THE EFFECTS OF POLICE SPENDING ON CRIME: EVIDENCE FROM A REGRESSION-DISCONTINUITY DESIGN

December 29, 2015

Philip J. Cook
Duke University and NBER

Jens Ludwig
University of Chicago and NBER

Douglas L. Miller
University of California, Davis and NBER

The authors are listed alphabetically, and all contributed equally to this study. This paper was supported by grant # 2011CKWX0098 from the U.S. Department of Justice Office of Community Oriented Policing Services, by operating support to the University of Chicago Crime Lab from the Joyce, MacArthur, and McCormick foundations, and by visiting scholar awards to Ludwig from the Russell Sage Foundation and Sciences Po LIEPP. Thanks to Matthew Scheider for very helpful discussions and to Roseanna Ander for all of her support for this project. We are grateful to Seth Bour, Stephen Coussens and particularly Samantha Edds for their outstanding help assembling and analyzing the data. All opinions and any errors are our own.

THE EFFECTS OF POLICE SPENDING ON CRIME:
EVIDENCE FROM A REGRESSION-DISCONTINUITY DESIGN

ABSTRACT

In this paper we estimate the effects on crime from increased policing spending by studying the 2009 competition for police hiring grants carried out by the US Department of Justice's COPS office. The question is important for understanding whether current levels of spending on police in the US are optimal. Our study takes advantage of the fact that the COPS office had substantial excess demand for grant funding in 2009, receiving $8 billion worth of requests for the $1 billion in grant funding available. Applications were scored using a formula that weighted together crime, financial distress and other local conditions, and then essentially rank-ordered and funded based on this score. This creates a regression discontinuity (RD) design that, we show, increased police spending per capita in the short term by 0.8 percent in jurisdictions with scores near the funding threshold, which remained at 0.7 percent several years later. We show that this discontinuity in police spending translated into a decline in total UCR part 1 offenses equal to between 1.1 and 1.8 percent, a decline in part 1 violent crimes equal to between 2.4 and 3.1 percent, and a decline in part 1 property crimes equal to 1.2 and 2 percent.

Philip J. Cook
Sanford School of Public Policy
Duke Box 90239
Duke University
Durham, NC 27708-0239
and NBER
pcook@duke.edu

Jens Ludwig
Harris School of Public Policy
1155 East 60th Street
University of Chicago
Chicago, IL 60637
and NBER
jludwig@uchicago.edu

Douglas Miller
Department of Economics
One Shields Avenue
University of California at Davis
Davis, CA 95616-8578
and NBER
dlmiller@ucdavis.edu

I. INTRODUCTION

Crime is a major social problem, with social costs in developed countries equivalent to as much as 10 percent of GDP (Anderson, 1999; Entorf and Spenger, 2002; Ludwig, 2006).[1] Despite the drop in crime rates during the last two decades, our rates of crime – particularly the most socially costly violent crimes – remain far above those of other developed nations. For example, the homicide rate per 100,000 in 2010 in the U.S. (4.8) was four times that of the U.K. (1.2) and fully six times that of Germany or Japan (0.8). Crime is also massively regressive in its impacts. Among males 15-24, the 2009 homicide rate for blacks was fully 18 times that of whites (74 versus 4/100,000).[2]

One of the primary ways in which every modern society tries to control crime is through policing, on which the U.S. currently spends over $110 billion each year.[3] The canonical economic model of crime from Becker (1968) predicts that police deter criminal behavior by increasing the expected costs of punishment to potential offenders. Police may also reduce crime by making arrests that result in the incarceration and hence incapacitation of active offenders, or by engaging in preventive problem-solving activities. While the US has increased spending on police over the past several decades, the growth in police per capita pales in comparison to the growth in the US incarceration rate. From 1980 to 2010 the number of people behind bars (prison or jail) increased from 500,000 to 2.2 million (Sourcebook of Criminal Justice Statistics, 2012), while total

[1] Crime remains an even bigger problem in many developing country contexts, with 2010 homicide rates in places like El Salvador and Honduras equal to 66 and 78 per 100,000 – that is, the country-wide homicide rate in those countries is about what we see among the highest-risk population in the U.S.
[2] http://www.cdc.gov/injury/wisqars/fatal_injury_reports.html
[3] http://www.albany.edu/sourcebook/pdf/t122006.pdf, converted to constant 2012 dollars by the authors.

spending on corrections more than tripled to $77 billion.[4] Given the possibility of diminishing marginal returns raises the possibility of shifting resources away from imprisonment towards police to reduce crime at no extra cost. Yet the effects of additional police spending in the US remain uncertain.

As part of a broader violence-reduction initiative in 1994, then-President Bill Clinton announced that the federal government would make grants to local jurisdictions to help them hire 100,000 additional police officers. The Community Oriented Policing Services (COPS) office was created to manage that program. The most thorough evaluation of that program found strong evidence that the grants were effective in stimulating additional hiring of uniformed officers, and that the result was a reduction in both property and violent crime (Evans and Owens 2007). After a latent period, COPS was revived under President Obama, and federal grants beginning in 2009 have helped a large number of jurisdictions hire more officers than their recession-limited budgets would have otherwise allowed. Even more than in the original version of COPS, Obama's version has been implemented in a systematic fashion that lends itself to a strong evaluation design to determine the causal effect of additional policing resources on crime rates. The results of this evaluation, which utilizes a regression-discontinuity (RD) design, are reported here.

Perhaps surprisingly, the claim that the police prevent crime remains somewhat controversial. Indeed, the conventional wisdom among criminologists as recently as 20 years ago was that the police have essentially no effective role in crime prevention (Bayley 1994). A large body of criminological research reported little systematic

[4] Trends in prison population taken from: http://www.albany.edu/sourcebook/pdf/t612010.pdf. Correctional spending is available through 2006 and is taken from: http://www.albany.edu/sourcebook/pdf/t122006.pdf, converted by the authors from nominal to constant 2012 dollars.

relationship between police spending and crime in cross-sectional data or over time (Chalfin and McCrary, forthcoming). If the political process tends to increase police budgets in response to an anticipated crime increase, then the simple correlation between crime and police may be positive even if police are in fact effective in crime prevention . The challenge of identifying the causal effect of police resources on crime from non-experimental data was well understood from the inception of the econometric study of crime (Fisher and Nagin, 1978).

One solution to this statistical challenge, implemented largely by the criminologists, has been field experiments. These experiments have not attempted to assess the marginal product of additional resources in reducing crime, but rather have tested the effects of changing how police use the resources they already have (Sherman and Rogan, 1995, Cohen and Ludwig, 2003, Skogan and Frydl, 2004; Braga, 2007, Kennedy, 2011, Sherman, 2011). Recent reviews have touted practices that concentrate resources on "hot spots" and put the police in a proactive problem solving role, while questioning the efficacy of some traditional deployments such as preventive patrol (Telep and Weisburd, 2012; Braga, Papachristos and Hureau 2012). One lesson some believe has been learned from these field experiments is that "…proactive, place-based and specific policing approaches appear much more promising in reducing crime than individual-based, reactive, and general ones" (Lum, Koper, and Telep 2011). One open question is whether these lessons are scalable in practice – that is, whether it is possible to change what police departments do on a day-to-day basis with their resources, and sustain those changes over time. In any case note that these field experiments do not speak directly to the question of whether it is worthwhile to provide police chiefs, who

may or may not follow the criminologists' guidance on resource allocation, with additional resources.

The other solution to possible simultaneity bias in estimating the effect of the police resources on crime requires exogenous sources of variation. Starting with the seminal paper by Levitt (1997), economists have tried to overcome identification problems by searching for various "natural experiments" in police resources, although as Durlauf and Nagin (2011) note (and as discussed further below), there necessarily remains some uncertainty with existing estimates. The recent paper with perhaps the strongest research design to date uses data from the re-allocation of police resources in London following a large-scale terrorist attack, which yields estimates that may or may not generalize to the US setting.

The contribution of the present paper is to present new (and we believe credible) estimates for the relationship between police spending and crime by taking advantage of the new wave of COPS funding. The COPS office received over $8 billion worth of funding requests for the roughly $1 billion of funding available. Proposals were scored based on a formula related to local crime rates and budget conditions, and for the most part funded in descending rank order until funding ran out.[5] We use this discontinuity in police spending across agencies to look for similar discontinuities in crime rates and, to determine whether deterrence is a key mechanism behind any crime impacts, on overall arrest rates as well.

Our paper thus builds on what we believe is one of the best previous natural-experiment studies of the police-crime relationship by Evans and Owens (2007) and

[5] There were actually some complications to this system from efforts to ensure adequate funding to each state and to small as well as large police agencies; we discuss these features of the scoring system and their implication for our analysis later in the paper.

Owens (2013), who examined the effects of COPS hiring grants under the Clinton Administration. Under President Clinton the COPS office did not use the same sort of funding formula that generates a discontinuity as in our study, and so Evans and Owens instead employ a difference-in-differences approach that compares before-after trends in crime across agencies that received relatively larger versus smaller COPS grants.

We show here that the 2009 COPS grant competition increased police spending per capita in the short term by 0.8 percent in jurisdictions with scores near the funding threshold, which remained at 0.7 percent several years later. Our estimates suggest that this discontinuity in police spending translated into a decline in total UCR part 1 offenses equal to between 1.1 and 1.8 percent, a decline in part 1 violent crimes equal to between 2.4 and 3.1 percent, and a decline in part 1 property crimes equal to 1.2 and 2 percent.

Our data are not as strong in helping us identify the mechanisms of action behind these crime drops. Following previous studies such as Owens (2013), we try to look at whether arrests increased or decreased as the addition of more police officers reduced crime, as a way to understand whether the relative size of the incapacitation effect from hiring more police outweighs or is outweighed by whatever deterrence and other crime-prevention effects comes from adding police officers. Unfortunately the arrest data contain even more noise and measurement error than do our crime data, with useful data on arrests being available for only about half as many agency-year observations as what we have for crimes themselves. We have not been able to produce results that are sufficiently "clean" as of this draft, in the sense of showing essentially zero "effects" in placebo regressions: that is, the effect of being above the COPS funding threshold in

2009 on outcomes measured *prior* to 2009. For that reason we do not report those results here but have them on the agenda as a high priority to finalize in the near future.

II. PREVIOUS STUDIES

A large literature in criminology dating back decades has found little systematic relationship between police spending and crime using cross-sectional data or repeated cross-sections. For example, Cameron's (1988) widely-cited review noted that 18 of the 22 studies examined found a null or positive relationship between police spending and crime. The National Academy of Sciences convened a blue-ribbon commission to examine the evidence on policing and wound up agnostic about whether police spending has any systematic relationship with crime (Skogan and Frydl, 2004). Franklin Zimring's (2011) account of the causes of New York City's miraculous crime drop since 1990 gave little credit to the one-third increase in the size of the police force that occurred over the 1990s, arguing that the research on police spending is "mixed."[6]

The obvious identification challenge to this line of research comes from the simultaneity that arises when policymakers respond to high levels or positive changes in crime rates by devoting more spending to police. The influential paper about this simultaneity problem by Franklin Fisher and Daniel Nagin as part of the National Academy of Science's 1978 review of deterrence research pushed most economists out of this business for decades.

A major advance in research on this topic came with Levitt (1997), who exploited variation in police spending due to the timing of electoral cycles for identification and

[6] For additional reviews see Eck and Maguire (2000), and Sherman (2011). Other studies have used time-series variation for identification. Marvell and Moody (1996) used a Granger-causality approach with panel data, while Corman and Mocan (2000) use monthly data from New York City, exploiting the fact that there is some delay between changes in crime and when the NYPD can have new officers hired, trained and deployed, to find a significant negative relationship between police and burglary (but not other crimes).

found a more pronounced negative relationship than previous studies. Levitt's estimated elasticity of violent crime with respect to police was around -1.0, and equaled -0.3 for property crime. While Levitt's estimates turned out to be less precisely estimated than originally reported (McCrary, 2002), and hence statistically insignificant, this seminal paper set off a search for other "natural experiments" to identify the effect of police on crime. Levitt (2002) himself utilized the number of firefighters as an instrument, with strong results. In this same spirit, DeAngelo and Hansen (2014) investigated the effects of an abrupt 36% reduction in the Oregon State Police capacity engendered by a voter referendum on taxation; the number of traffic citations issues fell in proportion, and highway traffic fatalities increased 12-14%.

Several subsequent studies have used changes in police deployment patterns following terrorist attacks to try to identify the relationship between police resources and crime. Di Tella and Schargrodsky (2004) use data from Buenos Aires, Argentina before and after a terrorist attack against a Jewish center that subsequently led to stepped-up police presence at Jewish and Muslim religious centers throughout the country. They find an elasticity of car thefts with respect to police on blocks where such religious centers were located of -0.33, with no such relationship found on nearby blocks. (They did not examine other crimes). Klick and Tabarrok (2005) used police redeployments in Washington, DC after terror alerts, but did not have data on local-area police presence and so cannot compute an elasticity estimate. More recently Draca, Machin and Witt (2011) use data from stepped-up policing in central London after two waves of bombings or attempted bombings, to estimate a crime-police elasticity of -0.3 to -0.4, with generally similar elasticities for violent crimes and property offenses.

A key challenge for this type of study design is the possibility of correlated shocks – in addition to the change in police presence after a terrorist attack, it is reasonable to believe that there will be changes in such crime-related variables as commercial activity, pedestrian presence and vigilance, or the use of private security guards. Draca et al. try to overcome this problem by taking advantage of a second source of variation that comes from the termination of the stepped-up policing program, which may be less likely to coincide with other changes in local criminogenic conditions. How the Draca et al. results would translate to the US context remains unclear. Moreover their paper does not generate a specific estimate for the elasticity of murder with respect to police, which as Chalfin and McCrary (forthcoming) note is particularly important in the US context given that a majority of the social costs of crime are accounted for by murder. This is particularly true for America's cities, where questions about how to improve public safety are most pressing.

A very different approach is taken by Evans and Owens (2007), who used variation across jurisdictions in the size of police-officer hiring grants awarded through the U.S. Department of Justice's COPS program under the Clinton Administration.[7] They show that the size of these grant awards are correlated with the prior levels but not trends in crime in recipient jurisdictions, and present difference-in-difference estimates that compare trends in crime before and after the hiring grants in areas that received larger versus smaller grants. Their preferred point estimates of elasticities for crime with respect to policing levels are -0.26 for property crimes and -0.99 for violent crime. In a follow-up study Owens (2013) shows that arrests as well as crime rates declined in areas with

[7] A related analysis using data from the Street Crime Initiative in England and Wales found that the police districts that received funding (10 of 43) enjoyed a reduction in robbery that was large enough to pass the cost-benefit test (Machin and Marie 2011)

relatively larger COPS hiring grants, suggesting that deterrence rather than incapacitation was the key mechanism through which police reduce crime. This is an important finding that suggests putting more resources into policing can help reduce crime without exacerbating the problem of "mass incarceration."[8]

With the publication of a new study by Chalfin and McCrary (forthcoming), the literature has in effect come full circle by returning to an analysis of naturally-occurring variation in police spending rather than focusing on just part of the variation that can be explained by some natural experiment. The authors argue that the null or perversely signed results from regressing crime on police result from the fact that statistics on the number of police employed by a city are highly error-prone. They compare two measures of the number of sworn officers, the Uniform Crime Report (UCR) measure (based on reports sent to the FBI by local police departments) and the Annual Survey of Government (ASG) conducted by the US Bureau of the Census and based on reports on city employment submitted by a city official outside of the police department. Chalfin and McCrary analyze the two series for 242 large cities for the period 1960-2000; while they are highly correlated in levels, the correlation in year-to-year percentage changes is just 0.22. They note that if the measurement errors in these two series are independent, that one can serve as an instrument for the other in a classic errors-in-variables analysis. Once they "solve" the errors-in-variables problem, they generate findings suggesting that growth in police manpower reduces both violent crime rates (elasticity = -.34) and property crime rates (elasticity = -0.17), both highly significant. While they admit that there may indeed be some simultaneity bias, the potential mechanism is not so obvious:

[8] Relatedly, Cook and MacDonald (2011) find that investments in private security following the creation of a business improvement district in Los Angeles has the effect of reducing crime and also reducing arrests.

in the regression they lag the independent variable (growth in police) and include state-year fixed effects.

The Chalfin-McCrary approach has the clear advantage of taking advantage of the full range of natural variation, rather than the more idiosyncratic and limited variation from a natural experiment (or true experiment). Their estimates tend to be smaller than those reported by Evans and Owens (2007), but are much more precisely estimated. If the only source of bias is simultaneity – that cities tend to hire more police in reliable anticipation of an increase in crime – then we can sign the bias (negative) and treat the results as a lower bound on the true prevention effects.

But in fact there is a more far-reaching problem with natural-variation data, namely that if we do not know the causes of the variation, it is hard to rule out correlated processes that may have either positive or negative effects on crime. For example, the growth in police may be the result in growth in city revenues, which may also be used to finance expansions of other programs that affect crime rates (such as regulation of alcohol outlets, or truancy-prevention programs). We conclude that while Chalfin and McCrary have made an important contribution, there remains great value from studying large-scale natural experiments that involve known processes that affect police resources.

III. COPS HIRING GRANTS UNDER THE OBAMA ADMINISTRATION

A. COPS Hiring Recovery Program Background

On February 13, 2009 under the American Recovery and Reinvestment Act of 2009, P.L.111-5, the Department of Justice Community Oriented Policing Services (COPS) Office was awarded $1 billion in funding for the COPS Hiring Recovery Program (CHRP) with the goal of creating roughly 5,000 law enforcement positions

funded for a period of three years at the state, local, and tribal levels.[9] As the fiscal climate continued to negatively affect law enforcement agencies (LEAs) across the nation, the CHRP was introduced to alleviate some of that pressure being felt by local policing systems: unlike previous iterations of the COPS hiring grant programs that included mandates regarding local match (25 percent) and salary caps ($75,000)[10], grants provided to LEAs under CHRP would cover the full entry-level salary and fringe benefits of every officer hired for a total of three years. Officers can be newly hired, previously laid off, or scheduled to be laid off. By increasing the capacity of local law enforcement through the hiring of additional officers, the CHRP sought to "increase their community policing capacity and crime prevention efforts."[11]

B. Program Eligibility

Between March 16 and April 14, 2009, the period of time for which the CHRP application was open, the COPS Office received 7,272 applications ($8,322,695,721 and 39,346 officers). While the CHRP was open to any state, local or tribal law enforcement agency only those agencies with primary law enforcement authority (the "first responders" for their jurisdiction) and full arresting and police powers were eligible to apply.[12] If otherwise eligible, a LEA could request funding for one of three types of "career law enforcement officers"[13]: (1) new-hire, full-time sworn officers, (2) rehiring of

[9] American Recovery and Reinvestment Act of 2009, H.R. 1—16
[10] http://www.justice.gov/opa/pr/Pre_96/October94/590.txt.html, "Grant Funding"
[11] http://www.cops.usdoj.gov/default.asp?Item=2267
[12] Start-up LEAs, consortiums of LEAs, and non-Federally Recognized Tribes' LEAs were also ineligible to apply
[13] As defined in the COPS FY2009 Application Guide: COPS Hiring Recovery Program, a career law enforcement officer is "a person hired on a permanent basis who is authorized by the law, or by a state, local, or tribal agency, to engage in or oversee the prevention, detection, or investigation of violations of criminal laws."

laid-off sworn officers (as a result of budgetary restrictions)[14], and (3) "rehiring" of sworn officers slated to be laid-off in the future (as a result of budgetary restrictions).[15] If awarded, the CHRP grant provided all necessary funding for officers for a total of three years at entry-level salary and benefits. [16]

While there was no cap on the number of officers a LEA could apply for funding from the CHRP, CHRP grantees that could not commit to continue funding CHRP grant-funded officers for a period of 12 months following the completion of the 36-month grant were ineligible to apply. Additionally, the CHRP non-supplanting requirement mandated any funds disbursed for the purposes of hiring additional officers must be used to supplement the LEA's budget for sworn officer.[17] Using CHRP funds to supplant law enforcement budgets was expressly prohibited; any additional officers hired as a result of a CHRP grant must be above and beyond the number of sworn officers that would have been employed should the grantee not have received COPS Office funding.

Other unallowable costs for which requests would not be funded included: requests for equipment, training, vehicles, existing locally-funded officers, salary and fringe benefits higher than an agency's entry-level salary and fringe benefits, civilian / non-sworn personnel, part-time officers, overtime costs, furloughed officers, and contractual arrangements for law enforcement services.

[14] The two main differences between the hiring of new entry-level officers and the re-hiring of previously laid-off officers was (1) the elimination of lag time between grant awards and officer deployment and (2) re-hires were generally not for entry-level positions

[15] Grantees were required to continue funding the positions slated for future lay-offs out of their local law enforcement budget until the date of the scheduled lay-offs.

[16] A LEA could hire officers at a higher rate than entry-level salary and fringe benefits, as was often the case with re-hires and preventions of future lay-offs, but any costs exceeding those levels must be covered by the LEA.

[17] http://www.cops.usdoj.gov/default.asp?Item=2282

Of the initial 7,272 applicants, 7,203 applications ($8,154,332,744 and 38,114 officers) contained eligible requests, as 58 agencies had been ineligible to apply and 11 agencies withdrew their funding requests (1 agency withdrew after final analysis, which is still included in these totals). The 7,202 eligible, competitive applicants is the sample of law enforcement agencies used for the purpose of this analysis.

C. Selection Criteria

1. Application Cleaning

In an effort to ensure accuracy in the CHRP grant distribution process, as the answers to the application questions were the sole basis of the ranking score, the COPS Office conducted extensive data cleaning to minimize errors in the application data.[18] Through data cleaning the applications, the COPS Office analyzed 276,336 values and utilized a variety of error-identification methods.

During the first round of data cleaning the COPS Office looked for internal contradictions in applicant responses. For example, if the number of sworn officers reported by a LEA was greater than the number of sworn officers reported in that LEA's budget the applicant was flagged for follow-up. Other methods utilized for identifying potential errors included unfitting 0's e.g. jurisdiction population, across all crime categories; large increases or decreases in budgets and/or personnel size; uncharacteristic poverty, unemployment, and/or foreclosure rates; and abnormally large crime numbers by population size (based on a predetermined formula.) The application data were also analyzed using standardized z-scores to help identify outlier values. As a result of these

[18] All LEA's whose information was flagged or changed were sent an e-mail notifying them of the change and instructing them to review the changes to their application online

data-cleaning processes, 1,709 agencies were contacted[19] to follow up on 2,864 potentially erroneous values.[20]

Through another round of cleaning, focused specifically on agency budgets and community policing narratives, an additional 912 applicants were contacted regarding 2,706 issues. The COPS Office also contacted agencies that may not have had primary law enforcement authority despite applying, such as Sherriff's Departments who solely oversee jails, natural resource police departments, State's Attorney's offices, etc., for verification of eligibility. In total, the COPS Office contacted over 2,600 agencies to correct and/or verify over 5,500 potentially erroneous data entries.

2. Application Questions, Weighting, and Scoring

To apply for CHRP funds, LEAs were required to respond to a series of questions regarding their locality. Basic information regarding the jurisdiction, its law enforcement capacity, and the population it served was collected. In the demonstration of need for federal assistance, each agency was also asked a series of questions related to its fiscal health, crime rates, and community policing efforts.

As funding for the CHRP was directed to the COPS Office through the American Recovery and Reinvestment Act, with the goal of creating and preserving jobs in those jurisdictions most impacted by the recession, significant weight was given to fiscal health questions, accounting for 50% of the overall ranking. Within the fiscal health section, questions were further broken down in to agency fiscal health (18% of overall), local government fiscal health (13% of overall), and community socio-economic health (19%

[19] The COPS Office would make four attempts (two e-mails, two phone calls) to get in contact with a given applicant. Failure to respond to data cleaning correspondence resulted in a ranking of last for the question in dispute, and was documented accordingly by CHRP administrators.

[20] A second round of cleaning using the same data cleaning criteria revealed an additional 236 issues to be addressed.

of overall). In line with the COPS Office mission of promoting public safety through community policing, crime rate questions accounted for 35% of the overall ranking and community policing questions accounted for 15%.

Using a 100-point scale and the relative weighting of the various sections as outlined above, each question in the application was also pre-assigned a maximum point value using the CHRP scoring rubric; the maximum total points a completed application could score was 100.

It is important to note that for the relative scoring and subsequent assignment of points to each question, agencies were first divided in to large (populations > 150,000) and small (populations < 150,000) groupings as determined by the size of population they served. Within these groupings by size of the locality's population, applicants' responses were compared and each question was ranked accordingly from most in need to least in need. The applicant demonstrating the least need, such as the applicant with the lowest homicide rate or the highest level of fiscal health, was assigned a ranking of 1; the applicant demonstrating the most need was assigned the maximal ranking for that question.[21] After each agency received a ranking for each question, the rankings were standardized using the maximum rank for that question to calculate the applicant's point value, or score, for that question.[22]

---

[21] Because of the possibility of ties e.g. identical homicide rates, the maximum ranking varied from question to question. For large agencies, the maximum ranking did not exceed 350, the total number of large agencies that applied. For small agencies, the maximum ranking did not exceed 6,853, the total number of small agencies that applied.

[22] For example if a LEA received a ranking X_i out of N_i agencies for a given question where N is equal to the total number of LEAs for that agency's size grouping i (large vs. small), the highest ranking received by any agency for that question is equal to $n_{max} \leq N_i$ (because of the possibility of ranking ties), and p_{max} represents the maximum number of points an applicant can receive for that question, then the point score P an agency received for that question is equal to:

$$P = \frac{X}{n\max} * p\max$$

For each question, applicants were first ranked based on their response relative to the other applicants' responses to the same question. For example, each applicant was asked to report their LEA's total operating budget for the current (2009) and previous two fiscal years, which was then used to calculate the average percent change in the LEA's budget across the three-year budget cycle. The agency that demonstrated the worst fiscal health and thus greatest fiscal need, specifically the LEA with the greatest average percent decrease in their total operating budget over the three-year period, would be assigned the highest rank for that question. After all responses for each question had been ranked, each applicant was then assigned a point-value for each question that corresponded to their relative ranking on that question.

After each question was assigned a score, the individual question scores were combined to create composite ranking scores for each LEA for each of the three ranking categories: fiscal health, crime rates, and community policing. The three composite scores were then summed to determine the overall ranking score for each applicant. The maximum overall ranking score for a small agency was 72.90. The maximum overall ranking score for a large agency was 75.77.

In the final allocation of CHRP funds, grantees were reordered based on their final ranked score and funds were awarded to the LEAs with the highest rankings until no funds remained. Because of the need to compare ranking scores across large and small agencies when distributing CHRP funds to satisfy the state minimum requirement (as explained in the following section), final ranking scores were calculated by taking the overall ranking scores in each applicant size grouping and standardizing them to a 100 point scale using the maximum ranked score from the respective grouping.

A total of 1,046 LEAs were awarded grants through the FY2009 CHRP resulting in the disbursement of $999,955,434 for the purposes of hiring and/or retraining 4,699 sworn law enforcement officers. Agencies were alerted of their awards in July 2009, with the official award start date as July 1st, and the public announcement coming July 28th. Agencies had to send a form to the COPS office signed by top government executives and law enforcement agencies with financial and programmatic authority in the jurisdiction within 90 days of receiving the award. Once the COPS office receives the signed documents, funds are available, although most rehiring and hiring began with the official grant announcement.

Since the FBI's Uniform Crime Reporting (UCR) system measures police force strength as of the end of October, any hiring by COPS grantees that occurred between July and October of 2009 could logically lead to a gain in police-force size that year in our analysis. That is, for the purposes of documenting the effects of the COPS program on the number of police in a jurisdiction, 2009 could be considered a "post-treatment year." We obtained quarterly grant progress reports and found that 47% of all agencies with 2009 awards hired all officers awarded by 9/30/2009, with 74% of agencies hired all officers award by the end of 2009, meaning many may have done so by the end of October. However because UCR crime rates capture outcomes measured over the course of the entire calendar, rather than a single snapshot as of the end of October, it is less clear whether we should count 2009 as a "post-treatment" year from the perspective of measuring impacts on our main crime outcome. Even jurisdictions that used their COPS grants to hire very quickly would have had a majority of calendar year under their "old" (pre-COPS) police staffing levels.

3. CHRP State Minimum and Population Size Requirements

As a stipulation of funding, the COPS Office was required to disburse at least half of 1% of the total available funds to each state or territory from which eligible applicants submitted proposals, or $5 million out of the available $1 billion.

To satisfy the state minimum requirement, the highest ranked agency or agencies from each state were identified first and awarded up to $5 million based on their final, standardized ranking score within their state. Only after each state minimum requirement had been met were the remaining funds allocated.

Another stipulation of CHRP funds was the distribution of half of all funds to agencies serving populations greater than 150,000 ("large agency") and half to agencies serving populations less than 150,000 ("small agency"). As the state minimum requirement must be met before further disbursement of funds, agencies in 27 states were awarded funds ahead of other higher scoring applicants from other states. In situations where agencies were funded ahead of higher scoring applicants from other states as a result of the state minimum requirement, funding across agencies in that state was capped at $5 million.[23] Generally speaking, the lower-ranked grantees were smaller LEAs in states with little policing capacity as-is. After allocating funds to satisfy the state minimum requirement, agencies were re-separated into groupings of large and small agencies and remaining funds were awarded to the highest ranking LEAs in each set, up to $500,000,000 total (or half of all available CHRP funds) across large agencies and $500,000,000 total across small agencies.

[23] In some cases, the last agency awarded funding in a "capped" state was awarded less than their requested amount to ensure funds were given to the agencies with the highest ranking scores once other distribution requirements had been met.

4. Award Capping

With no prior information on the number of officers a LEA might seek funding to hire/retain, the CHRP did not cap the number of officers an applicant could request. However to ensure the varying needs of law enforcement groups across the country were met, the COPS Office in consultation with the Attorney General's Office established a capping system that imposed a ceiling on the potential size of a given award. It was believed that without a cap on the number of officers a LEA could receive, a few agencies would have received nearly all of the available CHRP funds.

After all eligible applicants had received an initial ranking, a variety of cap options were applied to highlight the benefits and costs of utilizing different capping scenarios. Of the various cap alternatives, the most restrictive was chosen to maximize the distribution of funds across agencies: no agency would receive more than 5% of their current total force strength, up to a maximum of 50 officers.[24] Regardless of force strength size, every eligible agency qualified for at least one officer. Of the 85 applicants with more than 1,000 sworn officers, 24 requests were affected by the 50 officer cap.

IV. DATA

In this section we describe our different data sources on crime and police-force strength, which cover the years 2005-2012. While it has long been known that UCR crime data contain considerable measurement error, Chalfin and McCrary (2015) show that the same is true for UCR data on police-force strength as well. We describe below the different data cleaning procedures we apply before we use the data for estimation,

[24] For the purposes of determining total force strength, part-time officers were counted as one-half of a full time officer in the summation.

which includes a number of sensitivity tests to determine if our main findings hinge on how we handle data that are missing or potentially in error.

A. Sworn Officer Force Data

We follow Chalfin and McCrary (2015) and use two different data sources to create a measure the number of police officers in a given jurisdiction in a given year – the UCR together with the Annual Survey of Government Employee and Payroll (ASG). The two data sources are not entirely comparable – for example the UCR winds up covering more agencies and measures police as of October 31, while the ASG is a sample and measures employment as of end of March. Nevertheless as Chalfin and McCrary show there is useful "signal" that can be combined from both datasets.

Since the UCR has much greater coverage among our applicant cities than does the ASG, we rely on the UCR as our default data source. UCR captures data from all agencies that choose to voluntarily submit in a given year. In contrast the ASG sampling frame captures a census of all agencies in years ending in '2' and '7', and a sample in other years that has a sampling probability that is positively related to jurisdiction size (population). A new sample is selected every five years (ending in '4' and '9').[25] In the data we analyze here we find that among agencies that applied for COPS grants in 2009, 47% of the 850 agencies that served populations of 50,000 or more had ASG data for all analysis years, while 38% had ASG data for 5-7 analysis years. In contrast, among agencies serving populations of 10,000-50,000 (2,349 agencies in our sample), 16% had ASG data for all analysis years, while 33% had 5-7. Those numbers drop sharply for agencies serving populations under 10,000. In comparison, the UCR has sworn force data

[25] http://www.census.gov/govs/apes/about_the_survey.html

for all analysis years for jurisdictions with populations of at least 50,000, and 88% of all agencies (across all size categories).

To examine the sensitivity of our estimates to how we clean and combine these two different measures of police-force size, we create different versions of our measure that vary in the degree to which they accept missing data as missing in the UCR versus imputing missing values using different approaches. Our most conservative approach is to use only the UCR sworn force data, otherwise set to missing if the UCR data are either not available or are flagged as an outlier by our data-cleaning procedures.[26] We then create a second somewhat less-conservative measure that imputes values that are missing in the UCR (or set to missing from our data-cleaning procedure) using data from the ASG when available. This dataset also checks for and corrects outliers.[27] Finally we create a third version of the dataset that uses UCR and ASG data, and when neither data source is available, we impute missing values with either a regression or a spline model whenever at least 4 years of valid UCR (or ASG) data are available for that jurisdiction.[28]

[26] Our data-cleaning procedure defines outlier data points in the middle of our study period (not end years) as those with concurrent year-over-year changes of opposite signs that are at least 25% or larger in absolute value. For example, if the Anchorage Police Department has 100 sworn officers in 2007, 70 in 2008, and 105 in 2009, the 2008 year will be flagged as an outlier under this algorithm. For the boundary years of our sample period we use a different procedure and define outliers as those with a change of at least 35% in absolute value from the adjacent year in the sample. For example, if the Anchorage Police Department has 110 officers in 2011 and 150 officers in 2012, the 2012 year will be flagged as an outlier. We only apply these outlier rules if agencies have 15 or more sworn officers for 4 of the 8 study years, because in smaller jurisdictions we are more likely to see large swings in the actual data that are true (not data errors). The outlier algorithm was applied to the UCR data, with outliers set to missing initially, so adding the ASG data increases the useable data sample size.

[27] For the Anchorage Police Department example described above, if ASG data are available, it is imputed, and checked for outliers. Any new outliers are removed. If ASG imputes 90 officers for 2008, that would be accepted, because the concurrent year-over-year change is less than 25%. If ASG imputes 155 officers for 2012, it would be flagged again and removed. In our models no ASG imputed data were flagged.

[28] Wherever possible we impute missing values using a spline model with a knot between 2008-9 to allow for non-linearities in crime trends within a jurisdiction over time. However if we only have data for one side of the knot (that is, if data are missing for 2005-2008 or for 2009-2012) then we fit a linear regression model. When 5-8 years of sworn force data are missing, we leave them set as missing. Any predicted values from our imputation procedures that are negative are set to 0 if the median number of sworn force officers for all years is under 20, and set to missing if the median count of sworn force officers for all years

In practice the mean values of the three versions of our dataset are quite similar, as seen in the last column of Table 1 that reports average number of officers per agency in 2007 and 2009. Because the means seem so similar across different versions of the dataset, for our main analysis we use the version that imputes as many missing values as possible to maximize the information used in our estimation. In the appendix we show the results using our alternative data-cleaning rules, which reduce sample size, and find that our estimates are qualitatively similar.

B. Crime Data

Our main outcome measures are FBI Part 1 UCR offenses.[29] We create nine different versions of our dataset that vary along two dimensions- how aggressive we are

---

is 20 or over. Predictions that are positive, but for small agencies, defined as when the average sworn force for non-missing years that aren't imputed is 15 or less, are accepted as is. When the average sworn force for non-missing years is 15 or greater, predictions are considered for outlier testing. For example if the Anchorage Police Department is missing only 2006 sworn force data after the ASG imputation, and has an average of 95 officers for the non-missing years that aren't imputed, the 2006 predicted value will be checked for flagging. This baseline average number of officers again exists because with lesser denominators it is probable the variation in officers is true change. Then the predicted value is compared to the nearest value forward, or in the case of middle years, the average of the year before and after, so in the case of Anchorage, the predicted number of sworn officers for 2006 is compared to the average of the number of sworn officers in 2005 and 2007. We then assess if the predicted values differ more than 70% from the nearest sworn force number. If the 2006 predicted value is 90 officers and 2005 is 90 officers and 2007 is 100 officers, then the predicted value is within that 70% range of the average number of officers, 95, and used in the final dataset. If the predicted value differs more than 70% from the nearest crime/average nearest crimes, we look at whether the predicted value also differs more than 50% from the average sworn force for that agency. If the 2006 predicted value was 200 officers, it would be more than 70% different than the 2005/2007 average value, and more than 50% different than the average sworn force value of 100. In this case, the 2006 year would be set to missing. For agencies with average sworn force above 15, but are still small, there is more variation in number of officers year-to-year, so any given year may be quite different from the average. Particularly as some agencies grow and shrink over time, using local context allows us to keep the values that are reasonable in the local context, but appear to be outliers compared to the average. As a result, we chose to use the nearest sworn force year, or average sworn force years, as the first point of comparison. When data are otherwise missing, the nearest year forward is taken for comparison. Additionally, we add a second outlier check, related to the average sworn force, because in some cases agencies are missing a number of years in a row so the nearest crime comparison is not a useful comparison, in which case we check the predicted value against the average. These rigorous checks allow us to increase our sample size, while ensuring data quality. 4 outlier data points were found as a result of the regression/spline imputation and removed.

[29] Part 1 crime offenses consist of manslaughter, forcible rape, robbery, aggravated assault (summing to violent crimes), as well as burglary, larceny, and motor vehicle theft (summing to property crimes). Our models also use the index violent crimes, property crimes, and total crimes.

in flagging observations as outliers or errors, and how aggressive we are in imputing missing values or outliers that we set to missing.

The UCR data collection program was designed for national-level crime trend analysis,[30] and has some well-known sources of measurement error when analyzed at the level of the individual law enforcement agency. Maltz and Weiss (2006) describe some of the problems, including voluntary reporting that leads to substantial non-response, misreporting, and gaps in the data. Maltz generously shared with us his version of the UCR data through 2012.

Our most conservative approach to identifying errors in the UCR errs on the side of accepting UCR data as reported as correct. We then create additional measures that lower our tolerance for what we consider an outlier or error in the UCR data. The other dimension along which our measures vary is with respect to how aggressive we are in imputing observations that are missing (or that we have set to missing because we believe they are outliers).

Because of the measurement-error issues that are known to exist with UCR data, and based on our extensive graphing and diagnostic testing of these data, we subjected the data to extensive cleaning to detect errors. First, we identified agencies that had either missing crime data or reported 0's for all crimes for one or more years over our eight year study period, under the hypothesis that every jurisdiction would experience at least one Part 1 offense over the course of any given calendar year. Our second step was to identify values that were likely to be outliers. Our primary approach was to look for unusually large year-to-year fluctuations for that type of crime within a given jurisdiction. For "boundary" years (the years at the edge of our sample period), we look for unusually

[30] https://www.fbi.gov/about-us/cjis/ucr

large swings in value (in proportional terms) from the most adjacent year within our sample. For "internal" years in our sample we look for unusually large proportional declines followed by unusually large proportional increases the following year (or vice versa). We use three different methods that vary the thresholds for the size of the proportional changes that would lead to an observation to be flagged as an outlier.[31]

Once we have identified errors and outliers, there becomes the question of what to do with these observations. One approach we employ is very conservative and just leaves any values that have been flagged as outliers or errors as missing. We also try an alternative approach that takes advantage of the fact that most states publish on-line some

[31] The following are rules for all levels of outlier flagging, with different levels noted for conservative, baseline, and liberal methods. All numbers will be described in terms of conservative methods, with baseline and liberal noted in parentheses since all methods use the same rules, with different levels. Data are considered for being flagged as outliers and set to missing or imputed based upon each individual agency's crime type specific trend. Each of the seven part I crime types trends are considered separately, and as a trend of 8 years of data. For each crime type trend, if the crime counts are less than 30 for 4 of the 8 years, the trend will not be considered for cleaning and will be accepted as is (crime count <20 for 5 of 8 years for baseline and crime count <15 for 6 of 8 years for liberal). We did this because with low crime counts there are too few crimes to adequately tell apart whether any given pattern is driven by error or by true idiosyncrasy. If agencies have some years where all crime types are 0, this will initially be reflected as through each crime type as 0s and small data, but are accounted for later, which will be described. For example, if the Anchorage Police Department has crime counts that are more than 30 for 4 of 8 years for larceny, all years of larceny are eligible for flagging. However, if there are less than 30 murders for 4 of 8 years, all years of murder are accepted as is. For all our crime data, about 30.0% was eligible for flagging by the conservative method (37.7% for baseline and 43.3% for liberal). We then take all eligible data and check, by crime type trend, whether any of the three following rules occur, and if so flag the specific data point: 1. For interior years 2006-2011: Is change in crime rate the year prior, and year after are opposite directions and more than 70% different (60% for baseline, 50% for liberal)? If so, flag. If Anchorage has 50 larceny crimes in 2007, 150 in 2008 and 60 in 2009, then the 2008 value is flagged. This rule captures one year spikes and dips that are often seen in crime trends, but could be the result of not reporting for one year, entry errors, etc. 2. For edge years 2005 and 2012: For 2005, if 2006 is flagged, take the change between 2007 and 2005 (imagine that 2006 is a spike down or up and 2005 is close to level with 2007), if the change is greater than 60%, flag. Do the same for 2012, except if 2011 is flagged, take the change between 2010 and 2012 (50% for baseline and 40% for liberal). Because the edge years only have 1 nearby comparison year, we require the amount of change to be higher, and then account for when the closest year to the edge is flagged. If Anchorage has 40 motor vehicle thefts in 2005, 90 in 2006, and 50 in 2007, the 2005 number of motor vehicle thefts would be compared to the 2007 ones, and since it is within 60% it would be accepted as is. 3. Within an agency, within a year, across crime types are there 4 (3 for baseline and liberal) or more dips or spikes by 40-70% each (30-60% for baseline and 30-50% for liberal)? Crime types must be eligible to be analyzed in this rule. This is accounting for some entry errors we saw where all crimes dropped by similar amounts. We chose the baseline amount of change 40%, and then bound it at 70% where the data are flagged as outliers regardless. If larceny, motor vehicle theft, burglary, and aggravated assaults all dropped between 40-70% in 2011, these 4 data points would be flagged. If data points are already flagged before rule 3, they can still be considered in the analysis.

agency-level crime data, which sometimes differs from what was reported to the UCR. So we use these state data when available to impute placeholder values for observation values that are missing or we believe are in error, then re-apply our procedures described above for checking for outliers to see if these state values that we have imputed seem consistent with the rest of the data series or instead appear to be outliers or errors. Our final approach complements this state-data imputation procedure by doing further imputation when state data are not available using observations for that agency in the UCR data from other years. Specifically if we have a missing observation or one we believe is in error and cannot impute with state data, if that jurisdiction and crime type has at least four years of non-missing data within our study period we fit either a regression or spline model to data from the other years and impute a value. We apply some simple rules to eliminate predicted values that seem implausible.[32]

Table 2 compares the mean values and sample sizes of the nine versions of our dataset. The crime columns are average crime rates per 100,000. Particularly, the second and third rows (our two imputation strategies) are successfully preserving the number of observations across the columns as we move from a strategy that errs on the side of not flagging suspicious observations to a strategy that errs on the side of having a low bar to flag an observation, and also yields fairly similar overall crime counts as we move across columns. In the appendix we show the results using our alternative data-cleaning rules, which reduce sample size, are qualitatively similar.

[32] Our spline and linear regression model is very similar to the one described for sworn force cleaning, with a few differences. For each agency, each crime type trend is analyzed separately, as a stand-alone 8 years. This means for any splines or linear regressions to be fit, the crime type trend must have at least 4 of the 8 years' worth of data. Additionally instead of an average sworn force 15 or greater needed for a prediction to be considered for all non-missing years, that value is 20, and again, for each agency, for each specific crime type. Otherwise the percentages used are the exact same regardless of conservative, baseline, or liberal. The purpose of the spline/linear regression is to increase the sample, while the means remain the same.

C. Arrest Data

For purposes of understanding the mechanisms of action through which additional police may reduce crime, we also examine the effects of the COPS grant program on the number of arrests made within a jurisdiction. As Owens (2009) notes, a decline in arrests that accompanies a decline in crime provides suggestive evidence that more police reduce crime through some combination of general deterrence and whatever other activities police take to prevent crime from happening in the first place. This issue is also important for benefit-cost analysis, to determine whether the costs of additional incarceration accompany the costs of additional police.

We use FBI UCR on arrests for part 1 crimes as our primary measure of arrest activity within a jurisdiction. As with our approach to cleaning the crime data, we create nine different measures of our arrest variable that vary along the same two dimensions- how aggressive we are in flagging observations as outliers or errors, and how aggressive we are in imputing missing values or outliers that we set to missing.

The UCR arrest data collection is considered more error prone than the crime data. While crimes that occur are generally reported and recorded, only some of those crimes will result in arrest (which lowers the levels and so makes outlier detection and non-reporting more complicated in small jurisdictions). Moreover arrests are not always recorded in the UCR. As Maltz (1999) notes, the UCR only reports arrest data when it is sent in for both juveniles and adults. Additionally, arrest rates are also affected by police department priorities.[33]

[33] Arrest data problems are considered greater than crime because the percent of arrests reported by police is lower than % crimes reported

As with crime data we create different versions of arrest measures that vary in the degree to which they accept missing data as missing in the UCR, instead of using other methods of imputation. Our arrest datasets use most of the same rules for detecting missing or outlier observations, with some slight differences.[34]

D. Population Data

The final building block we need to construct our dataset is the number of people served in each agency, which is the denominator required to calculate any measure that is recorded on a per capita basis (crime, arrests, police, etc.) The UCR in principle captures population for each agency each year when they report to the UCR, although the accuracy of these data is less than perfect. Moreover much of the 2005 and 2009 UCR data on local-area population seems to be in error.[35] As a result we use cleaned UCR population data generously provided to us by Michael Maltz, which he obtained from the National Archive of Criminal Justice (NACJD) and subject to some additional data-

[34] All arrest outlier flagging methods consider observations for each agency, and for each specific arrest type, meaning each arrest trend has 8 years of data. We only consider observations for outlier testing if an agency has more than 15 arrests for 6 out of 8 of the years, regardless of outlier testing method; 25% of data are eligible for flagging. The most conservative outlier testing method has three parameters, which are the same as the three crime rules, with different percentages. These three parameters are the exact same for the baseline outlier testing method, which has the addition of a fourth parameter; while the liberal outlier testing method has all four parameters, with different levels. The three parameters that persist in all outlier testing methods are the same as the crime data flagging rules, with different percentages: 1. For interior years 2006-2011: Is change in crime rate the year prior, and year after are opposite directions and more than 60% different (50% for liberal)? 2. For edge years 2005 and 2012: For 2005, if 2006 is flagged, take the change between 2007 and 2005 (imagine that 2006 is a spike down or up and 2005 is close to level with 2007), if the change is greater than 50%, flag. Do the same for 2012, except if 2011 is flagged, take the change between 2010 and 2012 (40% for liberal). 3. Within an agency, within a year, across crime types are there 3 (same for liberal) or more dips or spikes by 30-60% each (30-50% for liberal)? The fourth parameter, which exists only in the baseline and liberal outlier testing methods is similar to the second parameter, but instead examines the end year change when the adjacent year is not flagged. Is the change between the adjacent year (2005-2006 or 2011-2012) greater than 70% for baseline (60% for liberal)? For example if Anchorage has 50 aggravated assault arrests in 2005, 140 in 2006, and 160 in 2007, the 2005 year is flagged.

[35] All UCR data was obtained directly from the FBI. When examining year-over-year population change most agencies exhibited changes from 2008-to-2009 at a magnitude over 200, with 2009-to-2010 showing drops of a similar magnitude. [35] In 2005 the mean population in the data from Maltz is 25% larger than in the FBI UCR data. But when we look at year-over-year changes we see a big "jump" in populations from 2005 to 2006 in the FBI data, and much more agreement between the Maltz and UCR population data starting in 2006 (with the exception of 2009).

cleaning procedures.[36] His data align with the UCR when comparing the raw differences[37], have fewer missing values, and has 2009 populations with year-over-year trends similar to other years in basic diagnostic testing.

We also explore the sensitivity of our analysis to imputing population values that are missing in Maltz's version of the UCR dataset by using population data from the American Community Survey (ACS) 5-year sample populations from 2009-2012[38]. The result is we can examine the sensitivity of our results to use of two different measures of jurisdiction population. Table 4 shows the sample sizes and population means for a few years. The addition of ACS data increases the sample size and produces qualitatively similar estimates.

V. ESTIMATION APPROACH

Our study provides new evidence on the effects of police spending on crime by exploiting a natural experiment generated by the COPS grants given out by the Obama administration. In what follows here we focus on the initial funding competition that was carried out by the administration in 2009, which disbursed a total of $1 billion. This was followed by smaller grant competitions that gave out $300 million in 2010, close to $200 million in 2011 and a little over $100 million in 2012. In future work we will incorporate results from those subsequent COPS funding rounds as well.

As discussed in Section III, agencies were ranked relative to their population-size group based on an algorithm that weighted together different application factors such as financial distress in the jurisdiction and local-area crime rates. After state funding was

[36] Correspondences with Michael Maltz show that filled in missing population data when possible.

[38] We chose ACS over Census estimates because Census estimates are linear and ACS data are based on a sample taken over 5 years.

disbursed (each state gets $5 million guaranteed), the remaining funds were allocated to agencies in descending order of their COPS grant algorithmic ranking. Funds were distributed until both small (population under 150,000) and large (population 150,000 or more) agency groups each received $500 million.

Because of the state minimum funding requirements and the separate grant competitions for large and small agencies, together with a few other idiosyncrasies of grant competition process, our setup is slightly more complicated than a standard regression discontinuity (RD) design. Through extensive conversations with the COPS office administrators, we were able to perfectly reconstruct the process through which agencies received funding.

The key to our research design is that the COPS office allocated funding based on each agency's rank-order as established by the application process, and funded agencies in descending order until the funding ran out. This leads to a situation in which the last agency funded will have a very similar application-algorithm score to the agency that is at the top of the "not funded" list, but this small difference in application score will lead to a large difference in COPS funding.

We use agencies' final ranked scores to recreate the funding cutoff to identify the effects of the COPS grants on agencies close to the cutoff using an RD design. Because state funds were disbursed first, we calculated the percent of total funding for each state, and marked all states with less than .55% of total funds. Within each ½ percent state (*i*) we found the maximum unfunded score (X_i) and the minimum funded score (X^*_i), then calculated a state cutoff as: $(((X_i - X^*_i)/2) + X_i)$. There is one cutoff for each ½ percent state. We then split agencies into small and large size categories based on their

application reported populations. For both small and large agencies there is a score beyond which all agencies are funded, and this is the general cutoff. We use the same formula as for state cutoffs, except it applies to the overall size category, not a specific state. Any agency over the small agency cutoff (65.75) or large agency cutoff (68.75) will be funded. Any agency below the size specific cutoff will not be funded, unless the agency is from a ½ percent state and above the state specific cutoff. We perfectly created the top down funding approach taken by COPS and accounted for the ½ percent state specific cutoffs.

We use the data described above to estimate equation (1) below, where Y_i is our dependent variable of interest, which could be police-force size, crime, or arrests per capita and T_i=1 indicates receipt of a COPS grant as part of the 2009 grant competition. We include a flexible control for a normalized version agencies' final ranked score, Z_i. We take each individual agency's score and subtract the relevant cutoff for that agency (½ percent specific if it is in a low-crime or low-population state where no agencies were above the COPS threshold, otherwise the actual COPS grant cutoff for that agency's jurisdiction-size group), then divide it by the standard deviation of all relative scores. This is also useful because there are COPS grants in later years which we can examine through this standardized lens.

$$(1) \qquad Y_i = \beta_0 + \beta_1 T_i + \beta_2 f(Z_i) + \varepsilon_i$$

Because of state minimum funding requirements, some agencies with application scores below the size-group funding threshold wind up getting COPS funding. For this reason we treat our as a "fuzzy" RD, where having a score above versus below the

funding threshold generates a discontinuity in the probability of receiving a COPS grant. Our estimate for β_1 in equation (1) is essentially an intention-to-treat (ITT) estimate.

Identification of the causal effects of winning a COPS grant—the ITT corresponding to a treatment of increased number of officers for an agency—comes from assuming smoothness in potential outcomes near the funding cutoff (Porter 2003). Intuitively, the challenge to identification is that agencies on either side of the funding threshold are not literally identical – they differ very slightly in their application score and hence in the underlying social conditions that help determine their application score. The assumption of smoothness in potential outcomes around the funding threshold means that we can use data on either side of the threshold (but "near" the threshold) to estimate the general relationship between small changes in the application score and outcomes. We then essentially net out the part of the difference in the outcome we see for agencies just above versus just below the threshold that is due to having slightly different application scores. The remaining difference is attributed to the effects of the treatment that is assigned on the basis of the funding threshold (in our case, COPS grants).

One estimation issue is how to control for the function $f(Z_i)$ in equation (1). Rather than imposing some functional form assumption, we use a nonparametric RD approach refined by Porter (2003). This model uses local linear regressions (Fan 1992) to produce estimated treatment impacts from the difference in the left and right side limits of discontinuity. We estimate this for a chosen bandwidth and use triangle kernel weights. This means we estimate a kernel-weighted linear regression once to the left of the COPS cutoff and once to the right of the cutoff, then difference the results for the treatment impact (in our case focusing on the means). Our use of triangle kernels is based on work

by, Cheng, Fan, and Marron (1997) which are shown to be optimal, if not very sensitive to choice of kernel (Fan and Gijbels 1996). Our standard errors are derived from Porters' work (2003).

VI. RESULTS

As noted above, we focus in this report on using the 2009 COPS funding competition as our "natural experiment" to estimate effects on crime. Because it is possible that there could be some lag in the effects of changing police resources (or practices) on crime, we examine the effects of winning a 2009 COPS award on crime in 2009 and in subsequent years for which we have crime data (2010-12). We begin by presenting estimates for how receipt of a COPS grant changes police spending, and then turn to estimated effects on crime.

A. Effects of COPS on police-force strength

In Figure 1, we present the results of being above versus below the COPS funding threshold on the likelihood of receiving an award. We show this figure simply to confirm that we have been able to reproduce the COPS funding process with our data and estimation model. Receipt of a COPS grant is on the y-axis in the figure and the running variable along the x-axis is our standardized 2009 COPS grant application score. The figure also shows the nonparametric estimate for the function relating the agency's normalized ranking score to the dependent variable, f(Z), as well as the implied discontinuity (that is, our estimate for β_1 from equation 1 above) estimated using a bandwidth of 0.8 standard deviations from the scoring cutoff. So that readers also have some sense for the raw data underlying these figures, the figures also include cell means

(shown in triangles in the figure) and their 95% confidence intervals (bars) from grouping the data into 10 evenly distributed bins on each side of the cutoff.

We see in Figure 1 that the probability of receiving an award jumps from essentially 0 (0.008) to almost 1 (0.968). The discontinuity does not go literally from 0 to 1 because of the way the COPS grant program was rolled out, as described in Section III above. This includes the effects of the state minimum funding rules, which gives funding to some agencies with scores below the cutoff ahead of agencies with higher scores from other (larger, higher-crime) states. Figure 2 reproduces this exercise but now using dollars of COPS funding awarded as the dependent variable shown on the y-axis. The mean funding amount for those below the cutoff is $1,921, and jumps to $620,165 at the funding threshold. Both estimates in Figures 1 and 2 are significant at the 1% cutoff. These graphs confirm that near the cutoff agencies that were very similar in their underlying social conditions (as captured by the COPS application score, which weighted together measures related to crime, financial distress and other factors) wound up receiving substantially different amounts of COPS funding. This is the key to our design.

In Figures 3 through 8 we present the results of estimating different "placebo" regressions, where being above the 2009 COPS funding threshold is the "treatment" and we measure police-force strength in the pre-treatment year of 2008.[39] (We also tried a similar exercise with data from the pre-treatment year of 2007 with qualitatively similar findings – available upon request). Figure 3 shows the results of using police-force size as a raw count, Figure 4 converts that to a rate per capita, Figures 5 and 6 show the

[39] Figure 3 estimated discontinuity=38.60 T-stat=0.94. Figure 4 estimated discontinuity=114.28 T-stat=1.19. Figure 5 estimated discontinuity=0.09 T-stat=0.74. Figure 6 estimated discontinuity=0.09 T-stat=1.73. Figure 7 estimated discontinuity=0.03 T-stat=2.04. Figure 8 estimated discontinuity=0.00 T-stat=1.23.

results of taking the logs of the outcomes used in Figures 3 and 4, and Figures 7 and 8 show the results of using the difference of the logs of the counts and rates per capita relative to the base years of 2005-6. Adjusting police counts to be per capita rates helps account for chance fluctuations in agency size on either side of the threshold. Logging these per capita rates helps further control for the influence of extreme values that were not flagged and corrected by our data-cleaning procedures, while taking differences from a pre-treatment baseline period (2005-6) essentially helps net out the influence of other time-invariant agency-level factors that may randomly vary near the threshold (and also help with any persistent idiosyncracies in how agencies report data).

On the basis of this type of placebo regression we choose the difference in log per capita rates in a given year relative to the pre-treatment baseline years of 2005-6 as our main estimating model. This specification tends to give us the cleanest set of "zero" impact estimates during the pre-treatment years of 2008 (and 2007) for our measures of police-force size, and (in results not shown) crime during the pre-treatment years as well. Selecting our model on the basis of fit to pre-treatment data, where we logically expect zero impacts on either the dosage (police) or response (crime), lets us pick a specification in a principled way that avoids risk of human data mining as can result when these decisions are made on the basis of seeing impacts in the post-treatment years.

In Figure 9, we show the results on police force strength (log officers per capita) as measured in the first "post-treatment" year of 2009. For the reasons mentioned above our measure is calculated as a difference between log officers per capita in 2009 from log officers per capita in 2005-6. Agencies were notified of their COPS application outcome in July 2009, which in principle gave them time to either hire new officers or prevent

layoffs before they reported their police-force strength to the UCR as of October 31, 2009. We see a "jump" in police per capita that is equal to 0.008.[40]

One question is the degree to which these COPS grant effects on police-force size persist over time. These effects may attenuate over time partly because local governments could in principle reduce their own spending commitments to police to enable them to essentially spread the new resources they receive across different government functions that local voters value. This is a standard prediction of public finance, although in practice many empirical studies have found a so-called "fly paper" effect where resources earmarked for a given use tend to disproportionately concentrate on that use. Another reason there could be some attenuation in the estimated effect on police resources over time is the subsequent COPS grant competitions carried out in 2010-12 may have provided additional resources to agencies that were in the 2009 "control group."

In Figure 10 we present the results of the estimated effect of being above the threshold on the log difference in police force size (y-axis) for different years. We see in 2007 and 2008 there is essentially no effect, as we would expect since these are "pre-treatment" (this just confirms the placebo test reported above). In 2009 we find an increase equal to 0.008 (about 1 percent) that is only slightly smaller in 2010 and 2012. (The estimated effect in 2011 is only about one-quarter as large but the confidence interval around this estimate is fairly wide).

The COPS program in 2009 wound up having the largest proportional effects on police resources in medium-sized jurisdictions. Figure 11, Panel A for instance shows that for small agencies serving populations under 50,000, we see a statistically significant gain in police resources equal to 0.7 percent – just slightly smaller than the full-sample

[40] Figure 9: Estimated discontinuity=.008 T-stat=2.15

average shown in Figure 10. We can also take advantage of the fact that agencies varied in the size of the police funding request they made and/or were eligible for, which is a baseline application variable and so one that we can construct for all agencies. We call this "potential impact" and is defined as the maximum percentage increase in police force size that an agency could expect from receiving a COPS grant. Specifically our potential impact variable is calculated as the maximum number of officers eligible to be awarded divided by the sworn force for 2009 reported in the application. Agencies could ask for any number of officers, as long as they could support the cost for 1 year after the 3-year grant concludes. The COPS office then imposed a cap of 5% for the size of the increase a department could experience as a result of receiving a COPS grant for agencies with up to 1,000 sworn officers. For larger agencies (24 were affected in total), the COPS office imposed an absolute cap of 50 officers maximum that could be supported by a COPS award. If agencies requested less than the COPS-imposed cap we code them as caped at the amount requested.

Panel B in Figure 11 shows that for small jurisdictions (under 50,000) with a potential impact score greater than 5%, the estimated change in police resources in 2009 as a result of receiving a COPS grant was 1.1 percent, about 50% larger than what we see for the full sample of small agencies. For small agencies with a potential impact score of 10% or greater the change in police resources in 2009 was fully 2.6 percent, nearly four times the small-agency overall average (panel C).

Figure 12 shows that the largest changes in police resources occurred among medium-sized agencies with populations between 50,000 and 150,000. Panel A shows

that among all agencies of this size the change in police spending in 2009 was 1.3 percent (compared to the 0.8 percent full-sample average).

In contrast for large cities with populations of 150,000 or more there is relatively little change in average police-force size as a result of winning a COPS grant, as perhaps might be expected given the absolute cap on the number of officers (Panel A of Figure 14). However if we look at large jurisdictions that had a larger "potential impact" from receiving a COPS grant, as suggested by baseline application data, we do see a statistically significant gain in police force strength for that sub-group of agencies.

So for starters we have documented that the 2009 COPS grant competition successfully increased the number of police on the streets, and that while these gains may attenuate a bit over time they do show some persistence through the end of our study period in 2012. The gains in police force strength were particularly large for sub-sets of the applicants that were either medium sized (populations 50,000 to 150,000) and/or were eligible for particularly large changes in police spending based on application data. The advantages of focusing on these sub-sets of agencies in our analysis of crime impacts will be larger underlying "first stage" changes in the police per capita "dose," traded off against a reduction in statistical power from using fewer observations.

B. Effects on Crime

Figures 15 and 16 show the estimated effects of being above the COPS threshold on the difference in the log of per capita total UCR part 1 crime rates in 2009 (Figure 15) and 2010 (figure 16) relative to the base years of 2005-6, for all applicants. [41] In 2009 the estimated effect on the log difference in total part 1 crimes is small and statistically

[41] Figure 15 estimated discontinuity= -.003 T-stat= -.441. Figure 16 estimated discontinuity= -.011 T-stat= -1.77.

insignificant. But by 2010 – the first full calendar year after the COPS grants were awarded – we see a decline in total part 1 crimes of 1.1 percent; given a standard error of 0.6 percent the estimate is not quite statistically significant at the usual threshold (the p-value is <.10 rather than <.05). But by 2011 the decline in total UCR part 1 offenses is equal to 1.8 percent, which is significant at the usual cutoff of p<.05 (Figure 17).

Figures 18 and 19 show the estimated effects year by year on the log difference of per capita violent and property crime rates, respectively. For part 1 violent crimes the estimated effect is statistically significant already in 2009, equal to a 2.4 percent decline, and basically persists through the end of our sample period in 2012 (getting as large as 3.1 percent in 2011). For part 1 property crimes the detectable decline in crimes occurs with a lag (starting only in 2010) and is somewhat smaller in magnitude, equal to 1.2 percent in 2010 and 2 percent in 2011. The fact that the estimated effect is larger for violent than property crimes is consistent with previous studies, and so reassuring in that sense. This pattern is also relevant for policy purposes since (as noted above) the social costs of crime are so disproportionately driven by violent crime.

Figures 20-28 report results separately for violent and property crimes for different sub-groups of jurisdictions defined by population group and our "potential impact" measure. These estimates have the disadvantage of substantially reducing the amount of data we are able to bring to bear to estimate impacts on our noisy outcome measure (crime rates). For the most part these estimates are imprecise and usually include zero within the confidence interval. An apparent exception to this pattern would initially seem to be Figure 21, which shows results for small agencies that were eligible for particularly large changes in police-force strength. While the post-COPS (2009 and later)

drops in crime are large and statistically significant, we see signs of these drops during the pre-COPS period as well – suggesting that with this small sample of agencies the identification assumptions of the RD model is somehow breaking down. So we do not put much weight on these sub-group estimates.

Finally, given the particular importance for the social costs of crime from murder, as shown by Chalfin and McCrary (forthcoming), it is worth examining whether the COPS-induced changes in police spending also have an effect on this crime. Figure 29 shows the 2009 estimated effects on the log difference of per capita murder crime rates.[42] These results are significant at the 10% cutoff.

C. Effects on Arrests

Our last set of results tries to learn a bit more about the potential mechanisms through which increased police spending reduce crime. As noted above, police may reduce crime by arresting and incapacitating offenders (either by catching them in the act or doing follow-up investigation work), deterring crime, or engaging in problem-solving or other preventive activities. We do not have direct data on any of these mechanisms but can look at what happens to arrest rates in these agencies. If arrests decreased alongside the reductions in crime that we have documented above, the implication would be that whatever deterrence and other prevention activities additional officers undertake outweigh whatever effect on crime comes from apprehending more offenders. On the other hand if the additional police resources reduce crime while increasing arrests, that would suggest that the apprehension / incapacitation effects dominate whatever deterrence and prevention activities occur. The distinction is important partly for

[42] Figure 29 estimated discontinuity= -.033 T-stat= -1.70.

strengthening our understanding of how police reduce crime, and partly because of policy reasons. If police reduce crime and arrests simultaneously the implication would be that increased police spending does not strain courts and correctional systems and actually alleviates (rather than exacerbates) concerns about America's incarceration rate.

We replicated our main estimates for crime and police-force strength now using arrests (in various forms) as the dependent variable, using being above the 2009 COPS funding threshold as the "treatment." We have not yet been able to identify a model specification with these arrest data that passes our specification check for valid inference, and so have this on the agenda for future research.

VII. DISCUSSION

This paper presents new estimates for the effects of increased police resources on crime in the US that, compared to most previous studies, has an unusually strong research design and the advantage of drawing on national data (rather than just from a case study of a single city). Specifically we exploit the fact that the US Department of Justice's COPS office rank-ordered applicants for funding in 2009 for police hiring grants based on a score that weights together local-area crime, financial distress and other factors, then essentially gave out grants until the funding was fully expended. This creates the opportunity for a regression discontinuity design.

We show that the COPS program increased police per capita by 0.8 percent for agencies near the funding threshold, which translates into a decline in overall UCR part 1 offenses of between 1 and 2 percent and a decline in violent crimes specifically of about 2 to 3 percent. These estimates on their face would seem to imply a very large elasticity of crime with respect to police - that is, percentage change in crime for a given

percentage change in police (%ΔC / %ΔP). However we would caution against using these point estimates at this point to calculate such an elasticity, partly because our estimates for both the change in crime and change in police have fairly wide confidence intervals. Because the elasticity is a ratio of two estimates that have wide uncertainty bands, the estimated elasticity will have even more uncertainty.

A key goal for our future work is to generate such an elasticity. This will first require trying to reduce the uncertainty in our estimates for the numerator and denominator of the elasticity. Partly this can be accomplished by taking the results from the 2009 COPS grant competition and aggregating results across the years for which we have outcome data. Folding in the results from subsequent (smaller) COPS funding competitions in 2010, 2011 and 2012 may help further. And to provide a sense for how much uncertainty there is in the resulting elasticity estimate we plan to carry out a boot-strapping exercise where we would draw random sub-samples of our data with replacement, estimate the numerator and denominator of the elasticity and calculate the elasticity value, then replicate this exercise thousands of times. The result would be an empirical distribution for the elasticity estimate that gives us some sense for how sensitive the estimate is to slight perturbations in the data.

It is nonetheless striking that the results taken at face value seem to imply that something less than a one percent increase in police per capita translates into more than a one percent decline in total crimes – particularly violent crimes. One speculative explanation for this surprisingly large responsiveness of crime with respect to marginal changes in police resources is suggested by a conversation one of our team had with a renowned police researcher several years ago, who noted that “99 percent of what every

police department does is driven by 911 calls." This is of course just an approximation (and perhaps a slight exaggeration), but it does suggest a deeper insight about the effects of incremental expansions of police resources. A one percent increase in police force strength would initially seem to be a very modest change in law enforcement capacity to control crime. But if the volume of 911 calls by citizens is held fixed when police resources increase, if anything like the 99% rule holds true then a one percent increase in police resources could potentially represent a very large proportional change in the flexible resources available to police to engage in pro-active or preventive activities. To say there may be increasing returns to police spending would probably be too strong, but the findings here in any case seem to support the view that additional resources for police represent a cost-effective approach for reducing crime.

Table 1 Number of Observations and Average Sworn Force per Agency

Imputation Method	# observations	Officers per agency 2007	2009
UCR data with outliers set to missing	54,103	74.03	77.26
UCR data with outliers imputed using ASG (where available)	56,040	69.15	76.95

Table 2 Number of Observations and Average Crime Rates per 100,000 for Different Imputation Approaches and Outlier Detection Methods

# Observations, Crime rate per 100,000 for Murder, Robbery, and Burglary												
Imputation Method	High threshold for flagging outliers	Murder	Robbery	Burglary	Medium threshold for flagging outliers	Murder	Robbery	Burglary	Lowest threshold for flagging outliers	Murder	Robbery	Burglary
Set to Missing	318,039	3	67	628	294,700	3	70	641	276,718	3	73	655
State Impute + Missing	346,526	3	67	24	345,842	3	67	625	343,113	3	66	627
Spline/Regression Impute + Missing	352,006	3	66	622	351,750	3	66	623	350,660	3	66	624

Table 3 Number of Observations and Average Arrest Rates per 100,000 for Different Imputation Approaches and Outlier Detection Methods

# Observations, Arrest rate per 100,000 for Murder, Robbery, and Burglary												
Imputation Method	High threshold for flagging outliers	Murder	Robbery	Burglary	Medium threshold for flagging outliers	Murder	Robbery	Burglary	Lowest threshold for flagging outliers	Murder	Robbery	Burglary
Set to Missing	231,687	3	27	104	230,270	3	27	104	228,591	3	26	104
State Impute + Missing	239,997	3	27	104	239,479	3	27	105	236,774	3	26	104
Spline/Regression Impute + Missing	251,210	3	26	105	250,558	3	26	105	249,482	3	26	104

Table 4 Number of Observations and Average Population per Agency

Imputation Method	# observations	Population per agency 2007	2009
Set to Missing (UCR only)	55, 181	31,689.92	32,066.08
ACS Impute + Missing	57,335	31,689.92	30,455.74

Figure 1: Discontinuity in receiving COPS grant

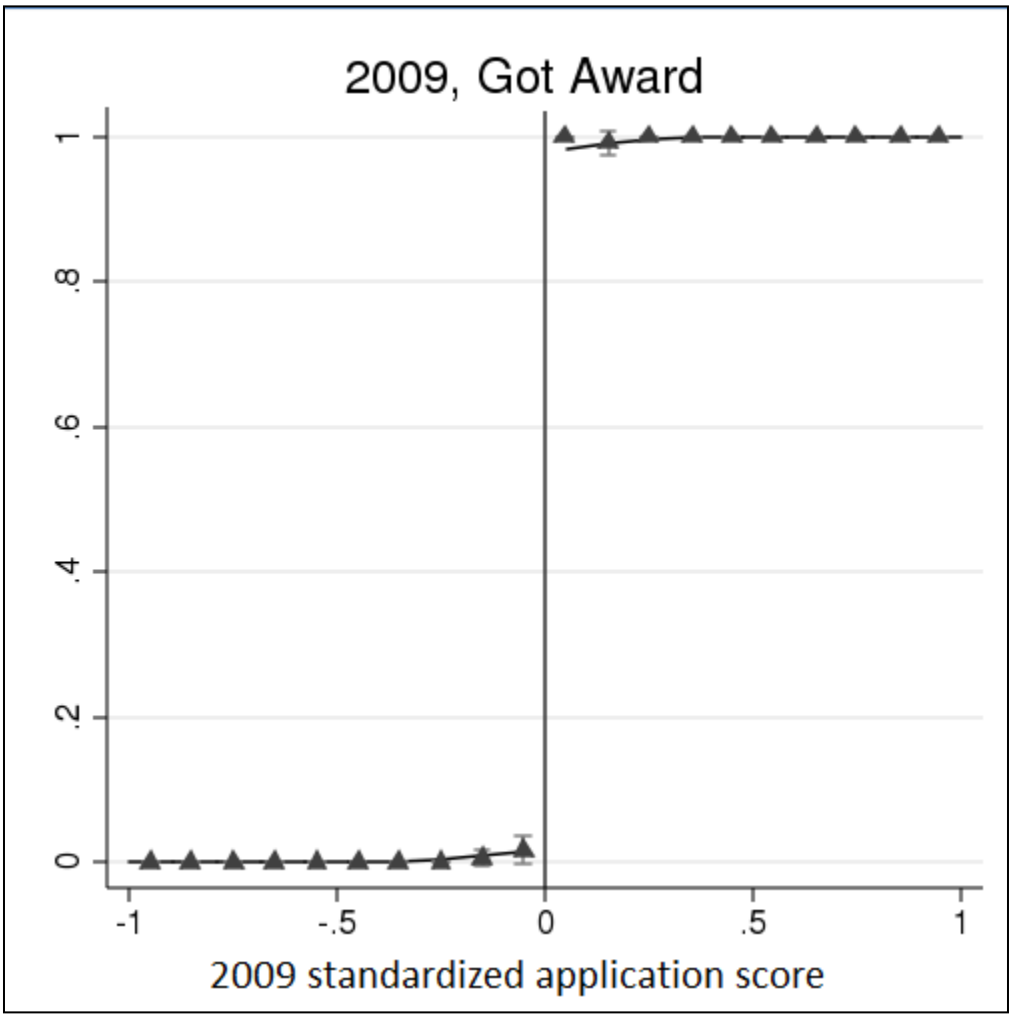


Figure 2: Discontinuity in receiving COPS award funding in dollars

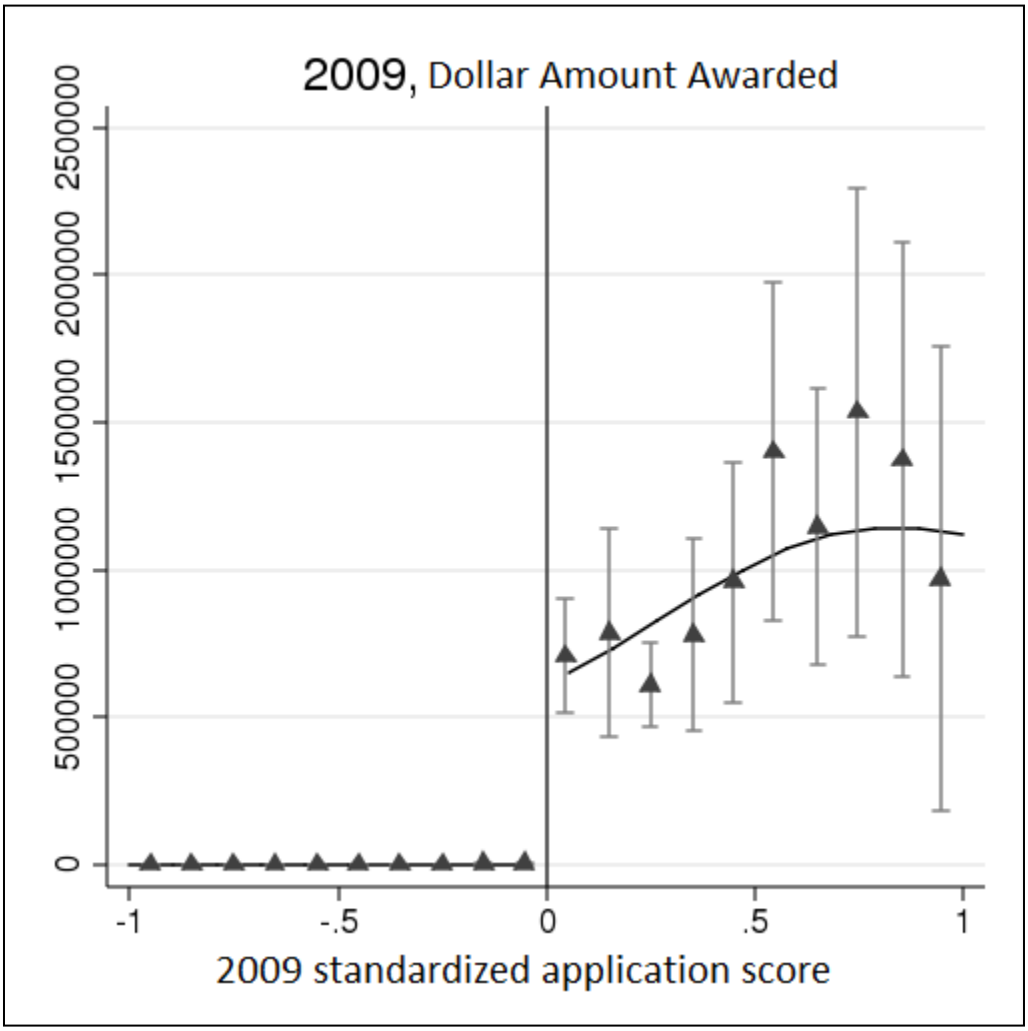


Figure 3: 2008 Sworn officers count

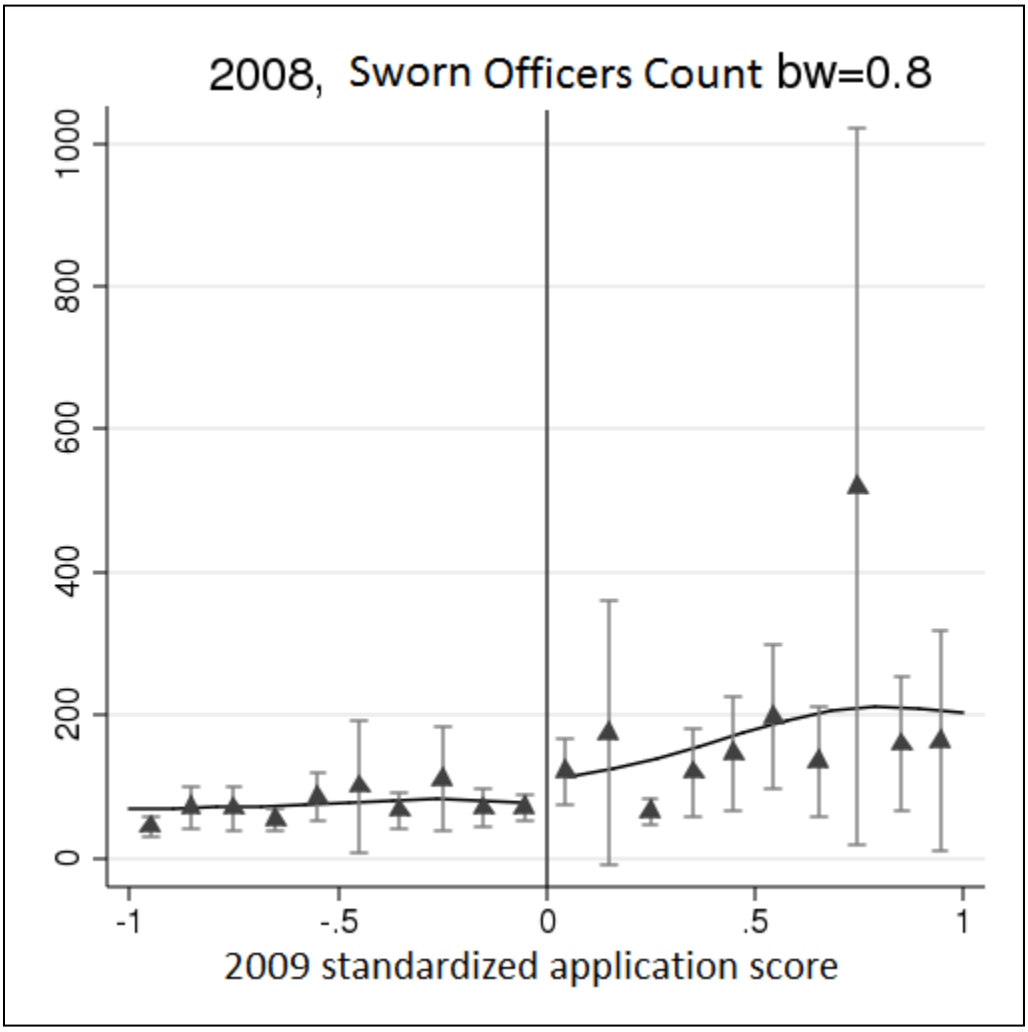


Figure 4: 2008 Sworn officers per 100,000 capita

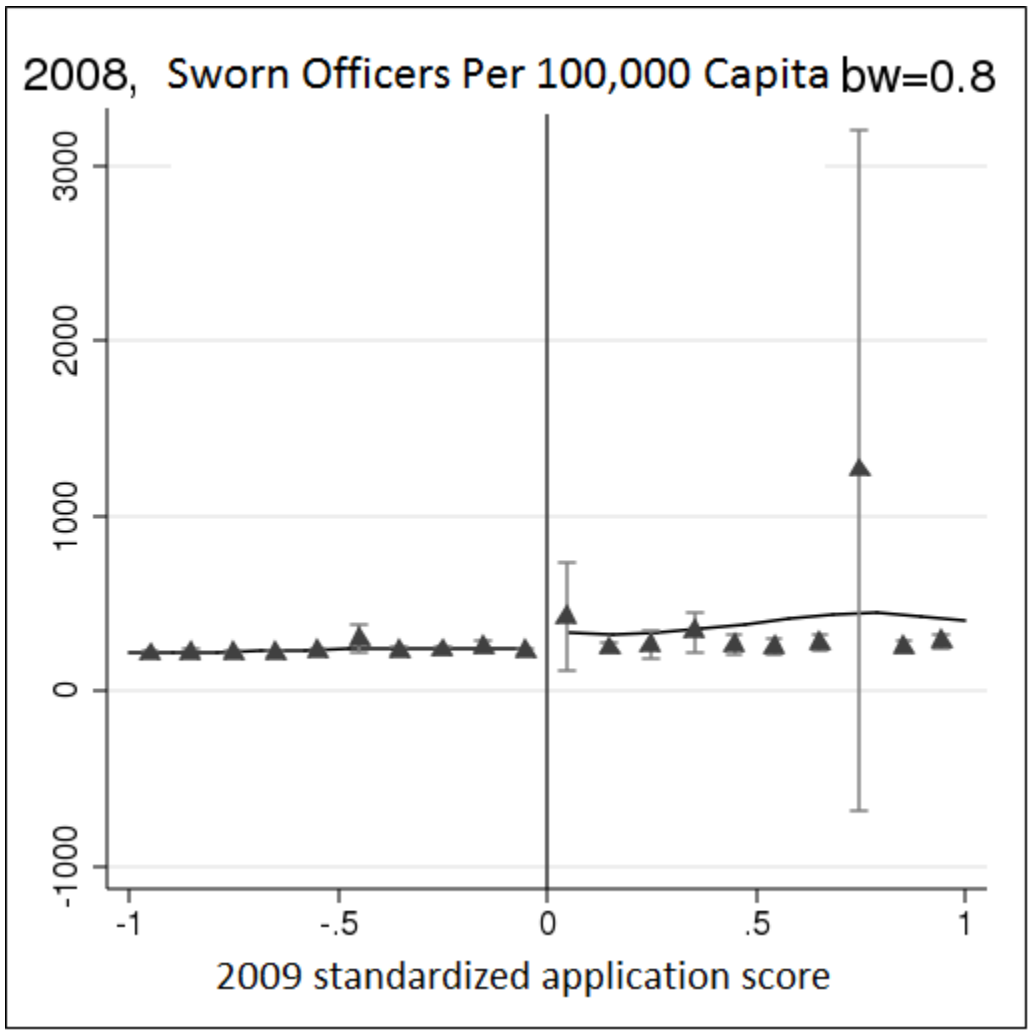


Figure 5: 2008 Log sworn officers count

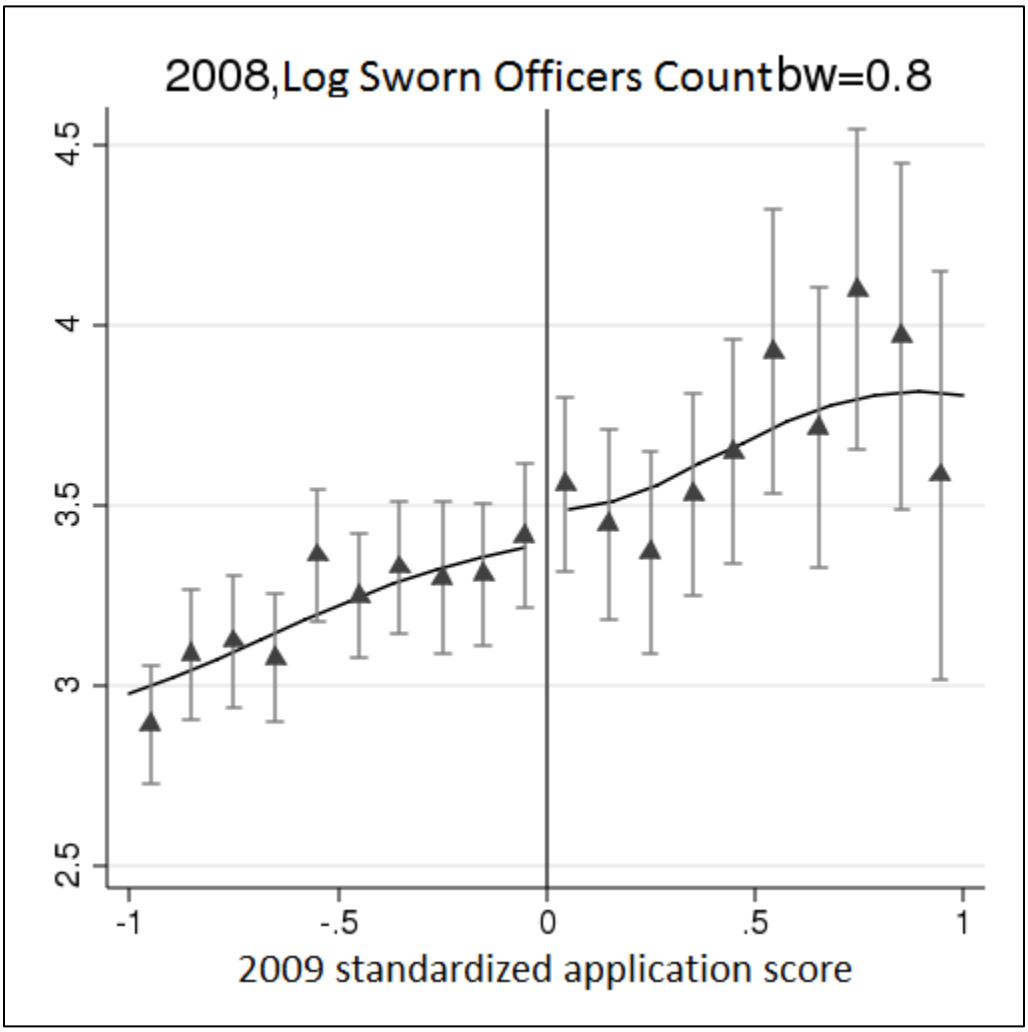


Figure 6: 2008 Log sworn officers per 100,000 capita

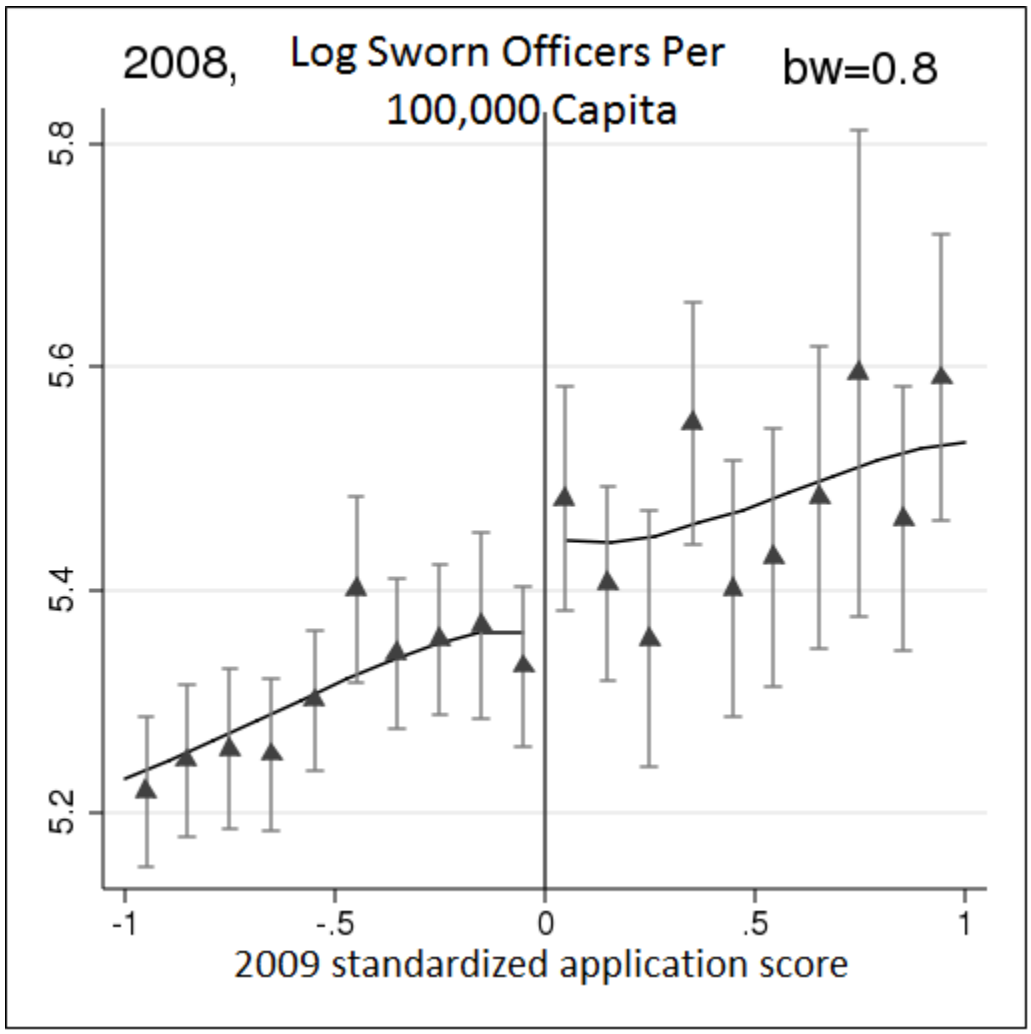


Figure 7: 2008 Difference of logs of sworn officers count relative to the base years of 2005-2006

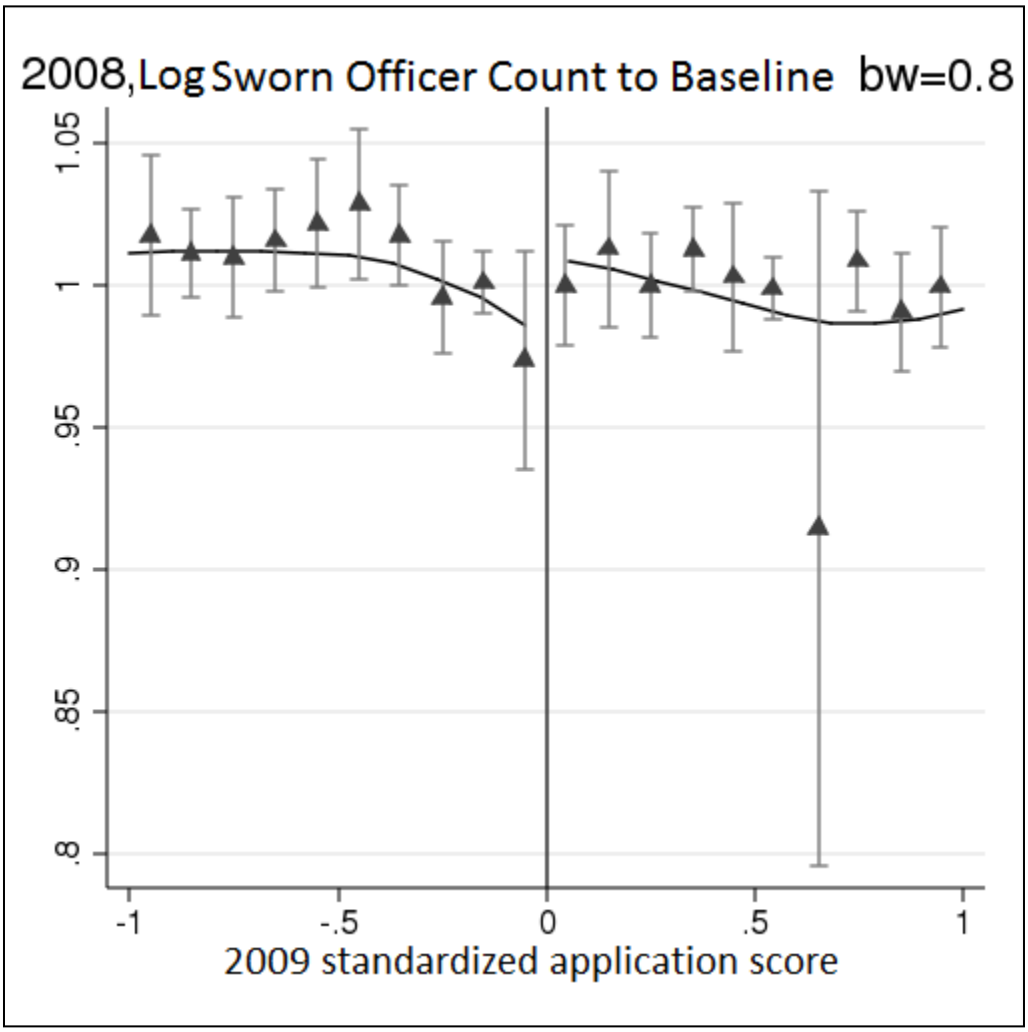


Figure 8: 2008 Difference of logs of sworn officers per 100,000 capita relative to the base years 2005-2006

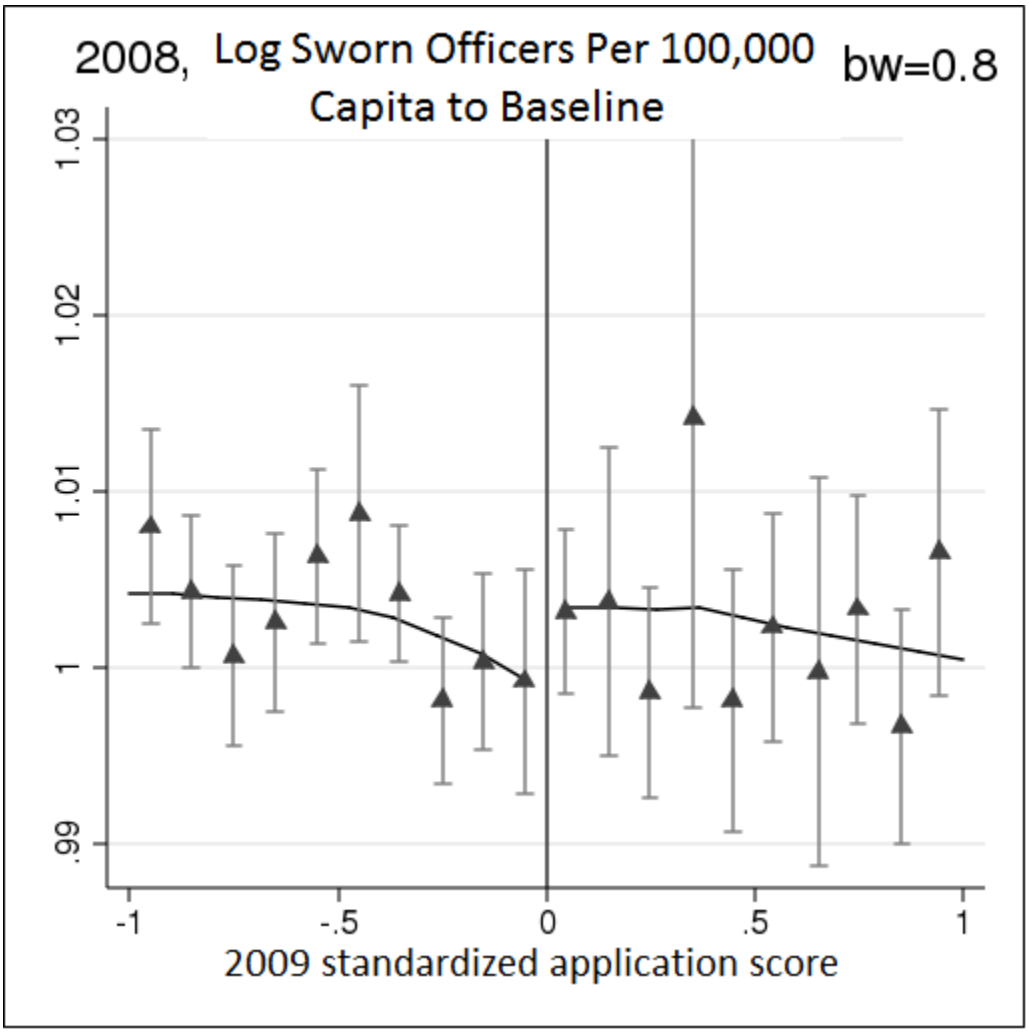


Figure 9: Log of sworn officers per 100,000 capita compared to base years 2005-2006

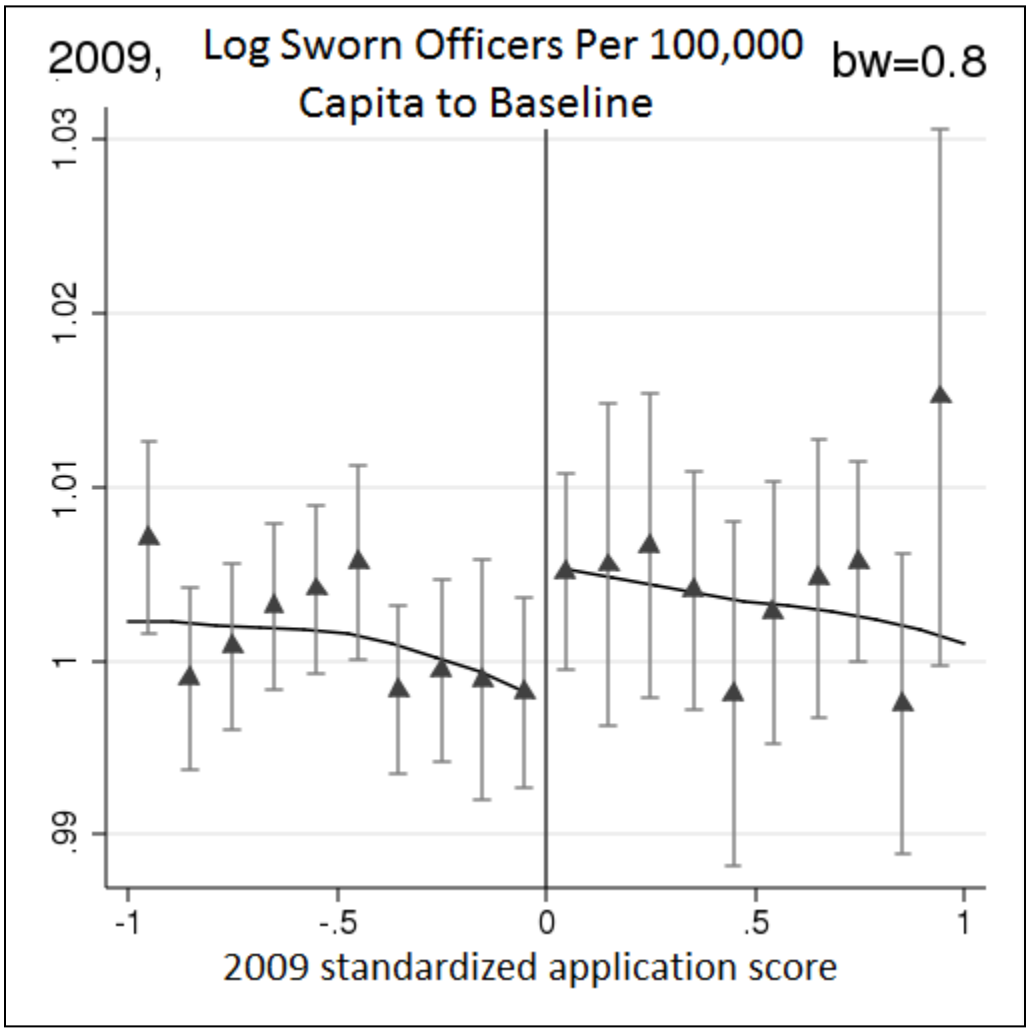


Figure 10: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006 for all agencies

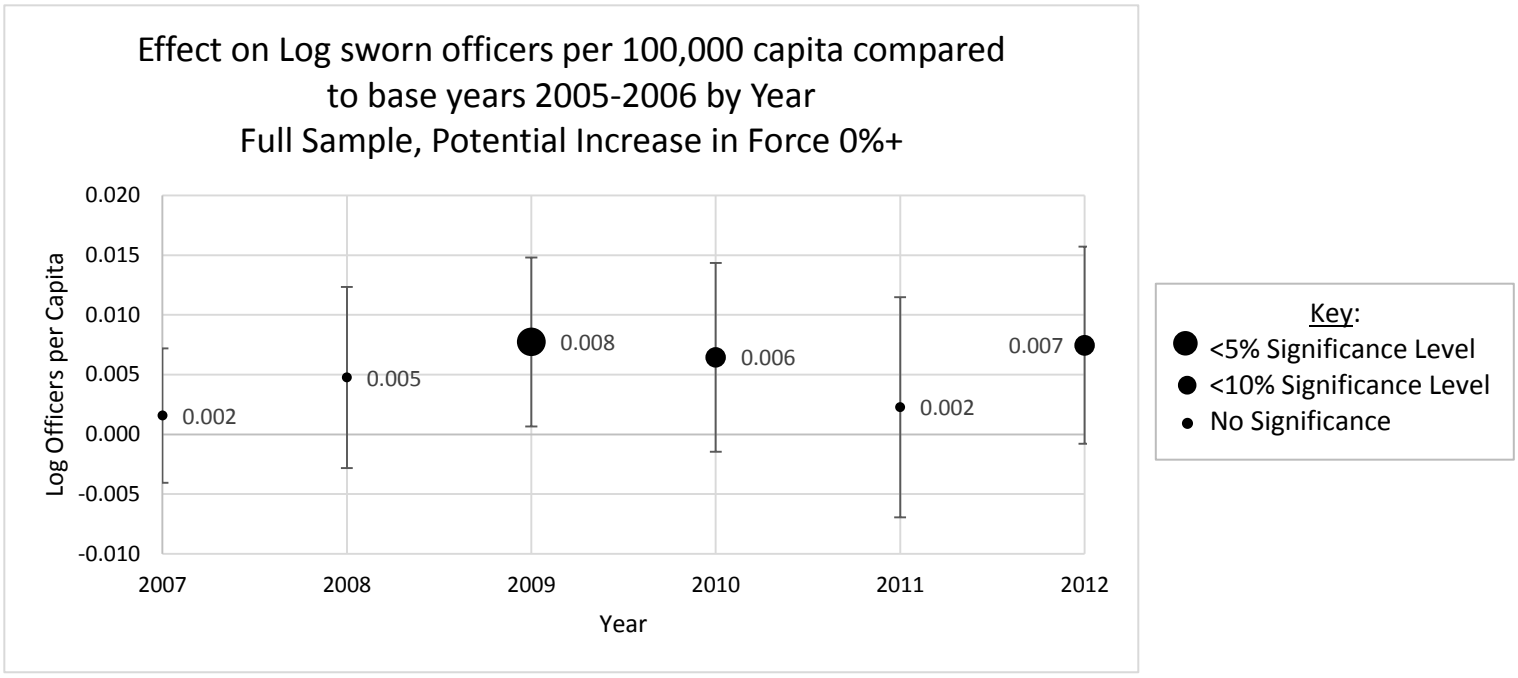


Figure 11: Panel A: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006, for agencies 0-50k, full sample

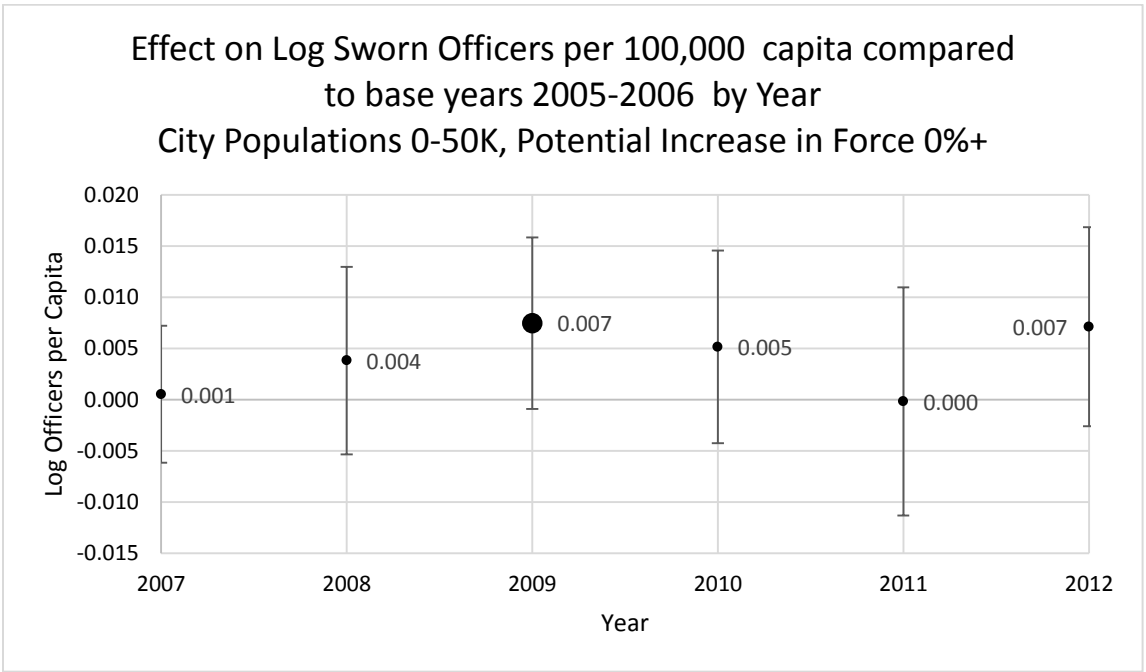


Figure 11: Panel B: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006, for agencies 0-50k, potential impact 5%

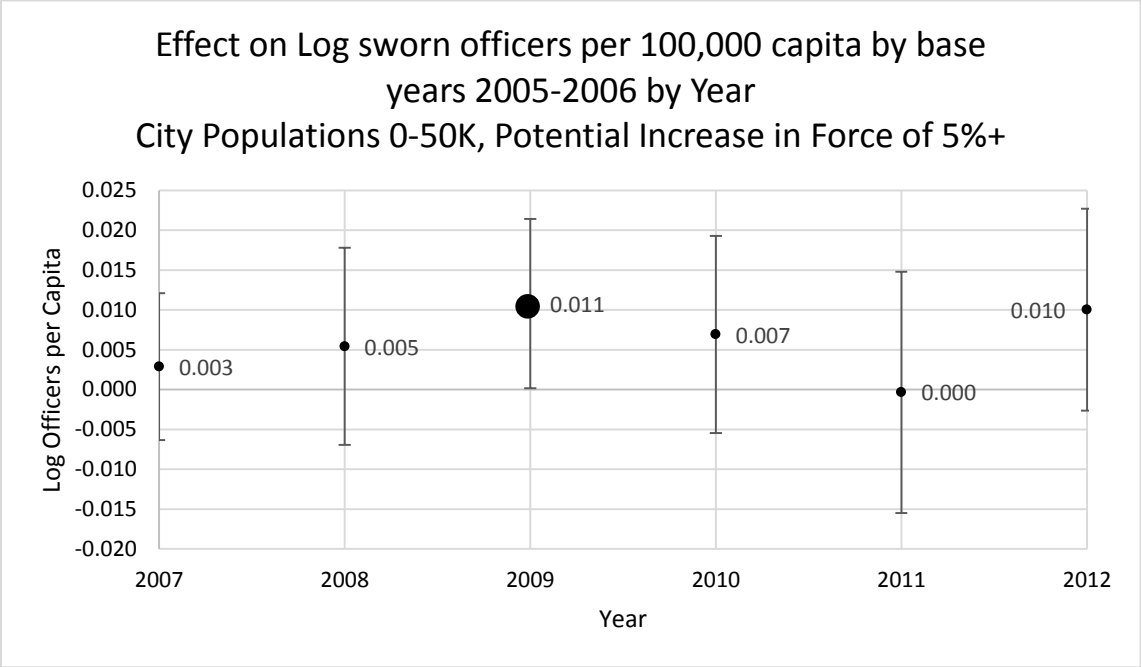


Figure 11: Panel C: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006, for agencies 0-50k, potential impact 10%

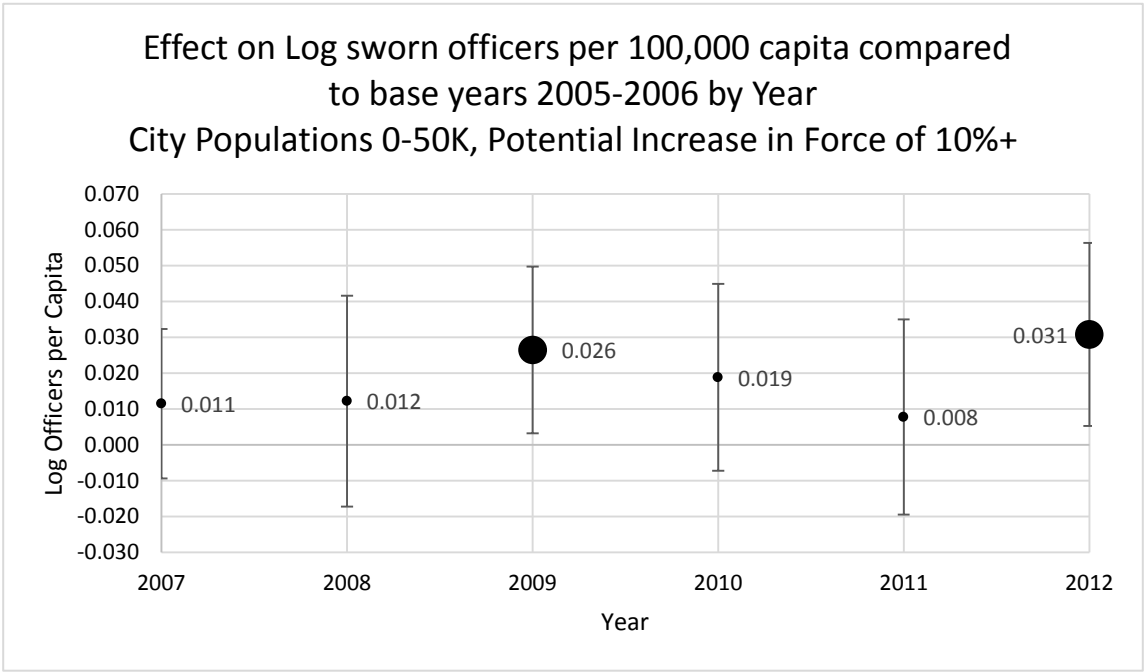


Figure 12: Panel A: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006, for agencies 50-150k, full sample

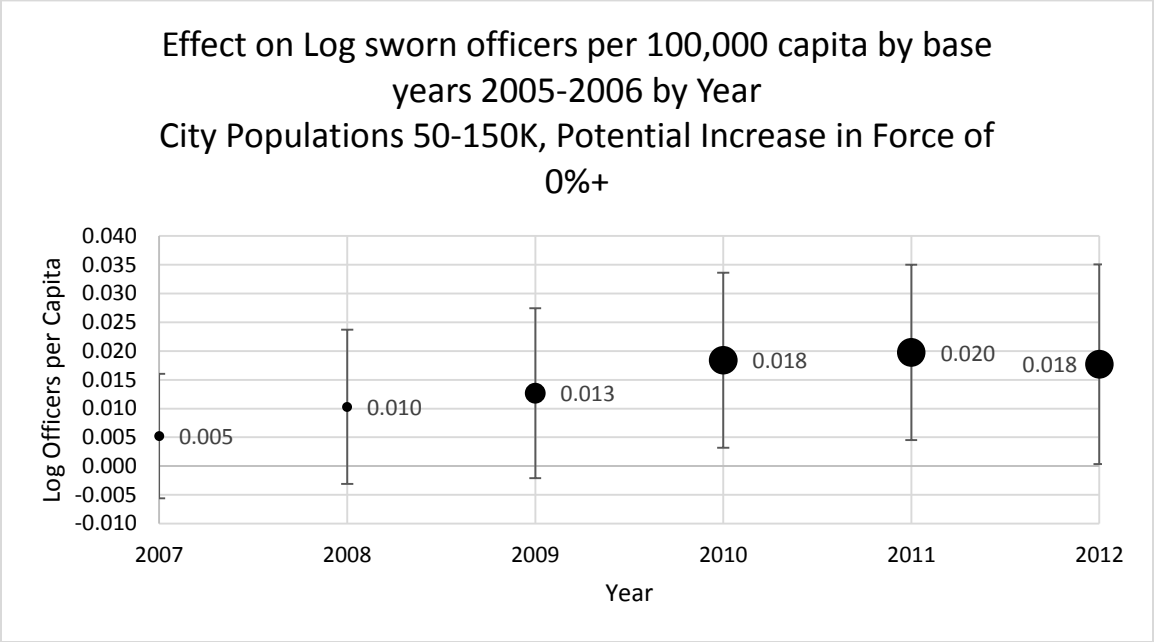


Figure 12: Panel B: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006, for agencies 50-150k, potential impact 5%

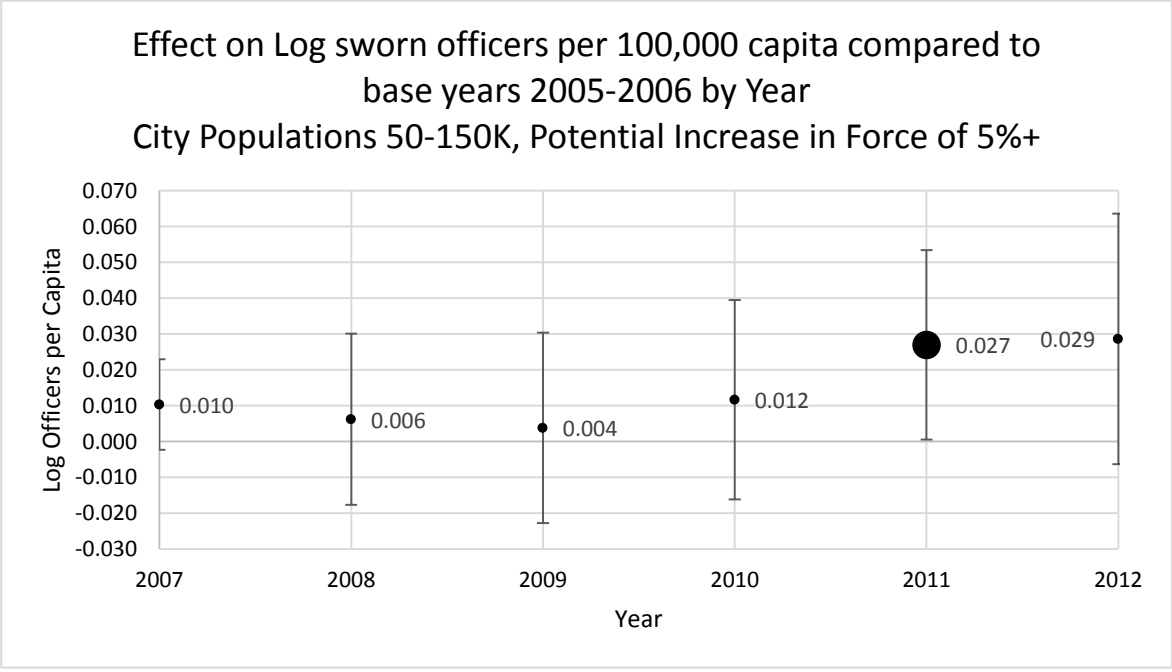


Figure 13: Panel A: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006, for agencies 150k+, full sample

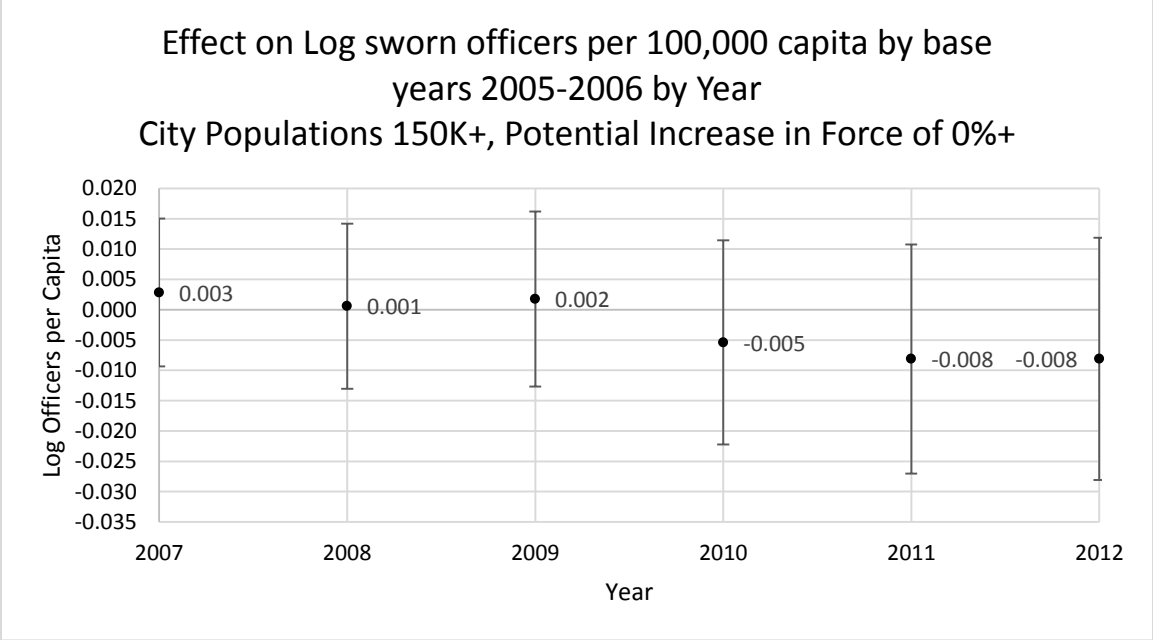


Figure 14: Panel A: Difference of logs of sworn force officers per 100,000 capita relative to the base years 2005-2006, for agencies 150k+, potential impact 5%

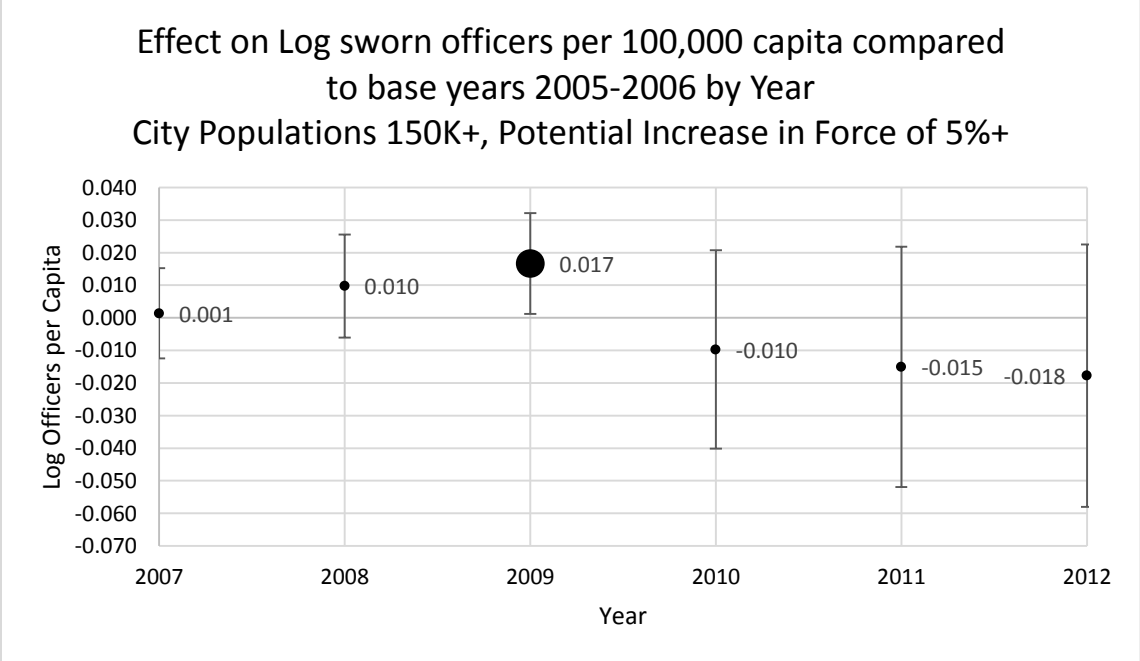


Figure 15: 2009 Total crime per 100,000 capita compared to base years 2005-2006

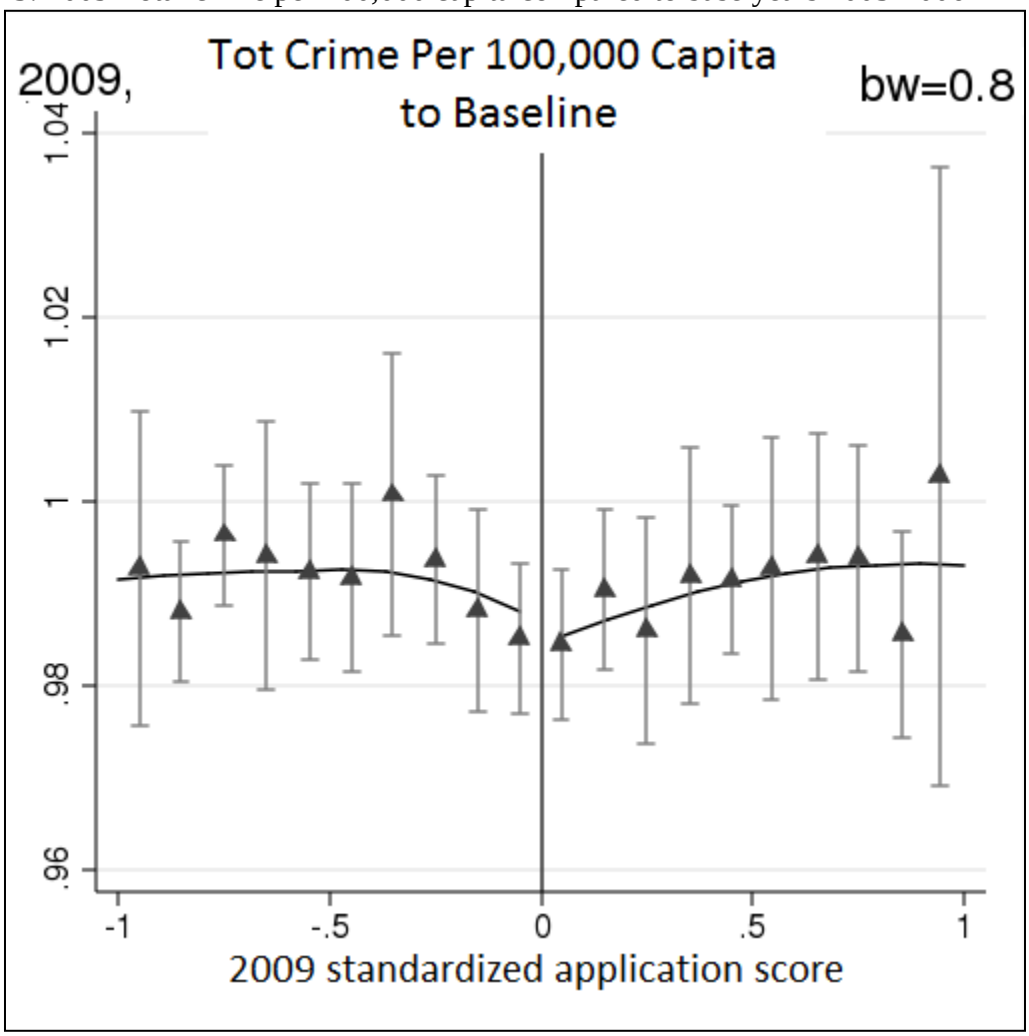


Figure 16: 2010 Total crime per 100,000 capita compared to base years 2005-2006

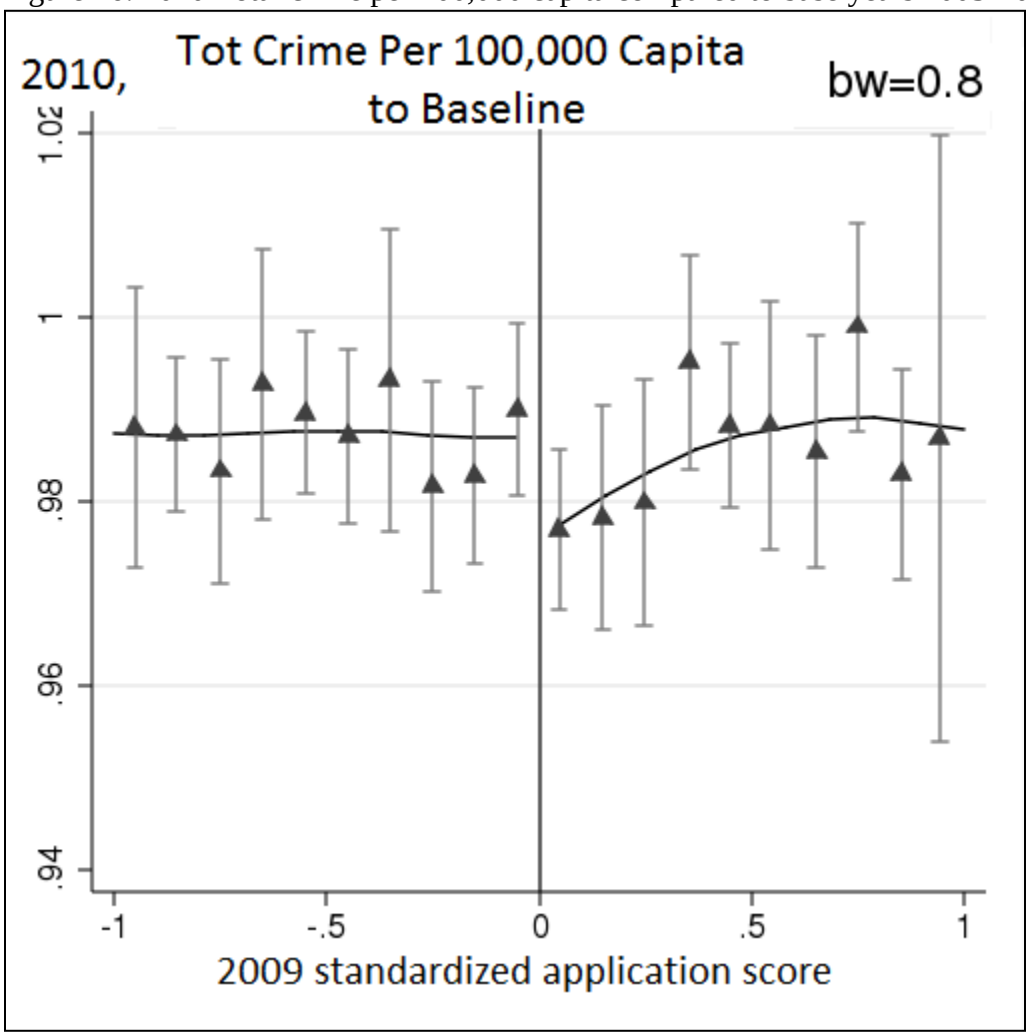


Figure 17: Difference of logs of total crime per 100,000 capita relative to the base years 2005-2006, full sample

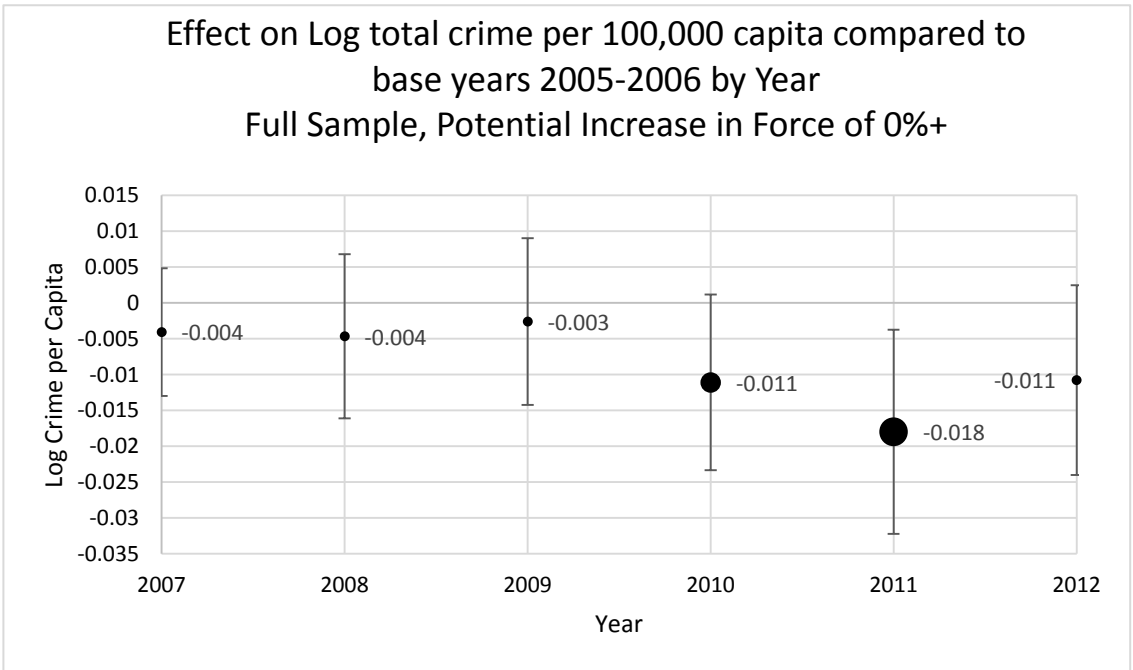


Figure 18: Difference of logs of violent crime per 100,000 capita relative to the base years 2005-2006, full sample

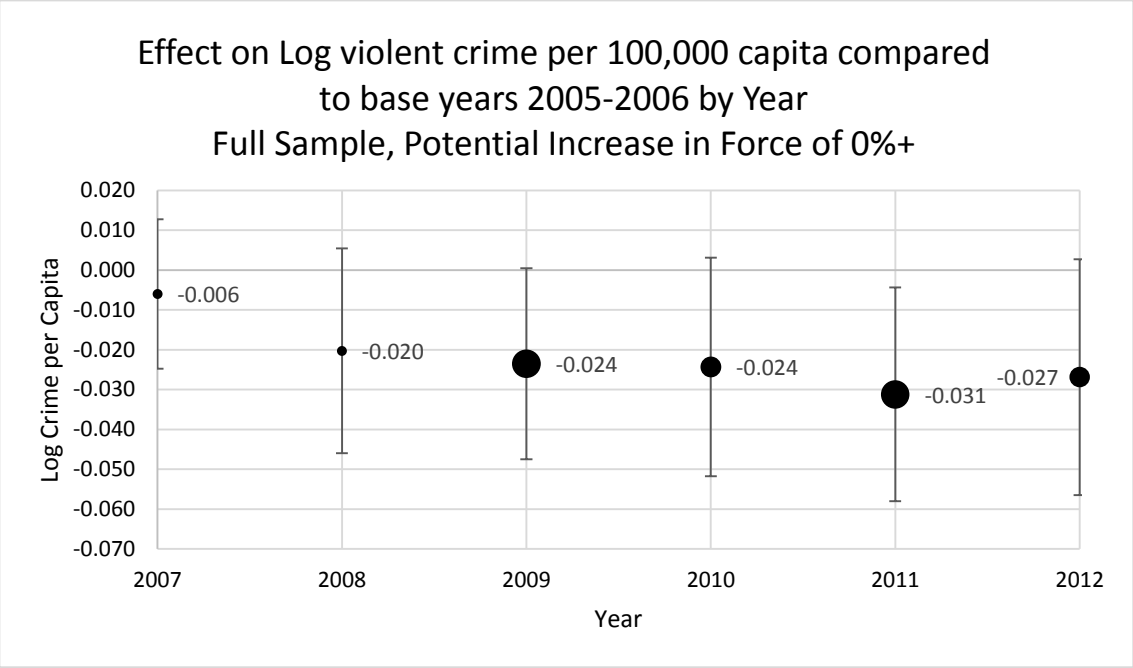


Figure 19: Difference of logs of property crime per 100,000 capita relative to the base years 2005-2006, full sample

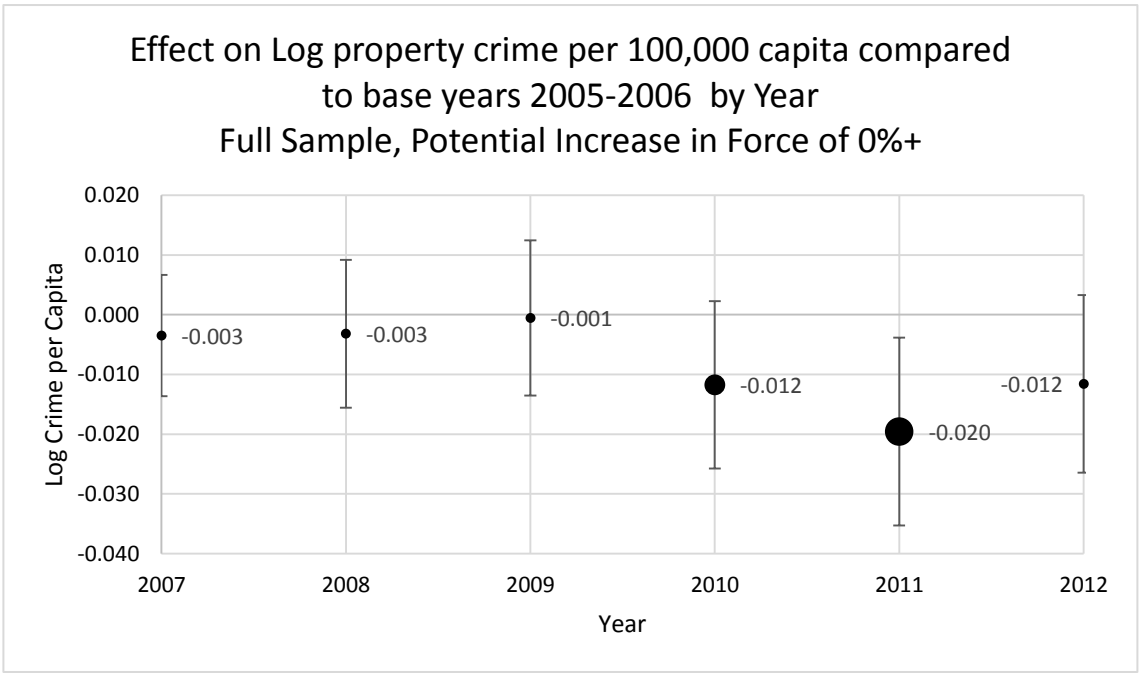


Figure 20: Difference of logs of total crime per 100,000 capita relative to the base years 2005-2006, for agencies 0-50k, potential impact 10%

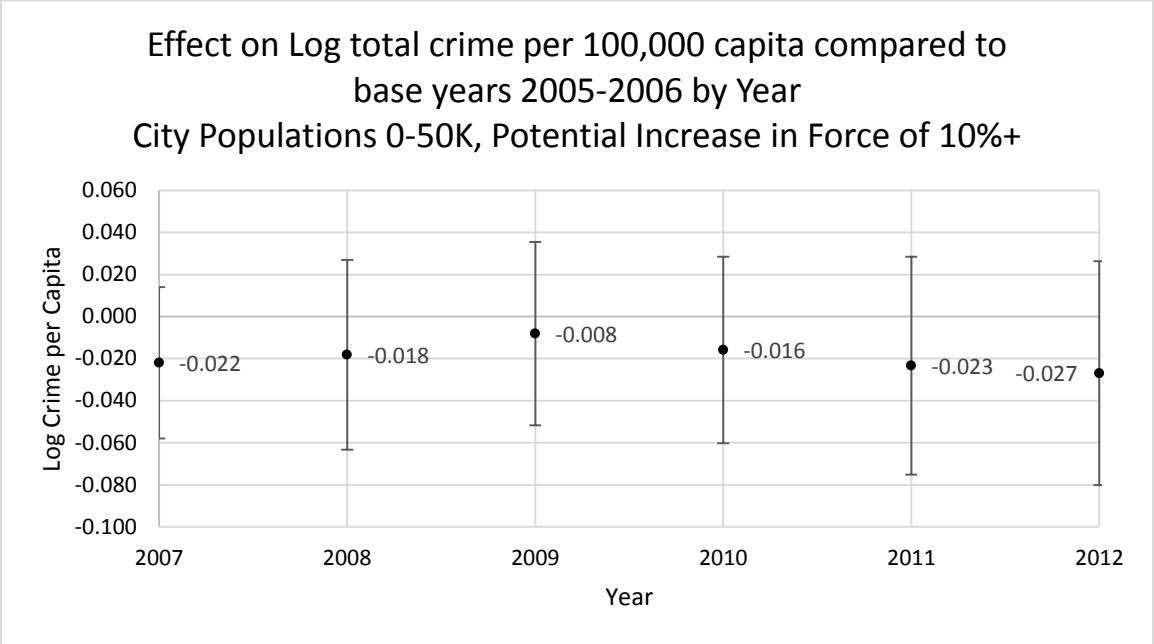


Figure 21: Difference of logs of violent crime per 100,000 capita relative to the base years 2005-2006, for agencies 0-50k, potential impact 10%

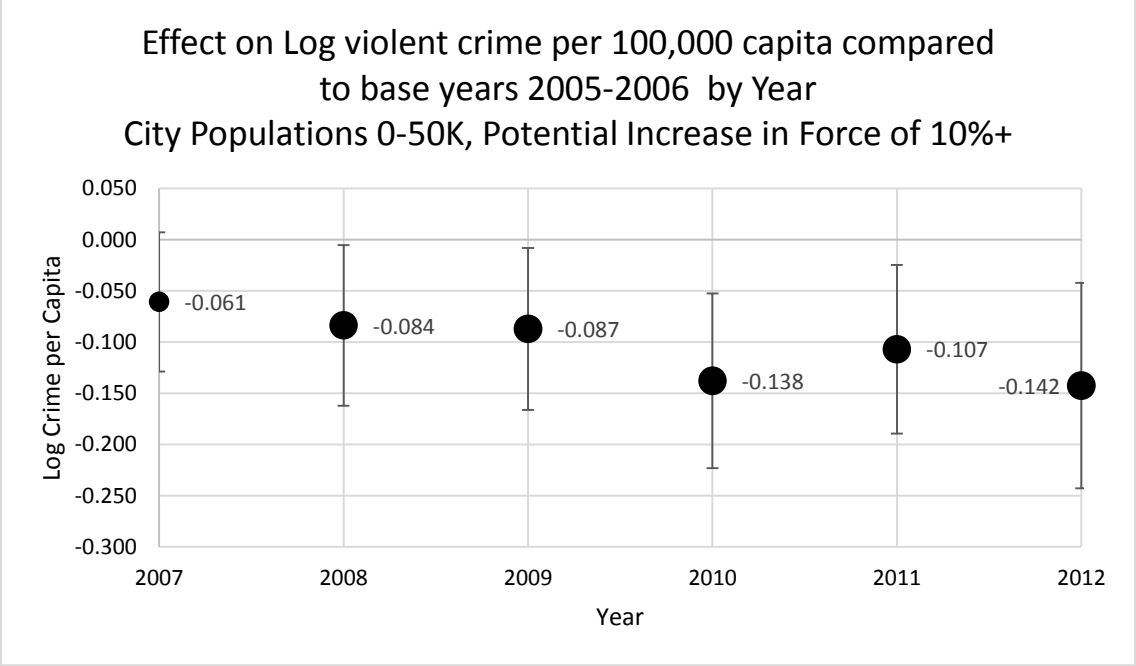


Figure 22: Difference of logs of property crime per 100,000 capita relative to the base years 2005-2006, for agencies 0-50k, potential impact 10%

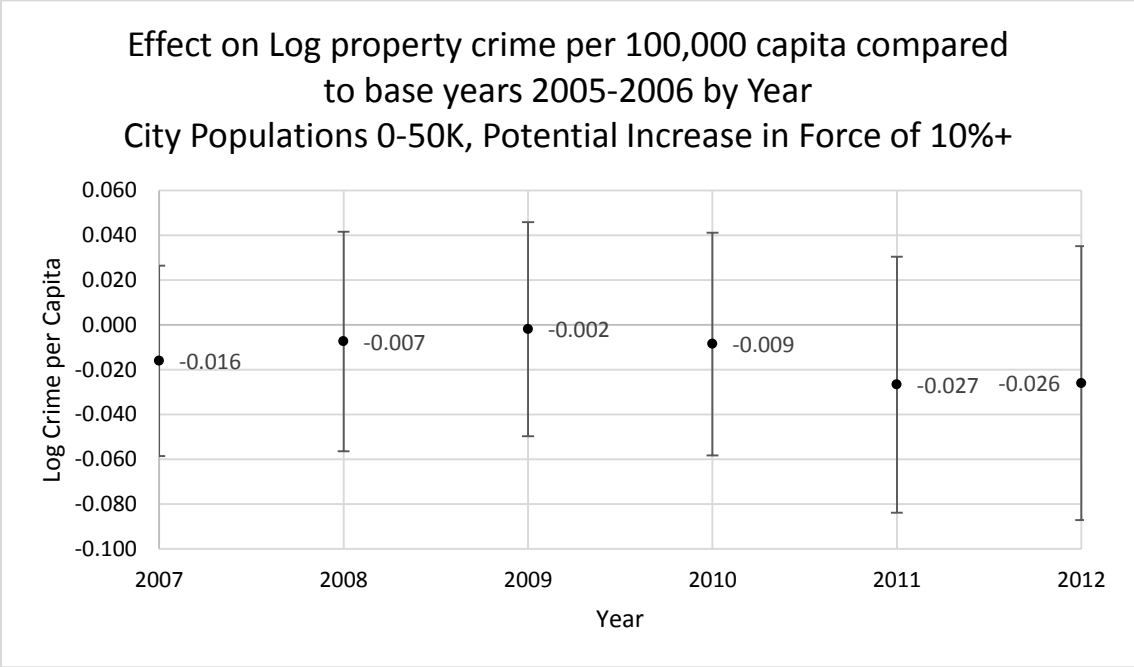


Figure 23: Difference of logs of total crime per 100,000 capita relative to the base years 2005-2006, for agencies 50-150k, potential impact 5%

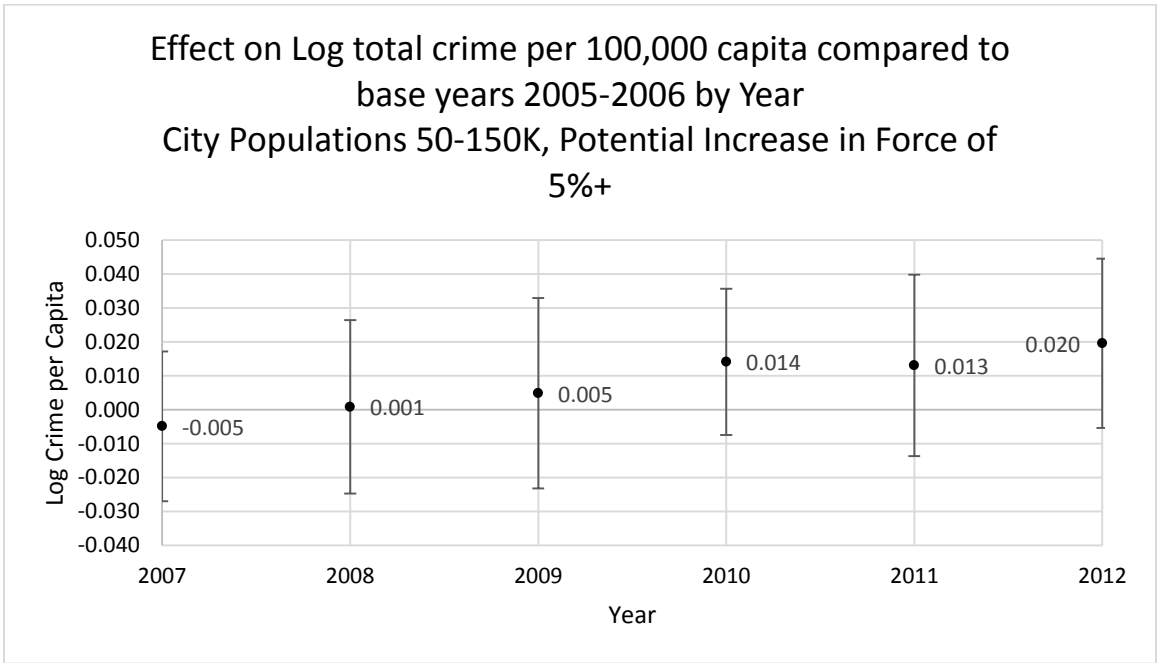


Figure 24: Difference of logs of violent crime per 100,000 capita relative to the base years 2005-2006, for agencies 50-150k, potential impact 5%

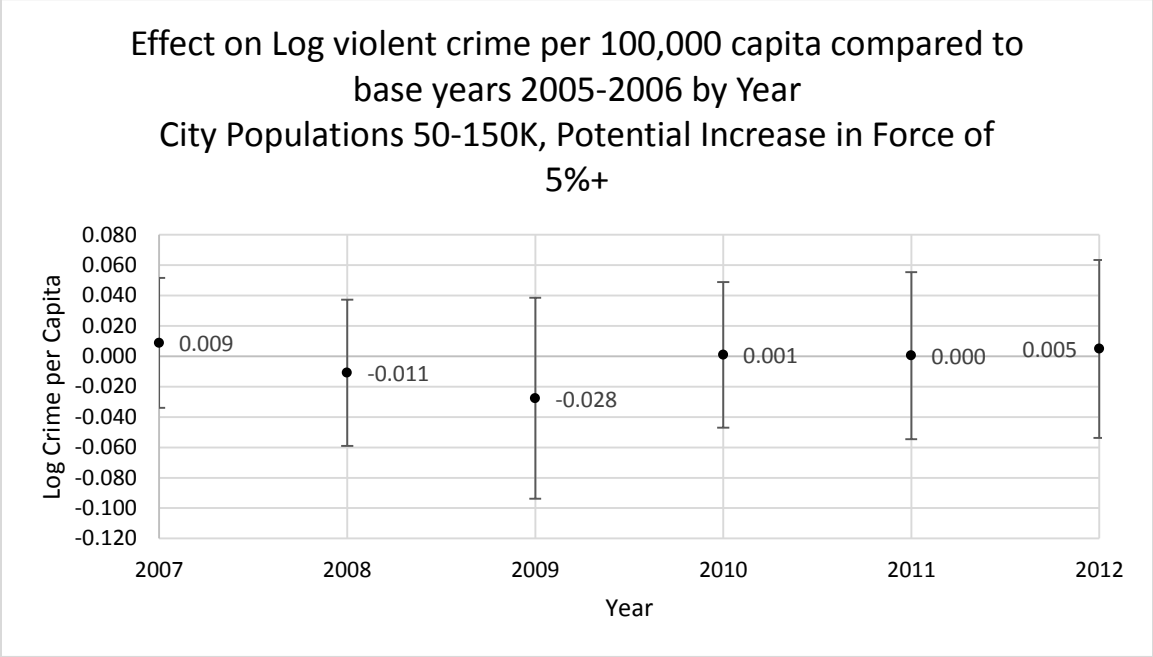


Figure 25: Difference of logs of property crime per 100,000 capita relative to the base years 2005-2006, for agencies 50-150k, potential impact 5%

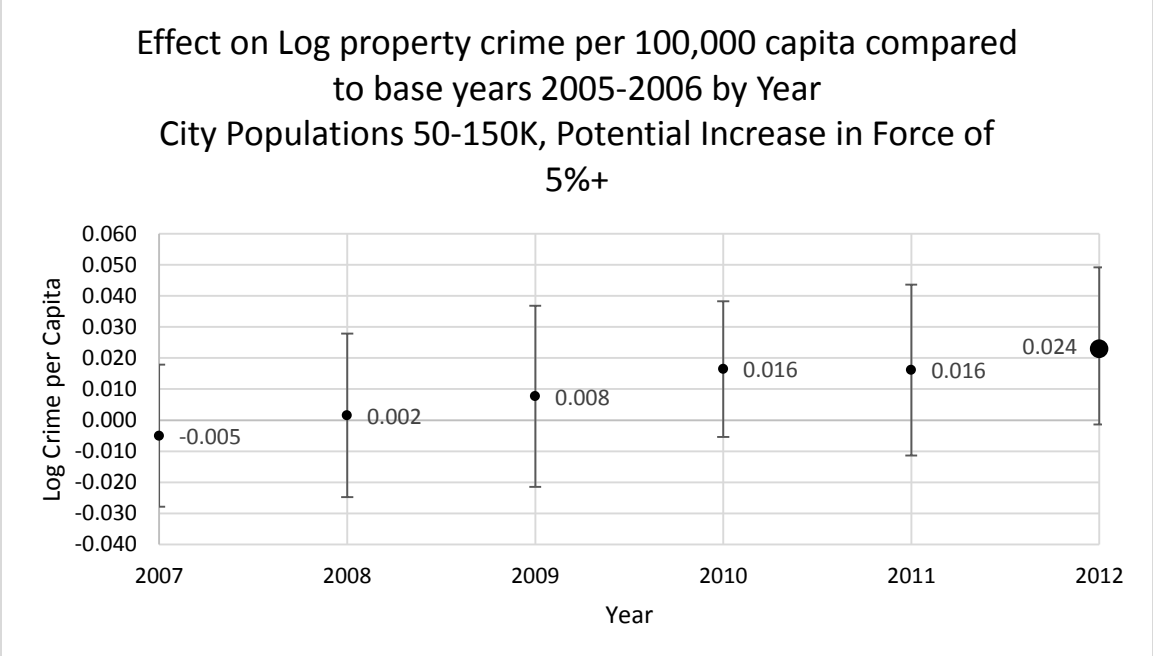


Figure 26: Difference of logs of total crime per 100,000 capita relative to the base years 2005-2006, for agencies 150k+, potential impact 5%

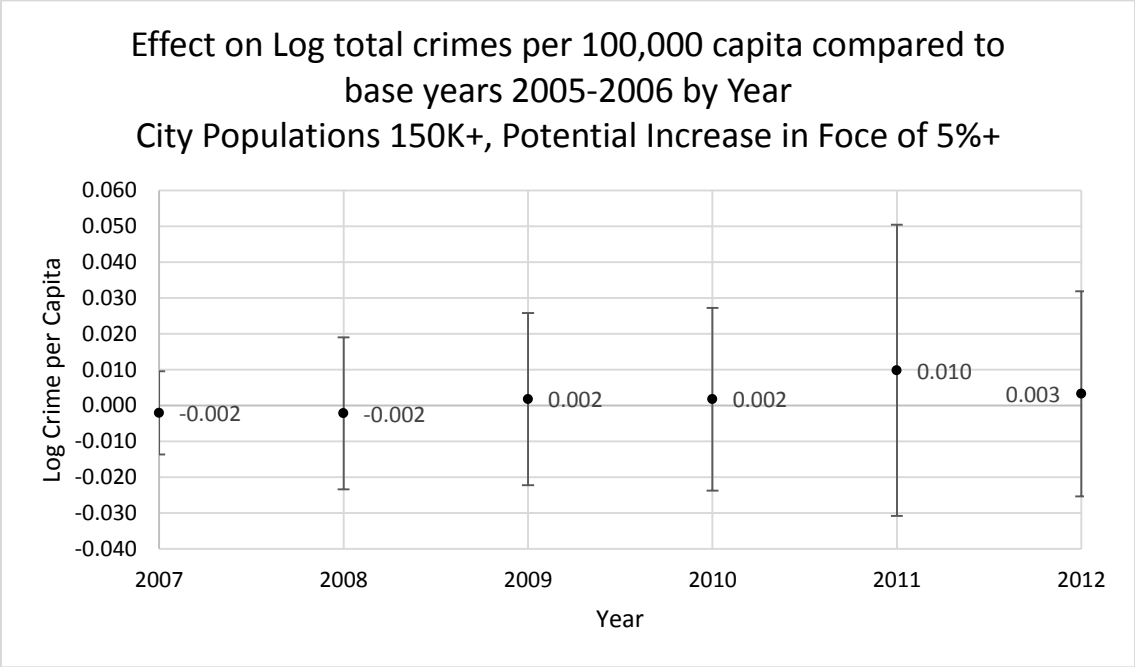


Figure 27: Difference of logs of violent crime per 100,000 capita relative to the base years 2005-2006, for agencies 150k+, potential impact 5%

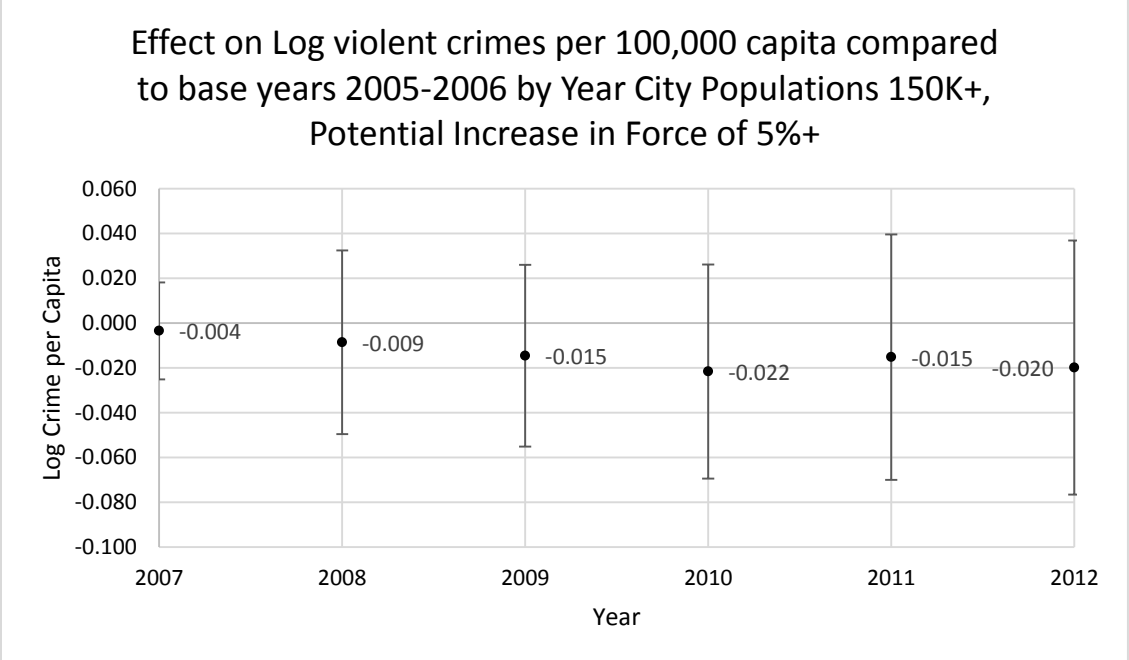


Figure 28: Difference of logs of property crime per 100,000 capita relative to the base years 2005-2006, for agencies 150k+, potential impact 5%

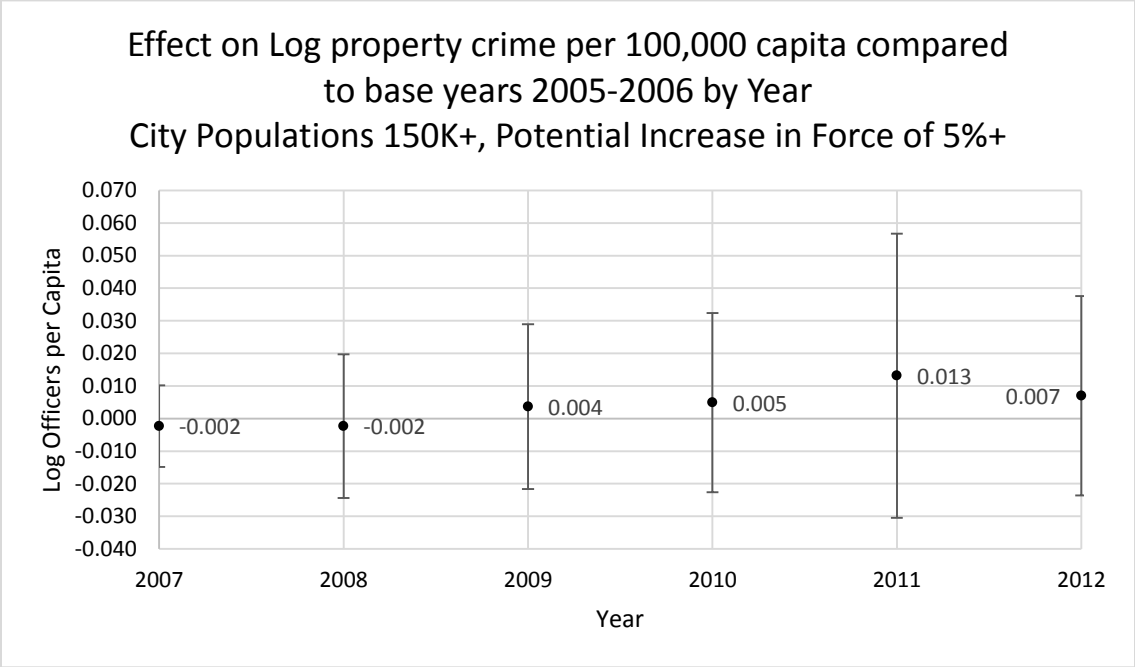


Figure 29: Difference of logs of murder per 100,000 capita relative to the base years 2005-2006, full sample

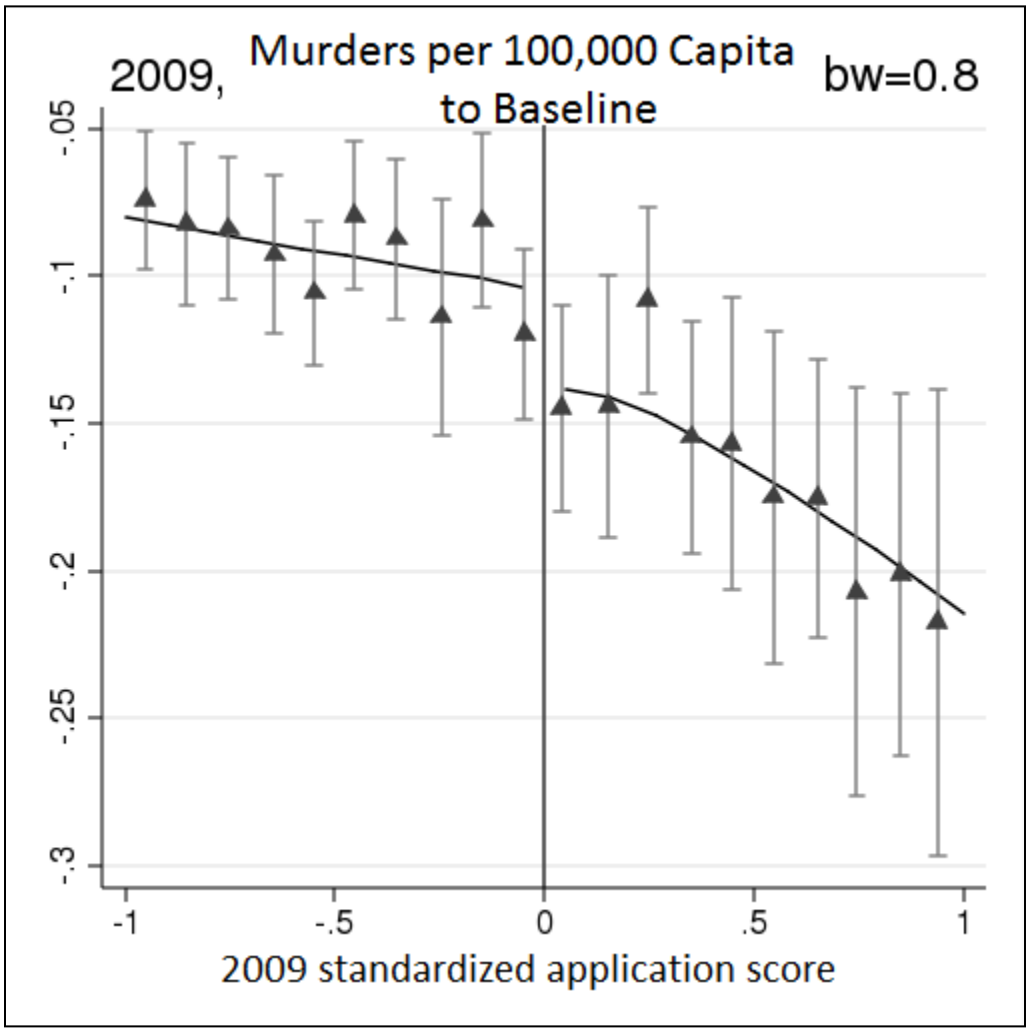


Figure 29: Panel A: Difference of logs of murder per 100,000 capita relative to the base years 2005-2006, full sample

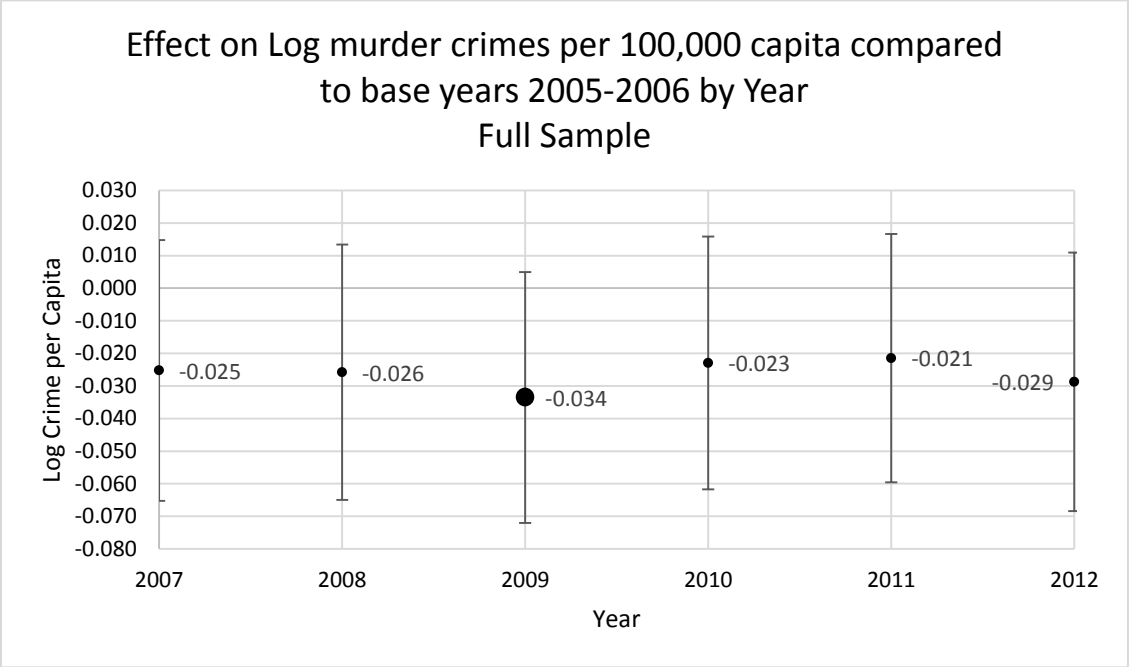


Figure 30: Difference of logs of total arrests per 100,000 capita relative to the base years 2005-2006, full sample

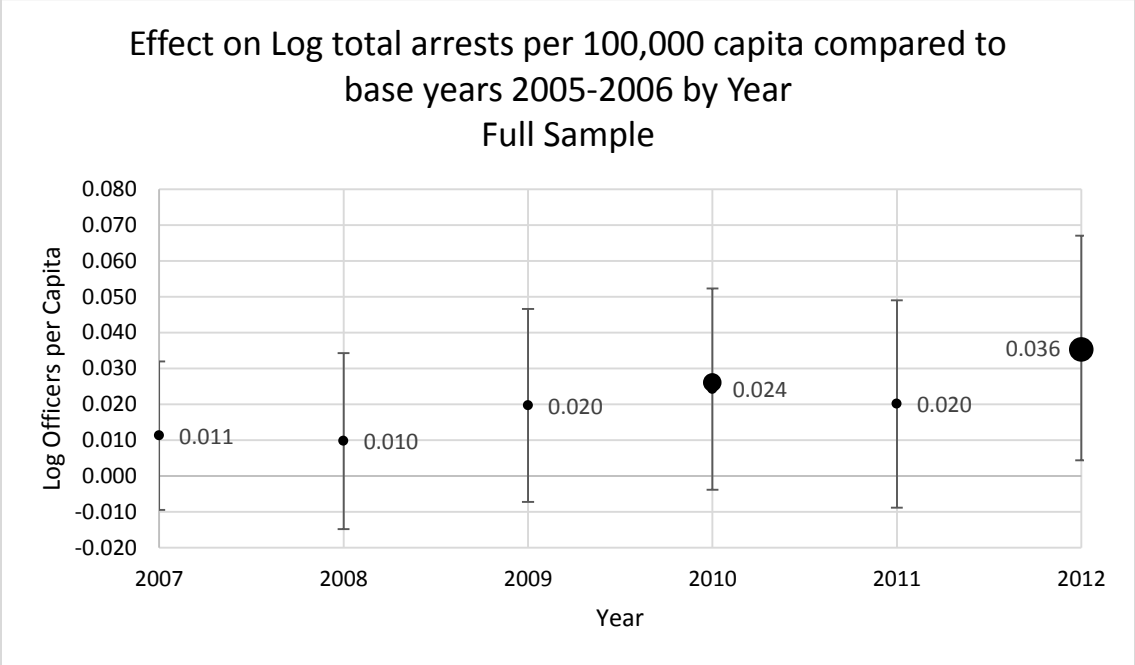


Figure 31: Difference of logs of violent arrests per 100,000 capita relative to the base years 2005-2006, full sample

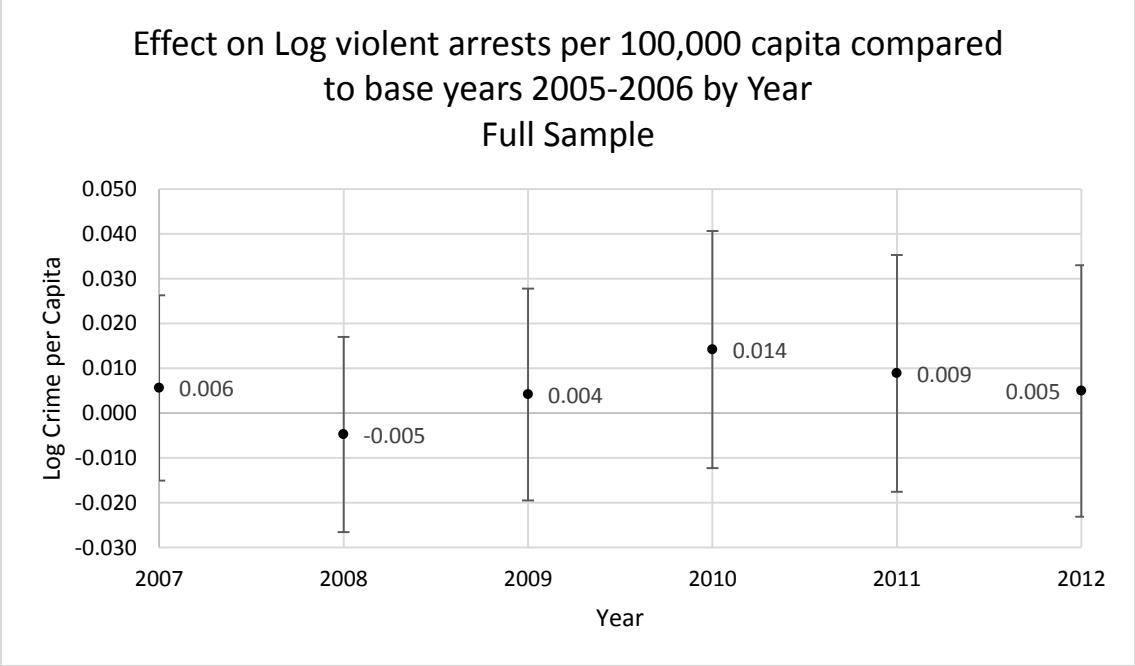
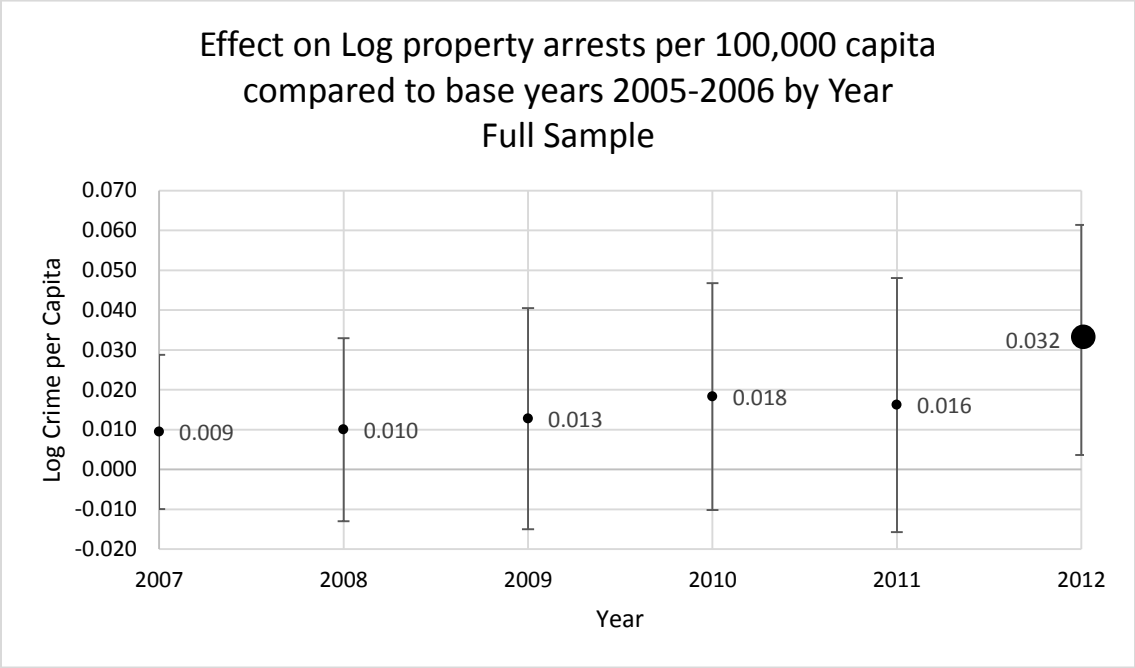


Figure 32: Difference of logs of property arrests per 100,000 capita relative to the base years 2005-2006, full sample



REFERENCES

Becker, Gary, 1968. "Crime and punishment: An economic approach." *Journal of Political Economy*. 76(2): 169-217.

Cameron, Samuel, 1988. "The economics of crime deterrence: A survey of theory and evidence." *Kyklos*. 41(2): 301-323.

Chalfin, Aaron and Justin McCray, 2012. The Effect of Police on Crime: New Evidence from U.S. Cities, 1960-2010. *NBER working paper 18815.*

Cheng, Ming-Yen, Jianqing Fan, and J. S. Marron,1997. "On Automatic Boundary Corrections," *Annals of Statistics*, XXV, 1691–1708.

Cook, Philip J., 1979. The Clearance Rate as a Measure of Criminal Justice System Effectiveness. *Journal of Public Economics*. 11(1):135-142.

Corman, Hope and H. Naci Mocan, 2000. A time series analysis of crime, deterrence, and drug abuse in New York City, *American Economic Review*. 90: 584–604.

Davis, Gareth, David B. Muhlhausen, Dexter Ingram and Ralph Rector, 2000. The facts about COPS: a performance overview of the Community Oriented Policing Services program, *Heritage Foundation Center for Data Analysis Report CDA00-10*.

DiTella, Rafeal, and Ernesto Schargrodsky, 2004. Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. *American Economic Review* 94: 115–133.

Draca, Mirko, Stephen Machin, and Robert Witt. 2011. "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks." *American Economic Review*, 101(5): 2157–81.

Durlauf, Steve N. and Daniel S. Nagin, 2010. The Deterrent Effect of Imprisonment. *NBER working paper c12078.*

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

Eck, John and Edward Maguire , 2000. "Have changes in policing reduced violent crime? An assessment of the evidence." In *The Crime Drop in America*, edited by Alfred Blumstein and Joel Wallman. New York: Cambridge University Press. pp. 207-265.

Evans, William N., and Emily G. Owens, 2007. COPS and Crime. *Journal of Public Economics* 91(2): 181-201.

Fagan, Jeffrey A., Amanda Geller, Garth Davies, and Valerie West, 2010. Street Stops and Broken Windows Revisited: The Demography and Logic of Proactive Policing in a Safe and Changing City. In Rice, Stephen K. and M. D. White, Editors, *Race, Ethnicity, and Policing: New and Essential Readings*, New York University Press: New York. 309-348.

Fan, Jianqing, 1992. "Design-Adaptive Nonparametric Regression," *Journal of the American Statistical Association*, LXXXVII, 998–1004.

Fan, Jianqing, and Irene Gijbels, 1996. *Local Polynomial Modelling and Its Applications*, New York: Chapman and Hall.

Fisher, Franklin and Daniel Nagin, 1978. "On the feasibility of identifying the crime function in a simultaneous equations model of crime and sanctions." In Alfred Blumstein, Daniel Nagin, and Jacqueline Cohen, Eds. *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*. Washington, DC: National Academy of Sciences.

Klick, Jonathan and Alex Tabarrok, 2005. Using terror alert levels to estimate the effect of police on crime. *Journal of Law and Economics* 48: 267–279.

Koper, Christopher, Edward E. Maguire, Gretchen E. Moore, and David E. Huffer, 2001. Hiring and Retention Issues in Police Agencies: Readings on the Determinants of Police Strength, Hiring and Retention of Officers, and the Federal COPS Program. Urban Institute Report to the National Institute of Justice.

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

Levitt, Steven D., 2002. Using electoral cycles in police hiring to estimate the effect of police on crime: reply. *American Economic Review* 92: 1244–1250.

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

Lin, Ming-Jen, 2009. More police, Less Crime: Evidence from US state data. *International Review of Law and Economics* 29: 73-80.

Maltz, Michael D., 1999. Bridging gaps in police crime data. *NCJ Publication No. 176365*.

Maltz, Michael D. and Harald E. Weiss, 2006. Creating a UCR Utility: Final Report to the National Institute of Justice. *NIJ Research Report 215341*.

Marvell, Thomas B., and Carlisle E. Moody, 1996. Specification Problems, Police Levels, and Crime Rates. *Criminology* 34: 609–646.

Miller, Ted, Mark Cohen, and Brian Wiersema. 1996. "Victim Costs and Consequences: A New Look." National Institute of Justices, U.S. Department of Justice.

Owens, Emily G., 2013. "COPS and Cuffs" in Philip Cook, Stephen Machin, Olivier Marie, and Giovanni Mastrobouni [Eds.] Lessons from the Economics of Crime: What Works in Reducing Offending? Cambridge: MIT Press, *forthcoming.*

Owens, Emily G., 2009. "More time, less crime? Estimating the incapacitative effect of sentence enhancements". Journal of law and economics. 52:551-580.

Porter, Jack, 2003. "Estimation in the Regression Discontinuity Model." Working Paper, Harvard University, Department of Economics.

U.S. Department of Justice, Office of the Inspector General, 1999. Police Hiring and Redeployment Grants: Summary of Audit Findings and Recommendations, October 1996–September 1998. Office of the Inspector General Special Report #99-14.

U.S. Government Accountability Office, 2003. Technical Assessment of Zhao and Thurman's 2001 Evaluation of the Effects of COPS Grants on Crime. Report Number GAO-03-867R.

U.S. Government Accountability Office, 2005. Community Policing Grants: COPS Grants Were a Modest Contributor to Declines in Crime in the 1990s. Report Number GAO-06-104.

Vollard, Benjamin and Joseph Hamed, 2011. Why the police have an effect on violent crime after all. Evidence from the British Crime Survey. *Mimeo*

Vollard, Benjamin and Pierre Koning, 2009. The effect of police on crime, disorder and victim precaution. Evidence from a Dutch victimization survey. *International Review of Law and Economics*, 29 (4):336-348.

Worrall, John L. and Tomislav V. Kovandzic, 2010. Police levels and crime rates: An instrumental variables approach. *Social Science Research* 39:506-516.