-3.59***
	(.05)	(.16)	(.36)	(1.02)	(1.69)	(4.53)	(.93)
	[32.22]	[34.54]	[36.07]	[32.99]	[32.01]	[31.91]	[35.25]
Pop >10,000 Only	-.07	-.25	-1.76**	-2.41	-.9	-9.52	-5.33**
	(.05)	(.28)	(.74)	(1.88)	(3.01)	(8.359)	(2.08)
	[13.57]	[14.09]	[17.1]	[11.91]	[14.19]	[12.9]	[16.65]
Flexible Score Controls	-.11*	-.28	-.88*	-1.81	-3.15	-7.71	-1.81
	(.06)	(.23)	(.51)	(1.69)	(2.8)	(6.53)	(1.23)
	[15.63]	[15.85]	[15.95]	[13.18]	[13.4]	[16.24]	[16.77]
Flexible + Close	-.12*	-.24	-1.16**	-2.74	-3.94	-8.43	-2.6*
	(.07)	(.25)	(.58)	(1.81)	(3.09)	(6.29)	(1.36)
	[14.49]	[15.4]	[14.36]	[12.88]	[12.92]	[18.06]	[15.64]

Notes: Each coefficient is from a separate IV regression. Standard errors clustered at the city level in parentheses. First-stage F-statistic for the corresponding regression in brackets. All regressions include controls, city $\times$ application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year $\times$ pre-program cell fixed effects. Cities are weighted by one over the number of applications. *Balanced Samples* uses only cities with valid data for that crime type in all years. *Balanced Panels* drops 2013 applications and uses only data between 4 years prior and 2 years after the application. *Flexible Score Controls* allows the event time $\times$ score effect to vary by whether the score exceeds the threshold. *Flexible + Close* repeats this specification using only applications within 1 point of the threshold.

Table 8: IV Estimates of the Effect of Police on Arrests

	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto Theft
Police	-0.0686	-0.00820	-0.326*	0.445	0.490	-0.538	-0.0890
	(0.0424)	(0.105)	(0.168)	(0.615)	(0.483)	(2.204)	(0.177)
Mean	.24	.9	2.68	14.53	10.22	48.85	3.54
Elasticitiy	-6.04	-.19	-2.57	.65	1.01	-.23	-.53
F-Stat	21.48	22.59	28.42	23.03	22.74	23.14	24.91
Crime Effect	-.05	-.22	-.97	-.3	-2.6	-8.43	-3.22
Cities	3507	3506	3506	3488	3496	3496	3497
City x Applications	7110	7105	7108	7078	7094	7087	7094
Observations	67172	66365	66452	65925	65857	66041	65692

Notes: Same as Table 6 except that the dependent variable is arrests per 10,000 residents. *Crime Effect* reports the coefficient on police from a regression where crimes per 10,000 residents is the dependent variable and the sample is the same as that used to estimate the arrest effect. The crime estimates for robbery, larceny, and auto theft are statistically significant. The p-values for the crime estimates (in order) are 0.224, 0.269, 0.006, 0.778, 0.165, 0.089, and 0.001.

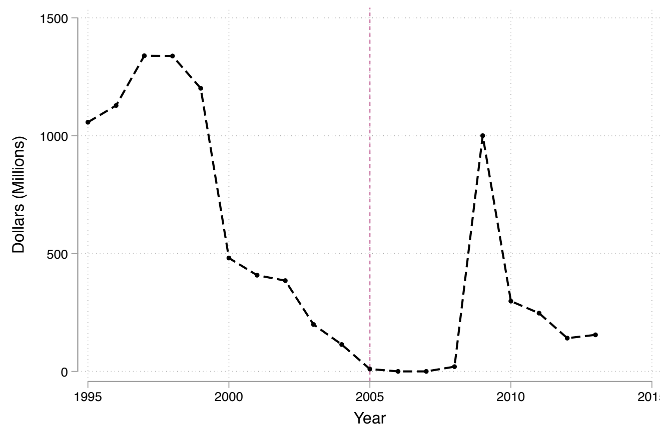
Table 9: Victimization Costs of Crime Types

Crime Type	*Victimization Cost*
Murder	$7,000,000
Rape	$142,020
Robbery	$12,624
Assault	$38,924
Burglary	$2,104
Larceny	$473
Auto Theft	$5,786

Notes: Costs of non-murder crime victimization taken from Cohen and Piquero (2009). Cost of murder taken from standard VSL estimates in the literature. $7m is the dollar value used in (Chalfin and McCrary 2016a).

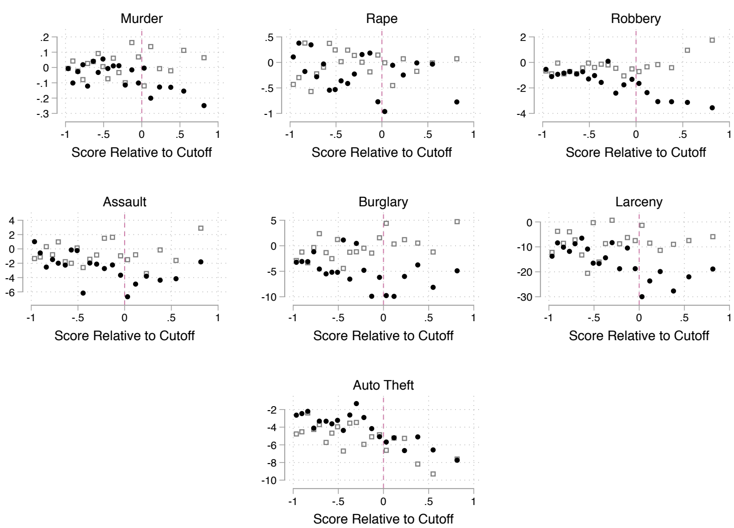
Appendix For Online Publication

Figure A-1: Yearly Appropriations for COPS Hiring Program, FY 1995-2013



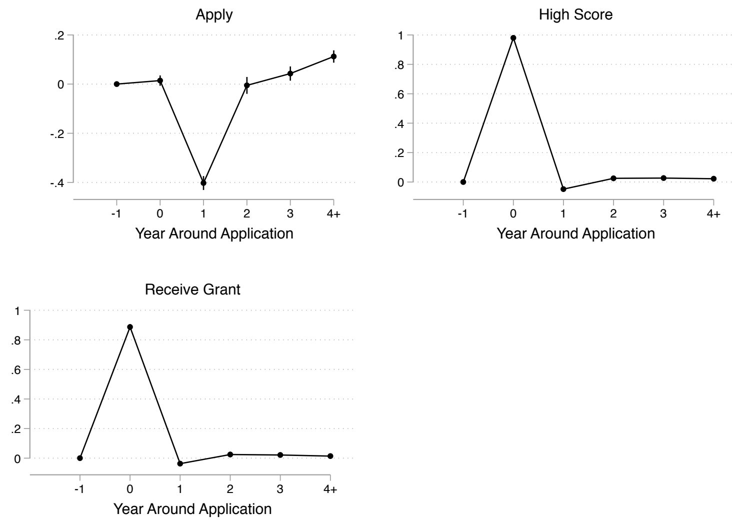
Notes: Appropriations data from James (2013). Dashed line denotes 2005, my first year of data used in the regressions.

Figure A-2: Simple RD Plots for Individual Crime Types



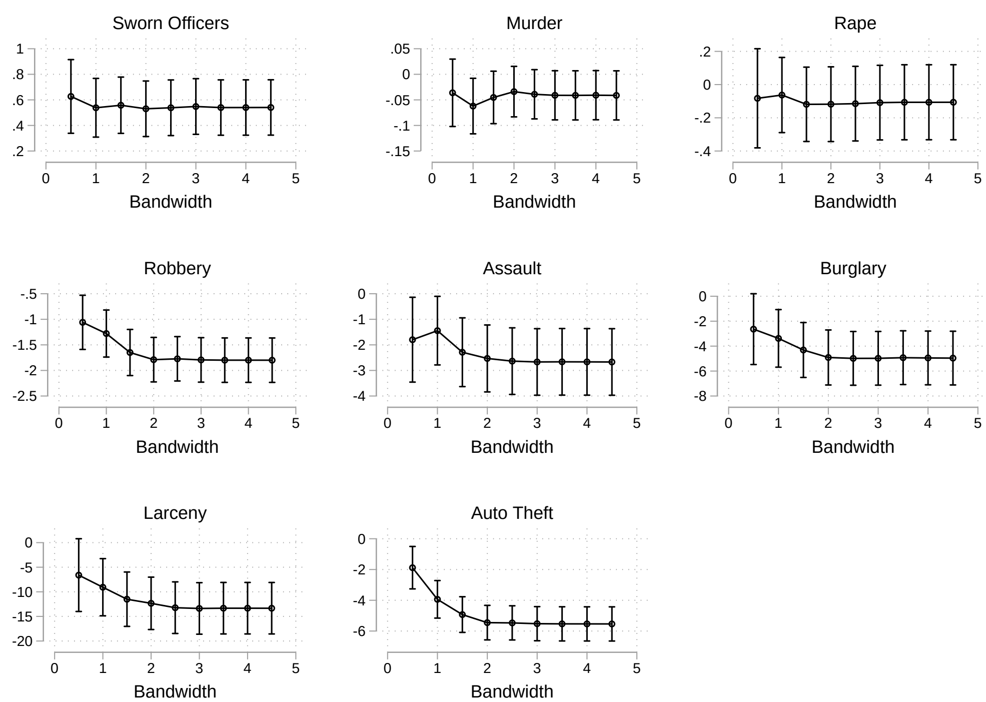
Notes: See Figure 3 for legend. An observation is a city × application. Bin width selected to have equal mass. Plot shows the mean two-year change (i.e. change from $t-2$ to t) in crimes per 10,000 residents for one year prior and one year after the application year. Cities are weighted by one over the number of applications.

Figure A-3: Effect of Exceeding the Threshold on Future Program Activity



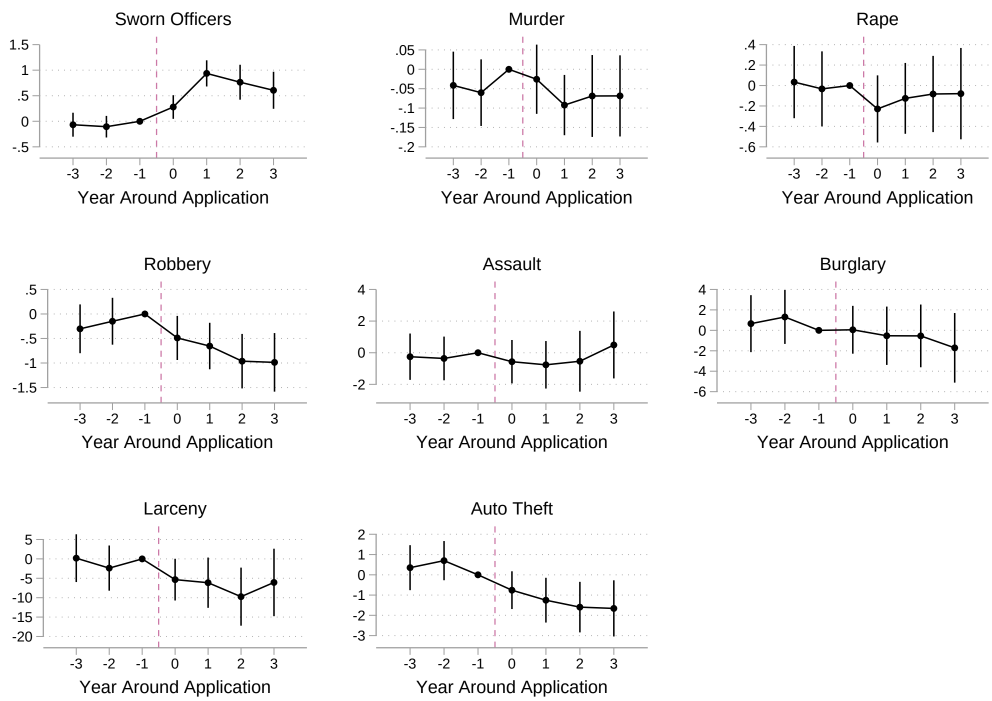
Notes: Figure plots coefficients on event time indicators interacted with whether the time zero application exceeded the threshold. Each regression includes controls, city × event fixed effects, event time fixed effects, and city size × year fixed effects.

Figure A-4: First Stage and Reduced Form Estimates without Score Controls, Varying Bandwidths



Notes: Same as Figure 6 except without event time × score interactions.

Figure A-5: Police and Crime Event Study Plots for Arrests Sample



Notes: Each figure plots the coefficients on *Post* × *High* from regressions where police (crimes) per 10,000 residents is the dependent variable and only applications with scores within the indicated bandwidth of the cutoff are used. 95% confidence bands are from standard errors clustered at the city level. All regressions includes controls, city × application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year × pre-program cell fixed effects. Cities are weighted by one over the number of applications.

Table A-1: Regression Discontinuity Specification Checks

	McCrary Test	Cov Index	Police	Violent	Property
Above Threshold	-0.072	-1.66	0.37	2.9	16.9
	(0.076)	(5.16)	(0.54)	(2.84)	(12.2)
Mean		357.87	21.58	44.83	378.93
IK Bandwidth	.67	.86	1.21	.83	1
Observations	2927	3686	4948	3405	4078

Notes: Column 1 reports coefficient (standard error) from the McCrary (2008) test for discontinuity in density. Columns 2-5 report RD estimates where y, measured one year prior to application, is the dependent variable. Robust standard errors from Calonico et al. (2014) in parentheses. Bandwidth choice is the optimal bandwidth from Imbens and Kalyanaraman (2012). Application score is controlled for via local linear regression with a triangular kernel.

Table A-2: Regression Discontinuity Estimates for Individual Crime Types

	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto Theft
Above Threshold	-0.007	-0.241	-0.327	-1.944	-1.19	-7.398	-0.331
	(0.072)	(0.335)	(0.446)	(1.43)	(2.3)	(5.66)	(1.617)
Mean	.39	3.98	10.46	31.74	84.92	304.93	36.7
IK Bandwidth	.87	.63	.87	.77	.8	.79	.77
Cities	2377	1931	2377	2199	2262	2245	2199
Observations	3314	2477	3366	3022	3179	3204	2973

Notes: Robust standard errors from Calonico et al. (2014) in parentheses. Bandwidth choice is the optimal bandwidth from Imbens and Kalyanaraman (2012). Dependent variable is the change in crimes per 10,000 residents between one year prior and one year after program application. Application score is controlled for via local linear regression with a triangular kernel.

Table A-3: Effect of Exceeding the Threshold on Individual Crime Types

	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto Theft
4+ Years Prior	-0.0477	-0.188	-0.170	-0.233	-0.970	-4.017	1.139*
	(0.0401)	(0.144)	(0.287)	(0.747)	(1.237)	(3.279)	(0.650)
3 Years Prior	-0.0421	0.0117	-0.519*	0.0329	-0.0444	-1.350	0.394
	(0.0464)	(0.167)	(0.271)	(0.759)	(1.365)	(3.032)	(0.551)
2 Years Prior	-0.0662	-0.0323	-0.254	-0.219	0.466	-1.881	0.663
	(0.0426)	(0.172)	(0.249)	(0.695)	(1.287)	(2.780)	(0.469)
Application Year	-0.0436	-0.203	-0.720***	-0.969	0.687	-4.440*	-1.214***
	(0.0432)	(0.160)	(0.277)	(0.702)	(1.231)	(2.648)	(0.452)
1 Year After	-0.120***	-0.0987	-1.143***	-1.274	-0.712	-7.791**	-1.622***
	(0.0391)	(0.168)	(0.289)	(0.776)	(1.547)	(3.303)	(0.542)
2 Years After	-0.0804	-0.0187	-1.509***	-0.948	-1.909	-12.08***	-1.804***
	(0.0521)	(0.184)	(0.296)	(0.946)	(1.576)	(3.809)	(0.601)
3 Years After	-0.0685	-0.0347	-1.301***	-0.0371	-2.240	-8.056*	-2.069***
	(0.0514)	(0.215)	(0.314)	(1.054)	(1.709)	(4.262)	(0.658)
4+ Years After	-0.168***	-0.437*	-1.365***	-0.948	-5.475**	-7.069	-2.628***
	(0.0546)	(0.246)	(0.337)	(1.133)	(2.131)	(5.655)	(0.740)
Mean	.42	4.08	10.69	33.16	86.82	311.12	36.52
Cities	4199	4197	4197	4085	4103	4106	4105
City x Applications	8457	8449	8452	8247	8281	8283	8283
Observations	79711	78329	78851	77046	78063	79225	77288

Notes: Standard errors clustered at the city level in parentheses. Dependent variable is police (crimes) per 10,000 residents. Table displays coefficient on event time indicators interacted with an indicator for whether the application score at time zero exceeds the threshold. Each regression includes controls, city $\times$ application fixed effects, event time fixed effects, event time fixed effects interacted with the application score, and year $\times$ pre-program cell fixed effects. Cities are weighted by one over the number of applications.

Table A-4: Balance Check

	High Score	High Score	High Score
Log Population	0.0396***	0.0650***	0.0440***
	(0.00559)	(0.0173)	(0.0158)
Log Per Capita Income	-0.0972***	-0.0767***	0.00477
	(0.0219)	(0.0225)	(0.0189)
Unemployment Rate	-0.00832***	-0.00831***	-0.0149***
	(0.00198)	(0.00196)	(0.00174)
Percent Age 15-24	-0.411***	-0.392***	-0.220
	(0.147)	(0.149)	(0.137)
Percent Black	-0.0914*	-0.0808*	-0.111***
	(0.0470)	(0.0472)	(0.0406)
Percent Hispanic	0.0364	0.0196	-0.0545*
	(0.0393)	(0.0397)	(0.0330)
Per Capita Sworn Officers	2.305	-0.356	-3.531
	(6.817)	(6.827)	(5.782)
Per Capita Violent Crimes	29.79***	29.06***	12.46***
	(2.251)	(2.236)	(1.760)
Per Capita Property Crimes	4.394***	4.431***	1.304***
	(0.408)	(0.414)	(0.342)
Application Score			0.218***
			(0.00608)
Joint F-Stat	129.25	94.26	24.22
Group FE	No	Yes	Yes
Observations	7984	7984	7984

Notes: Robust standard errors in parentheses.