NBER WORKING PAPER SERIES

# POLICE FORCE SIZE AND CIVILIAN RACE

Aaron Chalfin
Benjamin Hansen
Emily K. Weisburst
Morgan C. Williams, Jr.

Working Paper 28202
http://www.nber.org/papers/w28202

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020

We are grateful to David Autor, Bocar Ba, Shooshan Danagoulian, Aria Golestani, Jens Ludwig, Jacob Kaplan, John MacDonald, Dan O'Flaherty, Emily Owens, Rajiv Sethi and Yulya Truskinovsky as well as seminar participants at the Southern Economic Association Annual Meetings and Wayne State University for helpful comments. Any remaining errors are our own. Correspondence: Benjamin Hansen, Department of Economics, University of Oregon, E-Mail: bchansen@uoregon.edu. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Police Force Size and Civilian Race
Aaron Chalfin, Benjamin Hansen, Emily K. Weisburst, and Morgan C. Williams, Jr.
NBER Working Paper No. 28202
December 2020
JEL No. H72,J15,J18,K42

**ABSTRACT**

We report the first empirical estimate of the race-specific effects of larger police forces in the United States. Each additional police officer abates approximately 0.1 homicides. In per capita terms, effects are twice as large for Black versus white victims. At the same time, larger police forces make more arrests for low-level "quality-of-life" offenses, with effects that imply a disproportionate burden for Black Americans. Notably, cities with large Black populations do not share equally in the benefits of investments in police manpower. Our results provide novel empirical support for the popular narrative that Black communities are simultaneously over and under-policed.

Aaron Chalfin
University of Pennsylvania
achalfin@sas.upenn.edu

Benjamin Hansen
Department of Economics
1285 University of Oregon
Eugene, OR 97403
and NBER
bchansen@uoregon.edu

Emily K. Weisburst
UCLA Luskin School of Public Affairs
Department of Public Policy
337 Charles E Young Dr E
Los Angeles, CA 90095
weisburst@ucla.edu

Morgan C. Williams, Jr.
Robert F. Wagner Graduate School
of Public Service
New York University
Puck Building
295 Lafayette St.
New York, NY 10012-9604
mcw394@nyu.edu

# 1 Introduction

Following increased public attention on police shootings and the growth of social movements like *Black Lives Matter*, American support for law enforcement is currently at its lowest point in nearly thirty years despite the dramatic decline in crime since the 1990s.[1] The large drop in overall support for law enforcement is compounded by a widening race gap in support for police, with 19% of Black Americans expressing confidence in police relative to 56% of white Americans. This seminal and wide-ranging problem is strikingly described by Bratton and Anderson (2018) as the "great divide in American policing."[2]

For decades, activists, policymakers, and social scientists have debated the role of police presence, particularly in lower income neighborhoods where crime tends to be most prevalent. Given the over-representation of Black Americans among both homicide victims and civilians shot by the police, race remains a central fixture of public discourse on policing reform — in particular, reforms that are intended to decrease the exposure of low income minority communities to the collateral costs of policing. Proposed reforms emerging from recent public discourse include widespread calls for reductions in municipal funding for police departments. While there is now a strong consensus in the academic literature that the number of police officers (McCrary, 2002; Evans and Owens, 2007; Chalfin and McCrary, 2018; Mello, 2019; Weisburst, 2019b) combined with their presence and visibility (Sherman and Weisburd, 1995; Di Tella and Schargrodsky, 2004; Klick and Tabarrok, 2005; Braga et al., 2014; MacDonald et al., 2016; Weisburd, 2016) reduces crime, the extent to which the benefits of additional law enforcement accrue equally to Black and white Americans remains a surprisingly open question.

An extensive literature offers several possible explanations as to why homicide reductions that

[1] https://news.gallup.com/poll/317135/amid-pandemic-confidence-key-institutions-surges.aspx
[2] https://news.gallup.com/poll/317114/black-white-adults-confidence-diverges-police.aspx

are attributable to the expansion of law enforcement could differ across racial groups. First, the intense spatial concentration of street vice and homicide in Black neighborhoods provides more opportunities to address victimization through expanded policing efforts (Sampson et al., 1995; Cook et al., 2007; O'Flaherty and Sethi, 2010b). To the extent that an expansion of law enforcement successfully reduces the influence of illicit markets or dedicates additional resources to increasing the opportunity costs of offending, these interventions could lead to disproportionate decreases in Black homicide victimization (O'Flaherty and Sethi, 2010a; Williams Jr, 2020). Second, an increase in police manpower permits the deployment of additional resources to communities with higher homicide rates. If homicide serves as a particularly salient signal of criminal behavior, racial disparities in homicide rates could shape the allocation of policing resources. Finally, potential deterrence effects of more police may differ across Black and white neighborhoods if differences also exist in social norms or community perceptions of law enforcement legitimacy (Tyler, 2003; Gau and Brunson, 2010; Lovett and Xue, 2018).

Expanded law enforcement presence also raises concerns that policing strategies involving the use of directed patrol may create collateral costs and disproportionate burdens for disadvantaged communities (Weitzer et al., 2008; Bandes et al., 2019). Research finds that while concentrating police at crime hot spots improves public safety, such a strategy has not been effective in making community members feel safer or in improving perceptions of police legitimacy (Ratcliffe et al., 2015; Kochel and Weisburd, 2017). There is likewise evidence that mass enforcement policies have served to widen the net of the criminal justice system (Hagan and Dinovitzer, 1999; Kohler-Hausmann, 2018), leading to an increase in discriminatory practices which have had disproportionate impacts on minority communities (Gelman et al., 2007; Goel et al., 2016; Goncalves and Mello, 2020), including the use of violence in interactions with Black suspects (Fryer Jr, 2019). Indeed, over the life course, about 1 in every 1,000 Black men can expect to be shot by police (Edwards et al., 2019).

To what extent do police create racially disparate costs as well as benefits? Using national data on police employment for a sample of 242 large U.S. cities over a 38-year period, this research provides novel evidence on the racial differences in public safety returns to law enforcement expansion in the United States. We focus on two primary outcomes: homicide victimization and enforcement activity as proxied by various types of arrests.[3] By focusing on the size of a city's police force, we provide historical evidence on a critical policy estimand that is implicated by the "Defund" movement and which, for many years, has been the primary means by which municipal policymakers have invested in public safety. In focusing on police manpower and in keeping with the related literature, we note that we are implicitly holding fixed many additional sources of variation in police effectiveness — including police management styles and training (Mummolo, 2018; Owens et al., 2018; Ba and Rivera, 2019; Nagin and Telep, 2020; Wood et al., 2020) and the composition and quality of the police force (Donohue III and Levitt, 2001; McCrary, 2007; Miller and Segal, 2019; Harvey and Mattia, 2019) — each of which is worthy of independent study.

Given the potential endogeneity of police force size, we use two different instrumental variable strategies commonly employed in the policing literature. First, we predict police force size using variation in the timing of federal block grants provided by the U.S. Department of Justice's Community Oriented Policing Services (COPS) office (Evans and Owens, 2007; Mello, 2019; Weisburst, 2019b). Second, recognizing that cities operate under numerous constraints that make it difficult to get out ahead of crime waves, we follow an approach utilized in Chalfin and McCrary (2018) which argues that a primary driver of endogeneity bias in regressions of crime on police manpower is measurement error in police employment data. Using two distinct measures of police force size from different data sources, we derive estimates of the effect of police manpower on homicide victimization that are robust to measurement error.

[3] We consider fatal encounters between civilians and police officers in an auxiliary analysis.

We find that each additional police officer hired abates between 0.06 and 0.1 homicides with estimates that are strikingly similar across the two estimation strategies. The estimates suggest that investments in police manpower can save a life at a cost of between $1.3 million and $2.2 million while it is common for estimates of the value of a statistical life to exceed $7 million (Viscusi and Aldy, 2003; Chalfin and McCrary, 2018). Although the total reduction in homicide is roughly equal across Black and white victims, the decline in homicide is twice as large for Black victims in per capita terms. Next, we consider the extent to which investments in police manpower expand civilian interactions with the criminal justice system, or create "net widening" effects, focusing on differences by race in the burdens and benefits of enforcement activity. Here, we find that investments in police manpower lead to larger numbers of low-level "quality of life" arrests, with effects that imply a disproportionate burden for Black civilians who are arrested. At the same time, we find that arrests for the most serious offenses *fall* with investments police manpower. On a per capita basis, the decline in index crime arrests that we observe is between 4-6 times *greater* for arrests involving Black suspects. This finding is consistent with the idea that police hiring has the potential to create a "double dividend" (Bratton, 2011; Cook and Ludwig, 2011; Durlauf and Nagin, 2011) for both Black and white Americans by generating reductions in both crime and incarceration for serious offenses.

Critically though, the average effects described above mask important variation in the quality of policing across cities. In cities with relatively large Black populations, the returns to investments in police manpower are smaller and perhaps non-existent for Black civilians. Likewise, larger police forces lead to a greater number of arrests for "quality of life" offenses–in particular for Black civilians– without the reduction index crime arrests that we observe elsewhere. As such, the prospect for investments in police manpower to lead to a socially beneficial "double dividend" are far less compelling in these cities. The pattern of findings provides empirical support for two important propositions. First, given that we observe the largest increases in low-level “quality of life” arrests in the subset of cities that experience the smallest benefits of increases in police manpower, this research suggests that it is the presence of police officers rather than the number of arrests that they make which drives the public safety returns to investments in law enforcement. Second, by documenting that the cities with the largest Black population shares do not share equally in the benefits of policing while disproportionately sustaining the greatest burdens, we provide novel empirical support for the popular narrative that Black communities are simultaneously over- and under-policed (Leovy, 2015).

## 2 Data

Our analysis focuses on 242 large U.S. cities over the 1981-2018 period. The sample is restricted to cities which have populations greater than 50,000 in 1980 and regularly report data to the U.S. Census' Annual Survey of Government (ASG). We focus on municipal police departments serving these cities and on full-time sworn police employment. A detailed explanation of data sources and cleaning can be found in Appendix A3.

Our principal treatment variable is a measure of annual police employment collected as part of the Federal Bureau of Investigation's (FBI) Law Enforcement Officers Killed and Assaulted (LEOKA) series. Our first instrumental variables strategy uses a secondary measure of police employment collected independently by the ASG. A second instrumental variables strategy leverages federal grants for hiring police officers administered by the DOJ COPS office. Given that these grants began in 1994 as part of the Violent Crime Control Act, our analysis using COPS grants covers the period of 1990-2018.

Data on homicides come from the FBI's Supplementary Homicide Reports (SHR) dataset which assembles records of homicides reported from each police agency in the U.S. For each city-year, we aggregate homicides separately by race, focusing on homicides with either a non-Hispanic Black or a non-Hispanic white victim.[4] We exclude homicides committed in prisons or jails as well as felons killed in the commission of a crime as these are likely to fall under the legal definition of justifiable homicide. We also use the SHR data to calculate a homicide clearance rate–the proportion of homicides in which a suspect or perpetrator is identified.[5]

To assess the extent to which a larger police force widens the net of the criminal justice system, we use data on arrests collected by the FBI's Uniform Crime Reports (UCR). For much of the analysis, we group arrests into the FBI's definition of seven major "index crimes" (murder, rape, robbery, aggravated assault, burglary, grand larceny and motor vehicle theft), lower-level "quality of life" offenses (including disorderly conduct, liquor violations, loitering, loitering, and drug possession), and arrests for any other type of offense (see Appendix Table A10 and Appendix Table A11 for a full list of the components of these groups).[6] For each category, we track total arrests as well race-specific arrests.[7]

We supplement our analysis with additional data on city demographics and budgets from the U.S. Census to construct control variables. Demographic data for each analysis include population, resident race, gender, age shares, educational attainment, marital status, and income. Our budget data includes city expenditures, revenue, and tax receipts.

# 3 Econometric Methods

Our empirical strategy is motivated by the following least squares regression:

$$Y_{it}^{j} = \theta S_{it-1} + \gamma^{'} X_{it} + \rho_i + \psi_{st} + \varepsilon_{it} \tag{1}$$

[4]We consider Hispanic victims in an auxiliary analysis.

[5]This measure focuses on preliminary reports and will differ from clearance rates reported directly by police departments which include cases cleared in subsequent years.

[6]Notably included are "uncategorized" arrests. This means our estimates account for any potential improvements in reporting that could shift arrests recorded without a category into another of the arrest categories.

[7]As Hispanic/Latinx victims do not have their own category in the FBI's arrest data, these victims are classified either as white, Black, Asian or other. The "$\beta$/Pop." benchmarks we include for the arrest outcomes adjust white estimates for the combined Non-Hispanic white and Hispanic population in the U.S. Census to account for this uncoded category.

In (1), $Y_{it}^{j}$ is a given outcome of interest measured in city $i$ for individuals of race $j$ in year $t$. Given our central research question regarding the public safety returns to an increase in police manpower, we specify each of our models in levels with $\theta$ reflecting the marginal returns to employment of an additional officer within the policing production function.[8] $S_{it-1}$ is the number of sworn police officers measured in the previous year, a convention that is used in order to minimize endogeneity bias (Levitt, 1996, 2002; Chalfin and McCrary, 2018).[9]

The model conditions on city ($\rho_i$) and interacted state-by-year ($\psi_{st}$) fixed effects. The latter term accounts for annual variation in state-level policies including changes in incarceration levels and sentencing practices, as well as aggregate changes in policing technology. State-by-year fixed effects also account for changes in crime and arrest recording practices which could influence counts in the SHR and UCR data we use, to the extent that these change in states over time. We control for a vector of time-varying covariates, $X_{it}$, which includes a quadratic function of population and detailed demographic data including a city's racial composition, gender composition, age demographics, income, poverty, and the unemployment rate. Our models also account for each city's tax receipts, revenue and expenditures in order to directly study the effects of law enforcement expansion *holding municipal spending constant*. Accordingly $\theta$ represents the effect of hiring one additional police officer relative to the historical opportunity cost of using the funds for an alternative purpose. Our baseline specification weights the data by according to a city's 1980 population. Standard errors are clustered at the city-level.

There are primary two challenges to identifying a causal estimate of $\theta$, the impact of police employment. First, as shown by Chalfin and McCrary (2018), police employment is measured with

[8]Focusing on levels models presents several advantages. First, per capita models and other functional form assumptions do not directly address our main research question concerning the marginal public safety returns associated with hiring an additional officer. Second, the levels model permits greater flexibility in controlling fro the relationship between population and homicide or other key outcomes. Lastly, per capita models at the city-level are not easily translated to race-specific outcomes as covariates like city budget expenditures do not make sense when scaled by race-specific population.

[9]Estimates are extremely similar when a contemporaneous measure is used.

error. If measurement errors are classical, equation (1) will yield an estimate of $\theta$ that is attenuated towards zero–a problem that is likely made worse by the inclusion of covariates and fixed effects.[10] A second concern is that $\theta$ may be biased due to the omission of covariates or simultaneity bias between police hiring and crime (Levitt, 1996; Evans and Owens, 2007).

In order to obtain consistent estimates of $\theta$, we use two different instrumental variables strategies each of which has been employed in the prior literature. First, following Chalfin and McCrary (2018), we explicitly correct for measurement error bias in police force size using a second potentially independent measure of police manpower from the U.S. Census Annual Survey of Governments (ASG IV) as an instrument for the FBI measure of police manpower. As we show in Appendix A1, switching the role that each police measure plays in the IV framework leads to statistically identical estimates, consistent with the proposition that measurement errors are classical. Second, following Evans and Owens (2007), Mello (2019) and Weisburst (2019b), we instrument for police manpower with variation in federal "COPS" grants that were awarded to cities to facilitate police hiring. Previous work demonstrates that the likelihood of an agency receiving a grant in a given year remains plausibly exogenous conditional on covariates and fixed effects. Similar to Evans and Owens (2007), our specification uses the number of police officers eligible for hiring under an awarded grant as the instrumental variable. The model includes additional controls for the size of grant awards for non-hiring purposes and indicators for police department decisions to apply for grants over time (Weisburst, 2019b). Critically, controls for non-hiring grant awards and applications proxy for police department interest in and funding for other types of investments in police operations, including technology improvements.

In addition to estimating different local average treatment effects, each IV strategy has costs and benefits which can be characterized as a trade off between bias and variance. While models

[10] Conditioning on fixed effects removes some of the true signal in $S_{it}$ with the remaining variation left to include a larger share of error.

using COPS grants as an instrument credibly addresses both sources of inconsistency in OLS estimates (i.e., endogeneity *and* measurement error), these models retain only a small amount of the variation in police hiring and are less precisely estimated. Moreover, the commencement of the COPS program in 1994 restricts the study period for these analyses to the 1990-2018 period. On the other hand, while our measurement error instrument generates an extremely strong first stage and uses the full sample of data, the cost is that these models do not leverage a natural experiment to address endogeneity concerns. With respect to the latter point, we note that while concerns about simultaneity bias dominate the literature, similar to a famous result–that measurement error bias may be more important than ability bias in estimating a Mincer equation (Ashenfelter and Krueger, 1994)–simultaneity bias concerns may be less important than measurement error bias in our context. As discussed in Appendix A1, the political science and public administration literatures have detailed a variety of constraints faced by municipal leaders that make strategic police hiring difficult, at least over a one-year time period (Lewis, 1994; Joyce and Mullins, 1991; Poterba and Rueben, 1995; Shadbegian, 1998; Shavell, 1991; Koper, 2004; Rubin, 2016).

In practice, both the measurement error IV model and the COPS IV model lead to substantively similar outcomes which both narrows the scope for simultaneity bias to be a first-order problem and strengthens our confidence in the resulting estimates. Given the support for both identification strategies in the previous literature, we omit further discussion from the main body of the paper and refer readers to Appendix A1 where we provide additional details and evidence of the robustness of these strategies.

# 4 Results

## 4.1 Descriptive Statistics

Table 1 reports summary statistics for each of our key outcomes and control variables, weighted by 1980 population. On average, individuals living in the cities in our sample are 24% non-Hispanic Black, 19% Hispanic and 50% non-Hispanic white. The average city in our sample employs between 363 and 424 police officers per 100,000 residents depending upon the police measure used. This is higher than the national average, approximately 250 per 100,000 residents, but unsurprising given that our sample includes the largest cities in the U.S.

In an average city-year in our data, there are 242 homicide victims, of which 137 (57%) are non-Hispanic Black and 63 (26%) are non-Hispanic white. In per capita terms, Black residents are approximately 3 times as likely to be the victim of a homicide compared to white residents. Black Americans are also disproportionately arrested for both serious index crimes and low-level "quality of life" offenses. Black civilians make up over half of each of these types of arrests, and in per capita terms are arrested at 3 to 4 times the rate of their white counterparts.

## 4.2 Main Estimates

Our primary results are presented in Table 2. For the measurement error model, the $F$-statistic on the excluded instrument is over 500 indicating a very strong first stage relationship between the measures. For the COPS IV, the $F$-statistic on the excluded instrument is 16 which, while smaller, exceeds the critical value for maximal 10% bias as computed by Stock et al. (2002).

Next, we turn to our principal findings. For each outcome, we estimate the effect of a change in police force size separately for Black and white civilians. For each outcome variable, we present two useful benchmarks. First, in order to understand the proportional relationship between each outcome and police force size, we transform each coefficient into an elasticity. Second, because Black civilians make up a comparatively small share (24%) of the population in our sample, we present the estimate as a change per 100,000 residents of a given race. This allows us to comment more directly on the differential benefits and burdens of policing which accrue to Black versus white civilians.

Our first result is that an increase in police manpower reduces homicide victimization, in total and for each racial group. The marginal police officer abates between 0.06 and 0.1 homicides indicating that, on average, there is one life saved per 10-17 police officers hired.[11] In elasticity terms, these estimates imply that a 1% increase in police manpower leads to a $1.1 - 2.5\%$ decrease in Black homicide victimization and a $1.4 - 4.4\%$ decrease in white homicide victimization. On a per capita basis, police force expansion has a larger effect on homicide victimization for Black civilians ($0.006 - 0.012$ homicides per 100,000 population) than for whites ($0.003 - 0.007$ homicides per 100,000 population).[12]

Next, recognizing that police officers typically have broad discretion over whether or not to make arrests (Goldstein, 1963; Linn, 2009; Weisburst, 2017) and their level of proactivity in searching for and identifying criminal activity (Wu and Lum, 2017), we consider different types of arrests as markers of police activity. Using the ASG IV (COPS IV), we estimate that the marginal police officer makes approximately 7.3 (22) arrests for "quality of life" offenses. While approximately 60% of the marginal arrests accrue to white civilians, on a per capita basis, the burden of the additional low-level arrests falls upon Black civilians compared to white civilians. Using the COPS IV, this contrast is particularly apparent as point estimates imply that the burden of low-level arrests is

[11] As we note in Appendix Table A4, the fact that the COPS IV estimates are approximately twice as large as those obtained using the ASG IV model is largely an artifact of the restricted sample period for the COPS estimation strategy. Estimating the ASG model using the 1990-2018 period yields a point estimate for homicide ($-0.09$) that is very close to the estimate using the COPS instrument.

[12] The racial disparity in homicide rates, in per capita terms, is significant at conventional levels for both IV estimators ($p < 0.001$); In Appendix Table A7 and Appendix Table A8 we compute estimates which include more granular race and ethnicity categories where available. In Appendix Figure A5 we consider more granular demographic age-race-gender subgroups; the analysis shows police are considerably more effective at abating male homicides than female homicides.

70% greater among Black civilians than white civilians.[13]

We also consider the effects of police manpower on enforcement for more serious crimes. First, we examine whether a larger police force is able to clear more homicides–a critical metric of police productivity. Neither IV strategy produces any meaningful evidence on homicide clearance rates for victims of either race. Next, we consider the effects of police manpower on index crime arrests. Consistent with recent findings (Owens, 2013), we do not observe an increase in index crime arrests as a function of police manpower. Indeed the evidence suggests that index crime arrests *fall* (by between $-0.97$ and $-1.56$) with each additional police officer employed. Given that reductions in arrests are a function of both police behavior *and* offender behavior, we estimate the effect of police force size on index crimes for reference. Since larger police forces lead to reductions in index crimes, the decline in index crime *arrests* that we observe suggests that larger police forces reduce serious crime primarily through deterrence rather than by arresting and incapacitating additional offenders.

With respect to the racial incidence of index crime arrests, we observe that, relative to population, a larger police force leads to a reduction in index crime arrests that is between 4 and 6 times larger for Black suspects than for white suspects, a difference which is significant at conventional levels ($p < 0.001$). This result suggests that the deterrence value of police might be especially large for this sub-population and that investments in police employment potentially has the attractive quality of reducing both homicide victimization as well as imprisonment rates for this group. As such, despite elevated contact between police and Black civilians, police hiring does not automatically widen the net of the criminal justice system for Black Americans.

In Online Appendix A2, we subject each of the results reported in our main tables to greater scrutiny. We re-estimate the models without population weights, we condition on a number of more

[13] Using the COPS model, in per capita terms, this difference is significant at $\alpha = 0.1$. While the difference is not significant at conventional levels, we note that this test is conservative since, due to arrest data limitations, Hispanic arrestees are classed as white. As research indicates important Hispanic-white disparities with respect to policing outcomes (Sanga, 2009), the white which includes Hispanic arrestees estimate is likely to be larger than the non-Hispanic white estimate.

granular fixed effects and employ a variety of different functional forms. We also consider concerns regarding the reporting of crimes and arrests to the FBI. We re-estimate our models focusing on total arrests to account for the possibility that reductions in low-level arrests could be an artifact of the FBI's "hierarchy rule." In addition, we consider reporting along the extensive margin by considering sub-categories of arrests with zero reported arrests. In all cases, results are substantively similar to our preferred specification and do not suggest that reporting artifacts are a first order problem. Finally, we provide a host of supplemental results including an enhanced discussion of treatment effect heterogeneity with a focus on the the role of age and gender and an auxiliary analysis in which we study the effect of police force size on fatal encounters between police officers and civilians of different races.

### 4.3 Treatment Effect Heterogeneity

Racial differences in perceptions of law enforcement are an enduring feature of policing in the United States (Tuch and Weitzer, 1997). While survey data suggests that Black and white Americans do not differ markedly in their support for particular policing styles, there is large and longstanding gap in trust that civilians of different races have for law enforcement.[14] Racial differences in trust accord with a large body of research which finds that Black and white suspects are, on average, treated differently by individual police officers (Goncalves and Mello, 2020; Fryer Jr, 2019). One of the most salient drivers of the race gap in police behavior is geography, as different styles of policing tend to be applied in communities with different demographic compositions (Goel et al., 2016).

In Table 3 and Table 4, we allow the effect of police manpower to vary according to a city's 1980 Black population share. The sizable homicide reductions–and reductions in index crimes more generally–that are generated by a larger police force do not accrue to the same degree in cities with

[14]See: https://poll.qu.edu/new-york-city/release-detail?ReleaseID=2267.

more concentrated Black populations. We also observe that, in both absolute and per capita terms, the burdens of "quality of life" arrests resulting from police force expansions are especially large for Black civilians in these cities. Critically, the benefits of a reduction in index crime arrests do not accrue to Black civilians in these cities, in contrast with the average effects we observe across the pooled sample. These results indicate that the prospect for police hiring to create a "double dividend"–reducing both crime and serious arrests–does not reflect the reality experienced by Black Americans living in cities with relatively large Black populations.

# 5 Conclusion

This study reports the first estimates of the race-specific impacts of a larger police force. We find that larger police forces disproportionately abate homicides with Black victims. With respect to the prospect for police hiring to widen the net of the criminal justice system by subjecting larger numbers of people to human capital disruptions (Leslie and Pope, 2017; Dobbie et al., 2018) and adverse labor market outcomes (Pager, 2003; Agan and Starr, 2018; Doleac and Hansen, 2020), we report mixed conclusions. On the one hand, we find that larger police forces lead to more low-level "quality of life" arrests, in particular for Black civilians and especially for Blacks civilians who live in cities with a large Black population. On the other hand, our finding that index crime arrests fall with police manpower, and disproportionately fall for Black civilians, is consistent with the idea that police hiring has the potential to create a "double dividend" for society (Bratton, 2011; Cook and Ludwig, 2011; Durlauf and Nagin, 2011) by generating reductions in both crime and incarceration for serious offenses. While arrests for "quality of life" offenses have the potential to accumulate, and may have criminogenic effects either though jail sentences (Gupta et al., 2016; Leslie and Pope, 2017) or peer effects (Stevenson, 2017), the results imply that larger police forces are unlikely to be an important driver of lengthy prison sentences or mass incarceration, for both Black and white civilians.

Our research also shows that the marginal effects of police employment differ substantially across cities with different racial compositions. In cities with relatively large Black populations, the returns to police manpower in reducing homicides and index crime arrests are smaller or non-existent for Black civilians, while Black civilians experience especially large increases in low-level or "quality of life" arrests. These results show that the prospective benefits of larger police forces that we observe, on average, are not universal. These findings are notably inconsistent with at least some components of the theory of "broken windows" policing (Kelling et al., 1982), or the notion that aggressive policing of low-level offenses will either deter or incapacitate more serious crime — a finding which has also been called into question in research on order maintenance policing strategies by Harcourt and Ludwig (2006) and MacDonald et al. (2016). However, the findings are consistent with the idea Black communities are both simultaneously over- and under-policed, a theory that has received a great deal of attention in the public discourse (Leovy, 2015) but which has, to date, received little systematic inquiry in the scholarly literature. The absence of significant public safety returns to labor in cities with larger shares of Black residents, coupled with no evidence of changes in index arrests and clearance rates, suggests that improvements in policing productivity remain possible through technologies known to reduce homicide victimization.

Critically, our findings also highlight important channels that contribute to the "great divide" in policing in America that has been characterized as as the defining generational challenge for contemporary law enforcement Bratton and Anderson (2018). While we find that investments in law enforcement save Black lives, the number of averted homicides (1 per 10-17 officers hired) is modest and might even be zero in cities with large Black populations. Moreover, when they do accrue, abated homicides are also difficult, if not impossible, for the public to observe. In contrast, "quality of life arrests" and their antecedents, street and traffic stops, are considerably more common and are therefore likely to be far more salient for Black Americans, especially those living in cities with large Black populations. As shown by Weisburst (2019a) and Fryer Jr (2019), racial differences in the number of "quality of life" arrests may also be the most important driver of differences in the use of force by police against Black versus white civilians. While information on the use of force by police officers is not collected nationally, if we use the estimate in Weisburst (2019a)–that 2.7 percent of arrests lead to an incident in which force was used by a police officer–then hiring one additional police officer would yield between 3 and 6 use of force incidents per life saved through homicide abatement. The relative magnitudes of the effect of police force size on homicides and arrests likely translate to a difference in salience; individuals are more likely to observe arrest increases that result from police expansion than homicide increases that might result from police contraction.

Our estimates capture the historical opportunity cost of policing, by including controls that hold municipal spending fixed. In this vein, our results suggest that "de-funding" the police could result in more homicides, especially among Black victims. Of course, reducing funding for police could allow increased funding for other alternatives. Indeed an array of high-quality research suggests that crime can, in certain contexts, be reduced through methods other than policing or its by-product, incarceration. Among the many alternatives to police for which there is promising evidence are place-based crime control strategies such as increasing the availability of trees and green space (Branas et al., 2011), restoring vacant lots (Branas et al., 2016, 2018; Moyer et al., 2019), public-private partnerships (Cook and MacDonald, 2011), street lighting (Doleac and Sanders, 2015; Chalfin et al., 2019), and reducing physical disorder (Sampson and Raudenbush, 2001; Keizer et al., 2008). There is also evidence that social service-based strategies such as summer jobs for disadvantaged youth (Heller, 2014; Gelber et al., 2016; Davis and Heller, 2017), cognitive behavioral therapy (Blattman et al., 2017; Heller et al., 2017), mental health treatment (Deza et al., 2020) and local non-profits more generally (Sharkey et al., 2017) can have important crime-reducing effects. While social service interventions are often difficult to scale (Mofiitt, 2006; Ludwig et al., 2011), the increasing number of studies which show that there are ways to reduce crime outside the deterrence channels of the traditional model of Becker (1968) is encouraging.

At the same time, our findings on low-level arrests highlight the potential benefits of changing the priorities of law enforcement. This could occur through changes in policy like the decriminalization of drug possession or via efforts to recruit a larger number of Black or female police officers (Donohue III and Levitt, 2001; McCrary, 2007; West, 2018; Miller and Segal, 2019; Harvey and Mattia, 2019; Ba and Rivera, 2020; Linos and Riesch, 2020). Moreover, there is growing evidence to support the efficacy of de-escalation training (Engel et al., 2020) and procedural justice training (Owens et al., 2018; Nagin and Telep, 2020; Wood et al., 2020), federal oversight of police agencies (Powell et al., 2017; Goh, 2020), and the use of and training in non-lethal weapons (MacDonald et al., 2009; Sousa et al., 2010). There is likewise support for the idea that reforms to police unions may be effective (Dharmapala et al., 2019) especially if unions can be incentivized to "self-regulate," which could take the form of transferring the burden of liability insurance from municipalities to unions (Ramirez et al., 2018; Ba and Rivera, 2019). Finally, police officers tend to be highly responsive to managerial directives (Mummolo, 2018), which suggests that procedural reforms could meaningfully alter officer behavior even holding police force size fixed.

Whether communities should invest less in law enforcement and more in alternative strategies to maintain public safety continues to remain an open question, as such a material change in our society's approach to public safety has yet to be implemented at scale. Our research focuses on one crucial aspect of this current policy debate–the effect of reducing police employment–an outcome which would likely result if proposals to reduce funding for municipal police departments are adopted in the future. This study provides an estimate of the historical trade-offs of investments in law enforcement and, critically, the resulting implications for communities of color.

## References


Agan, A. and S. Starr (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Ashenfelter, O. and A. Krueger (1994). Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review*, 1157–1173.

Ba, B. A. and R. Rivera (2019). The effect of police oversight on crime and allegations of misconduct: Evidence from chicago. *University of Pennsylvania, Institute for Law & Economics Research Paper* (19-42).

Ba, Bocar, K. D. M. J. and R. Rivera (2020). Diversity in policing: The role of officer race and gender in police-civilian interactions in chicago.

Bandes, S. A., M. Pryor, E. M. Kerrison, and P. A. Goff (2019). The mismeasure of terry stops: Assessing the psychological and emotional harms of stop and frisk to individuals and communities. *Behavioral Sciences & the Law 37*(2), 176–194.

Barber, C., D. Azrael, A. Cohen, M. Miller, D. Thymes, D. E. Wang, and D. Hemenway (2016). Homicides by police: comparing counts from the national violent death reporting system, vital statistics, and supplementary homicide reports. *American Journal of Public Health 106*(5), 922–927.

Becker, G. S. (1968). Crime and punishment: An economic approach. *The Journal of Political Economy 76*(2), 169–217.

Blattman, C., J. C. Jamison, and M. Sheridan (2017). Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in liberia. *The American Economic Review 107*(4), 1165–1206.

Braga, A. A., A. V. Papachristos, and D. M. Hureau (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly 31*(4), 633–663.

Branas, C. C., R. A. Cheney, J. M. MacDonald, V. W. Tam, T. D. Jackson, and T. R. Ten Have (2011). A difference-in-differences analysis of health, safety, and greening vacant urban space. *American Journal of Epidemiology 174*(11), 1296–1306.

Branas, C. C., M. C. Kondo, S. M. Murphy, E. C. South, D. Polsky, and J. M. MacDonald (2016). Urban blight remediation as a cost-beneficial solution to firearm violence. *American Journal of Public Health 106*(12), 2158–2164.

Branas, C. C., E. South, M. C. Kondo, B. C. Hohl, P. Bourgois, D. J. Wiebe, and J. M. MacDonald (2018). Citywide cluster randomized trial to restore blighted vacant land and its effects on violence, crime, and fear. *Proceedings of the National Academy of Sciences 115*(12), 2946–2951.

Bratton, W. and B. C. Anderson (2018). Precision policing. *City Journal*.

Bratton, W. J. (2011). Reducing crime through prevention not incarceration. *Criminology & Public Policy 10*(1), 63–68.

Chalfin, A., B. Hansen, J. Lerner, and L. Parker (2019). Reducing crime through environmental design: Evidence from a randomized experiment of street lighting in new york city. Technical report, National Bureau of Economic Research.

Chalfin, A. and J. McCrary (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature 55*(1), 5–48.

Chalfin, A. and J. McCrary (2018). Are us cities underpoliced? theory and evidence. *Review of Economics and Statistics 100*(1), 167–186.

Cook, P. J. and J. Ludwig (2011). *More prisoners versus more crime is the wrong question.* Brookings Institution.

Cook, P. J., J. Ludwig, S. Venkatesh, and A. A. Braga (2007). Underground gun markets. *The Economic Journal 117*(524), F588–F618.

Cook, P. J. and J. MacDonald (2011). Public safety through private action: an economic assessment of bids. *The Economic Journal 121*(552), 445–462.

Davis, J. M. and S. B. Heller (2017). Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. *The Review of Economics and Statistics*, 1–47.

Deza, M., J. C. Maclean, and K. T. Solomon (2020). Local access to mental healthcare and crime. Technical report, National Bureau of Economic Research.

Dharmapala, D., R. H. McAdams, and J. Rappaport (2019). Collective bargaining and police misconduct: Evidence from florida.

Di Tella, R. and E. Schargrodsky (2004). Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *The American Economic Review 94*(1), 115–133.

Dobbie, W., J. Goldin, and C. S. Yang (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *The American Economic Review 108*(2), 201–40.

Doleac, J. L. and B. Hansen (2020). The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden. *Journal of Labor Economics 38*(2), 321–374.

Doleac, J. L. and N. J. Sanders (2015). Under the cover of darkness: How ambient light influences criminal activity. *The Review of Economics and Statistics 97*(5), 1093–1103.

Donohue III, J. J. and S. D. Levitt (2001). The impact of race on policing and arrests. *The Journal of Law and Economics 44*(2), 367–394.

Durlauf, S. N. and D. S. Nagin (2011). Imprisonment and crime: Can both be reduced? *Criminology & Public Policy 10*(1), 13–54.

Edwards, F., M. H. Esposito, and H. Lee (2018). Risk of police-involved death by race/ethnicity and place, united states, 2012–2018. *American Journal of Public Health 108*(9), 1241–1248.

Edwards, F., H. Lee, and M. Esposito (2019). Risk of being killed by police use of force in the united states by age, race–ethnicity, and sex. *Proceedings of the National Academy of Sciences 116*(34), 16793–16798.

Engel, R. S., H. D. McManus, and T. D. Herold (2020). Does de-escalation training work? a systematic review and call for evidence in police use-of-force reform. *Criminology & Public Policy*.

Evans, W. N. and E. G. Owens (2007). Cops and crime. *Journal of Public Economics 91*(1-2), 181–201.

Finch, B. K., A. Beck, D. B. Burghart, R. Johnson, D. Klinger, and K. Thomas (2019). Using crowd-sourced data to explore police-related-deaths in the united states (2000–2017): The case of fatal encounters. *Open Health Data 6*(1).

Fryer Jr, R. G. (2019). An empirical analysis of racial differences in police use of force. *Journal of Political Economy 127*(3), 1210–1261.

Fuller, W. (1987). Measurement error models, new york: John wiley.

Gau, J. M. and R. K. Brunson (2010). Procedural justice and order maintenance policing: A study of inner-city young men's perceptions of police legitimacy. *Justice Quarterly 27*(2), 255–279.

Gelber, A., A. Isen, and J. B. Kessler (2016). The effects of youth employment: Evidence from new york city lotteries. *The Quarterly Journal of Economics 131*(1), 423–460.

Gelman, A., J. Fagan, and A. Kiss (2007). An analysis of the new york city police department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association 102*(479), 813–823.

Goel, S., J. M. Rao, R. Shroff, et al. (2016). Precinct or prejudice? understanding racial disparities in new york city's stop-and-frisk policy. *The Annals of Applied Statistics 10*(1), 365–394.

Goh, L. S. (2020). Going local: Do consent decrees and other forms of federal intervention in municipal police departments reduce police killings? *Justice Quarterly*, 1–30.

Goldstein, H. (1963). Police discretion: The ideal versus the real. *Public Administration Review*, 140–148.

Goncalves, F. and S. Mello (2020). *A Few Bad Apples?: Racial Bias in Policing*, Volume 2020.

Gupta, A., C. Hansman, and E. Frenchman (2016). The heavy costs of high bail: Evidence from judge randomization. *The Journal of Legal Studies 45*(2), 471–505.

Hagan, J. and R. Dinovitzer (1999). Collateral consequences of imprisonment for children, communities, and prisoners. *Crime and Justice 26*, 121–162.

Harcourt, B. E. and J. Ludwig (2006). Broken windows: New evidence from new york city and a five-city social experiment. *The University of Chicago Law Review*, 271–320.

Harvey, A. and T. Mattia (2019). Reducing racial disparities in crime victimization.

Heller, S. B. (2014). Summer jobs reduce violence among disadvantaged youth. *Science 346*(6214), 1219–1223.

Heller, S. B., A. K. Shah, J. Guryan, J. Ludwig, S. Mullainathan, and H. A. Pollack (2017). Thinking, fast and slow? some field experiments to reduce crime and dropout in chicago. *The Quarterly Journal of Economics 132*(1), 1–54.

Joyce, P. G. and D. R. Mullins (1991). The changing fiscal structure of the state and local public sector: The impact of tax and expenditure limitations. *Public Administration Review*, 240–253.

Kaplan, J. (2019a). Jacob kaplan's concatenated files: Uniform crime reporting (ucr) program data: Supplementary homicide reports, 1976-2018. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Kaplan, J. (2019b). Uniform crime reporting program data: Law enforcement officers killed and assaulted (leoka) 1975-2016. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*, 07–08.

Kaplan, J. (2019c). Uniform crime reporting program data: Offenses known and clearances by arrest, 1960-2017.

Keizer, K., S. Lindenberg, and L. Steg (2008). The spreading of disorder. *Science 322*(5908), 1681–1685.

Kelling, G. L., J. Q. Wilson, et al. (1982). Broken windows. *Atlantic Monthly 249*(3), 29–38.

King, W. R., A. Cihan, and J. A. Heinonen (2011). The reliability of police employee counts: Comparing fbi and icma data, 1954–2008. *Journal of Criminal Justice 39*(5), 445–451.

Klick, J. and A. Tabarrok (2005). Using terror alert levels to estimate the effect of police on crime. *The Journal of Law and Economics 48*(1), 267–279.

Knox, D., W. Lowe, and J. Mummolo (2020). Administrative records mask racially biased policing. *American Political Science Review 114*(3), 619–637.

Knox, D. and J. Mummolo (2020). Making inferences about racial disparities in police violence. *Proceedings of the National Academy of Sciences 117*(3), 1261–1262.

Kochel, T. R. and D. Weisburd (2017). Assessing community consequences of implementing hot spots policing in residential areas: Findings from a randomized field trial. *Journal of Experimental Criminology 13*(2), 143–170.

Kohler-Hausmann, I. (2018). *Misdemeanorland: Criminal courts and social control in an age of broken windows policing*. Princeton University Press.

Koper, C. S. (2004). *Hiring and keeping police officers*. US Department of Justice, Office of Justice Programs, National Institute of ....

Leovy, J. (2015). *Ghettoside: A true story of murder in America*. Spiegel & Grau.

Leslie, E. and N. G. Pope (2017). The unintended impact of pretrial detention on case outcomes: Evidence from new york city arraignments. *The Journal of Law and Economics 60*(3), 529–557.

Levitt, S. D. (1996). The effect of prison population size on crime rates: Evidence from prison overcrowding litigation. *The Quarterly Journal of Economics 111*(2), 319–351.

Levitt, S. D. (2002). Using electoral cycles in police hiring to estimate the effects of police on crime: Reply. *The American Economic Review 92*(4), 1244–1250.

Lewis, C. W. (1994). Budgetary balance: The norm, concept, and practice in large us cities. *Public Administration Review*, 515–524.

Linn, E. (2009). *Arrest Decisions: What Works for the Officer?* Number 5. Peter Lang.

Linos, E. and N. Riesch (2020). Thick red tape and the thin blue line: A field study on reducing administrative burden in police recruitment. *Public Administration Review 80*(1), 92–103.

Loftin, C., D. McDowall, K. Curtis, and M. D. Fetzer (2015). The accuracy of supplementary homicide report rates for large us cities. *Homicide Studies 19*(1), 6–27.

Lovett, N. and Y. Xue (2018). Do greater sanctions deter youth crime? evidence from a regression discontinuity design. *Evidence from a Regression Discontinuity Design (October 25, 2018)*.

Ludwig, J., J. R. Kling, and S. Mullainathan (2011). Mechanism experiments and policy evaluations. *Journal of Economic Perspectives 25*(3), 17–38.

MacDonald, J., J. Fagan, and A. Geller (2016). The effects of local police surges on crime and arrests in new york city. *PLoS one 11*(6).

MacDonald, J. M., R. J. Kaminski, and M. R. Smith (2009). The effect of less-lethal weapons on injuries in police use-of-force events. *American Journal of Public Health 99*(12), 2268–2274.

MacDonald, J. M., J. Klick, and B. Grunwald (2016). The effect of private police on crime: evidence from a geographic regression discontinuity design. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 179*(3), 831–846.

McCrary, J. (2002). Using electoral cycles in police hiring to estimate the effect of police on crime: Comment. *The American Economic Review 92*(4), 1236–1243.

McCrary, J. (2007). The effect of court-ordered hiring quotas on the composition and quality of police. *The American Economic Review 97*(1), 318–353.

Mello, S. (2019). More cops, less crime. *Journal of Public Economics 172*, 174–200.

Miller, A. R. and C. Segal (2019). Do female officers improve law enforcement quality? effects on crime reporting and domestic violence. *The Review of Economic Studies 86*(5), 2220–2247.

Mofiitt, R. A. (2006). Forecasting the effects of scaling up social programs: An economics perspective. *Scale-up in education: Ideas in principle 1*, 173.

Moyer, R., J. M. MacDonald, G. Ridgeway, and C. C. Branas (2019). Effect of remediating blighted vacant land on shootings: A citywide cluster randomized trial. *American Journal of Public Health 109*(1), 140–144.

Mummolo, J. (2018). Modern police tactics, police-citizen interactions, and the prospects for reform. *The Journal of Politics 80*(1), 1–15.

Nagin, D. S. and C. W. Telep (2020). Procedural justice and legal compliance: A revisionist perspective. *Criminology & Public Policy 19*(3), 761–786.

Owens, E., D. Weisburd, K. L. Amendola, and G. P. Alpert (2018). Can you build a better cop? experimental evidence on supervision, training, and policing in the community. *Criminology & Public Policy 17*(1), 41–87.

Owens, E. G. (2013). Cops and cuffs. lessons from the economics of crime: What reduces offending.

Ozkan, T., J. L. Worrall, and H. Zettler (2018). Validating media-driven and crowdsourced police shooting data: a research note. *Journal of Crime and Justice 41*(3), 334–345.

O'Flaherty, B. and R. Sethi (2010a). Homicide in black and white. *Journal of Urban Economics 68*(3), 215–230.

O'Flaherty, B. and R. Sethi (2010b). The racial geography of street vice. *Journal of Urban Economics 67*(3), 270–286.

Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology 108*(5), 937–975.

Poterba, J. M. and K. S. Rueben (1995). The effect of property-tax limits on wages and employment in the local public sector. *The American Economic Review 85*(2), 384–389.

Powell, Z. A., M. B. Meitl, and J. L. Worrall (2017). Police consent decrees and section 1983 civil rights litigation. *Criminology & Public Policy 16*(2), 575–605.

Ramirez, D., M. Wraight, L. Kilmister, and C. Perkins (2018). Policing the police: Could mandatory professional liability insurance for officers provide a new accountability model. *American Journal of Criminal Law 45*, 407.

Ramirez, O. A., C. B. Moss2, and W. G. Boggess2 (1994). Estimation and use of the inverse hyperbolic sine transformation to model non-normal correlated random variables. *Journal of Applied Statistics 21*(4), 289–304.

Ratcliffe, J. H., E. R. Groff, E. T. Sorg, and C. P. Haberman (2015). Citizens' reactions to hot spots policing: impacts on perceptions of crime, disorder, safety and police. *Journal of Experimental Criminology 11*(3), 393–417.

Rubin, I. S. (2016). *The politics of public budgeting: Getting and spending, borrowing and balancing.* CQ Press.

Sampson, R. J. and S. W. Raudenbush (2001). *Disorder in urban neighborhoods: Does it lead to crime.* US Department of Justice, Office of Justice Programs, National Institute of ....

Sampson, R. J., W. J. Wilson, J. Hagan, and R. D. Peterson (1995). Toward a theory of race, crime, and urban inequality. *1995*, 37–54.

Sanga, S. (2009). Reconsidering racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy 117*(6), 1155–1159.

Shadbegian, R. J. (1998). Do tax and expenditure limitations affect local government budgets? evidence from panel data. *Public Finance Review 26*(2), 118–136.

Sharkey, P., G. Torrats-Espinosa, and D. Takyar (2017). Community and the crime decline: The causal effect of local nonprofits on violent crime. *The American Sociological Review 82*(6), 1214–1240.

Shavell, S. (1991). Specific versus general enforcement of law. *Journal of Political Economy 99*(5), 1088–1108.

Sherman, L. W. and D. Weisburd (1995). General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial. *Justice Quarterly 12*(4), 625–648.

Sousa, W., J. Ready, and M. Ault (2010). The impact of tasers on police use-of-force decisions: Findings from a randomized field-training experiment. *Journal of Experimental Criminology 6*(1), 35–55.

Stevenson, M. (2017). Breaking bad: Mechanisms of social influence and the path to criminality in juvenile jails. *The Review of Economics and Statistics 99*(5), 824–838.

Stock, J. H., J. H. Wright, and M. Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics 20*(4), 518–529.

Tuch, S. A. and R. Weitzer (1997). Trends: Racial differences in attitudes toward the police. *The Public Opinion Quarterly 61*(4), 642–663.

Tyler, T. R. (2003). Procedural justice, legitimacy, and the effective rule of law. *Crime and Justice 30*, 283–357.

Viscusi, W. K. and J. E. Aldy (2003). The value of a statistical life: a critical review of market estimates throughout the world. *Journal of Risk and Uncertainty 27*(1), 5–76.

Weisburd, S. (2016). Police presence, rapid response rates, and crime prevention. *The Review of Economics and Statistics*, 1–45.

Weisburst, E. (2017). Whose help is on the way? the importance of individual police officers in law enforcement outcomes. *Working Paper*.

Weisburst, E. K. (2019a). Police use of force as an extension of arrests: Examining disparities across civilian and officer race. In *The American Economic Review*, Volume 109, pp. 152–56.

Weisburst, E. K. (2019b). Safety in police numbers: Evidence of police effectiveness from federal cops grant applications. *American Law and Economics Review 21*(1), 81–109.

Weitzer, R., S. A. Tuch, and W. G. Skogan (2008). Police–community relations in a majority-black city. *Journal of Research in Crime and Delinquency 45*(4), 398–428.

West, J. (2018). Racial bias in police investigations.

Williams Jr, M. C. (2020). Gun violence in black and white: Evidence from policy reform in missouri. Technical report, Working Paper, MIT.

Wood, G., T. R. Tyler, and A. V. Papachristos (2020). Procedural justice training reduces police use of force and complaints against officers. *Proceedings of the National Academy of Sciences 117*(18), 9815–9821.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data mit press. *Cambridge, MA 108*.

Wu, X. and C. Lum (2017). Measuring the spatial and temporal patterns of police proactivity. *Journal of Quantitative Criminology 33*(4), 915–934.

Table 1: Summary Statistics

| | Mean | S.D. | | Mean | S.D. |
|---|---|---|---|---|---|
| **Outcomes** | | | **Covariates** | | |
| Homicide Victims | 242.51 | (397.50) | Population | 1593676 | ( 2402359) |
| Black | 137.35 | (214.74) | Total Government Expenditure | 12781422 | ( 29128508) |
| White | 63.38 | (143.28) | Total Government Revenue | 12775643 | ( 28812641) |
| Homicide Clearance Rate | 66.12 | ( 22.52) | Total Taxes | 5011031 | ( 12037901) |
| Black | 63.51 | ( 24.93) | % Black | 24.23 | ( 18.23) |
| White | 70.24 | ( 23.59) | % White | 48.57 | ( 19.75) |
| Index Crime Arrests | 16419 | ( 26176) | % Hispanic | 18.98 | ( 16.99) |
| Black | 9100 | ( 15555) | % Male | 48.26 | ( 1.29) |
| White | 7058 | ( 11114) | % Age <14 | 20.37 | ( 2.90) |
| Quality of Life Arrests | 58393 | ( 132575) | % Age 15-24 | 16.00 | ( 2.76) |
| Black | 29960 | ( 71569) | % Age 25-44 | 31.20 | ( 3.15) |
| White | 27752 | ( 61415) | % Age >45 | 32.43 | ( 4.30) |
| Index Crimes | 93928 | ( 145967) | % Female Head of Household | 16.34 | ( 4.58) |
| | | | % Never Married | 36.96 | ( 7.09) |
| | | | % Education < High School | 24.46 | ( 8.94) |
| | | | Unemployment Rate | 8.68 | ( 3.08) |
| | | | Poverty Rate | 34.23 | ( 21.95) |
| | | | Median Household Income | 36315 | ( 7750) |

| | Mean | S.D. | | N |
|---|---|---|---|---|
| **Policing** | | | **Sample Counts** | |
| UCR Employment | 5831 | ( 10288) | Number of Cities | 242 |
| ASG Employment | 6647 | ( 12447) | N: ASG Models | 9438 |
| *COPS Grants (Per Grant)* | | | N: COPS Models | 7018 |
| Eligible Hires | 143.05 | ( 346.68) | Number of Hiring Grants | 1125 |
| Hiring Grant Award | 21035428 | ( 49469310) | Number of Non-Hiring Grants | 1460 |
| Non-Hiring Grant Award | 6622148 | ( 26744427) | | |

Note: Summary statistics are weighted by population of each city in 1980. Civilians Shot by Police are available for 2010-2018. COPS IV Models cover the period 1990-2018, ASG IV models cover the period of 1981-2018.

Table 2: Marginal Impact of Police Employment

| | A. ASG IV | | | | | | B. COPS IV | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | S.E. | Elasticity | $\beta$/Pop. | Mean | N | Coeff. | S.E. | Elasticity | $\beta$/Pop. | Mean | N |
| **First Stage** | | | | | | | | | | | | |
| Police Employment | 0.961*** | ( 0.041) | - | - | 6047.0 | 8645 | 3.200*** | ( 0.797) | - | - | 6390.7 | 6623 |
| | *(F-Test = 559.17)* | | | | | | *(F-Test = 16.13)* | | | | | |
| **Homicides** | | | | | | | | | | | | |
| Victims | -0.058*** | ( 0.004) | -1.42 | -0.003 | 249.0 | 8554 | -0.102*** | ( 0.010) | -2.95 | -0.006 | 223.3 | 6531 |
| Black | -0.026*** | ( 0.003) | -1.13 | -0.006 | 140.3 | 8524 | -0.050*** | ( 0.004) | -2.48 | -0.012 | 130.0 | 6503 |
| White | -0.016*** | ( 0.002) | -1.39 | -0.003 | 65.5 | 8503 | -0.044*** | ( 0.001) | -4.43 | -0.008 | 59.2 | 6490 |
| Clearance Rate | 0.001 | ( 0.001) | 0.06 | - | 65.2 | 7676 | 0.001 | ( 0.001) | 0.15 | - | 60.4 | 5767 |
| Black | 0.001 | ( 0.001) | 0.08 | - | 62.6 | 6067 | 0.001 | ( 0.001) | 0.18 | - | 56.8 | 4600 |
| White | -0.001 | ( 0.001) | -0.06 | - | 69.5 | 7046 | 0.000 | ( 0.002) | 0.03 | - | 66.4 | 5224 |
| **Arrests** | | | | | | | | | | | | |
| Quality of Life | 7.32*** | ( 0.88) | 0.55 | 0.54 | 60244 | 7804 | 21.88*** | ( 5.00) | 1.92 | 1.73 | 49908 | 5839 |
| Black | 2.28*** | ( 0.53) | 0.34 | 0.70 | 30896 | 7768 | 8.10*** | ( 1.60) | 1.43 | 2.77 | 24807 | 5831 |
| White | 5.10*** | ( 0.48) | 0.81 | 0.56 | 28827 | 7779 | 13.95*** | ( 3.42) | 2.48 | 1.65 | 24674 | 5818 |
| Index | -0.97*** | ( 0.28) | -0.27 | -0.07 | 16349 | 7797 | -1.56*** | ( 0.32) | -0.51 | -0.12 | 13364 | 5834 |
| Black | -0.69*** | ( 0.20) | -0.35 | -0.21 | 8928 | 7755 | -1.11*** | ( 0.19) | -0.70 | -0.38 | 7006 | 5810 |
| White | -0.45*** | ( 0.09) | -0.28 | -0.05 | 7212 | 7772 | -0.53*** | ( 0.16) | -0.38 | -0.06 | 6135 | 5813 |
| **Index Crimes** | -17.97*** | ( 1.43) | -1.12 | -1.08 | 96892 | 8645 | -23.35*** | ( 1.82) | -1.79 | -1.38 | 83209 | 6623 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Panel A covers 1981-2018 and Panel B covers 1990-2018. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. In Panel A, the instrument is police employment from the U.S. Census. In Panel B the instrument is the number of eligible hires awarded through a COPS Hiring grant. Models include covariates in Table 1; Panel B also controls for non-hiring grant award size and whether a city applied for a hiring or non-hiring grant (lagged). "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). F.B.I. UCR data on arrests does not include sub-categories for Hispanic residents; as a result, white population share includes Hispanic residents for these outcomes in calculating the "$\beta$/Pop." measure. All estimates pass a Bonferroni multiple hypothesis correction of 20.

Table 3: Results by City Racial Composition, ASG IV

| **ASG Employment IV** | Coeff. | S.E. | Elasticity | $\beta$/Pop. | Mean | N |
|---|---|---|---|---|---|---|
| (1) **% Black Residents - Bottom Quartile** | | | | | | |
| *(25th Percentile= 3.05%, First Stage F-Test = 34.27)* | | | | | | |
| Homicide Victims | 0.001 | ( 0.017) | 0.04 | 0.001 | 5.0 | 1834 |
| Black | 0.018** | ( 0.009) | 6.57 | 0.501 | 0.6 | 1830 |
| White | -0.005 | ( 0.007) | -0.42 | -0.007 | 2.6 | 1821 |
| Quality of Life Arrests | 2.90 | ( 7.76) | 0.21 | 2.36 | 2856 | 1707 |
| Black | -0.37 | ( 1.61) | -0.21 | -10.10 | 359 | 1676 |
| White | 4.08 | ( 6.25) | 0.35 | 3.85 | 2363 | 1705 |
| Index Arrests | 4.54 | ( 2.88) | 0.67 | 3.68 | 1383 | 1690 |
| Black | 1.08 | ( 0.97) | 0.99 | 29.38 | 225 | 1664 |
| White | 3.38 | ( 2.43) | 0.63 | 3.19 | 1104 | 1688 |
| Index Crimes | -14.24 | ( 12.74) | -0.47 | -11.51 | 6360 | 1850 |
| (2) **% Black Residents - Interquartile Range** | | | | | | |
| *(50th Percentile= 11.96%, First Stage F-Test = 4407.95)* | | | | | | |
| Homicide Victims | -0.066*** | ( 0.003) | -1.83 | -0.003 | 298.2 | 4204 |
| Black | -0.030*** | ( 0.001) | -1.67 | -0.006 | 147.4 | 4184 |
| White | -0.019*** | ( 0.002) | -1.64 | -0.002 | 88.9 | 4180 |
| Quality of Life Arrests | 6.00*** | ( 0.88) | 0.41 | 0.33 | 89220 | 3806 |
| Black | 1.25** | ( 0.56) | 0.17 | 0.35 | 44223 | 3794 |
| White | 4.88*** | ( 0.43) | 0.67 | 0.38 | 44058 | 3799 |
| Index Arrests | -1.59*** | ( 0.12) | -0.45 | -0.09 | 21583 | 3790 |
| Black | -1.10*** | ( 0.07) | -0.63 | -0.31 | 10675 | 3783 |
| White | -0.67*** | ( 0.06) | -0.38 | -0.05 | 10577 | 3784 |
| Index Crimes | -19.79*** | ( 0.70) | -1.33 | -0.87 | 122698 | 4267 |
| (3) **% Black Residents - Top Quartile** | | | | | | |
| *(75th Percentile= 27.26%, First Stage F-Test = 462.50)* | | | | | | |
| Homicide Victims | 0.012 | ( 0.029) | 0.13 | 0.002 | 154.8 | 1898 |
| Black | 0.021 | ( 0.023) | 0.35 | 0.009 | 106.8 | 1894 |
| White | -0.002 | ( 0.005) | -0.17 | -0.001 | 25.4 | 1884 |
| Quality of Life Arrests | 9.51*** | ( 3.07) | 0.81 | 1.45 | 21405 | 1674 |
| Black | 6.76*** | ( 1.80) | 0.99 | 2.59 | 12399 | 1674 |
| White | 2.95** | ( 1.23) | 0.59 | 0.81 | 8985 | 1654 |
| Index Arrests | 0.75 | ( 1.23) | 0.20 | 0.12 | 7011 | 1689 |
| Black | 0.90 | ( 0.94) | 0.34 | 0.34 | 4769 | 1677 |
| White | -0.09 | ( 0.31) | -0.08 | -0.03 | 2196 | 1674 |
| Index Crimes | -1.88 | ( 5.27) | -0.06 | -0.30 | 52186 | 1921 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Quartiles of cities by racial composition are created using Black population share in 1980. The sample covers 1981-2018. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files; the instrument is police employment from the U.S. Census. Models include covariates in Table 1. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). F.B.I. UCR data on arrests does not include sub-categories for Hispanic residents; as a result, white population share includes Hispanic residents for these outcomes in calculating the "$\beta$/Pop." measure. All estimates pass a Bonferroni multiple hypothesis correction of 20, except for "Black Homicide Victims" (Specification 1), "Black Quality of Life Arrests" (Specification 2), and "White Quality of Life Arrests" (Specification 3).

Table 4: Results by City Racial Composition, COPS IV

| **COPS Eligible Hires IV** | Coeff. | S.E. | Elasticity | $\beta$/Pop. | Mean | N |
|---|---|---|---|---|---|---|
| (1) **% Black Residents - Bottom Quartile** | | | | | | |
| *(25th Percentile= 3.05%, First Stage F-Test = 1.78)* | | | | | | |
| Homicide Victims | -0.414* | ( 0.243) | -19.01 | -0.324 | 4.7 | 1398 |
| Black | -0.044 | ( 0.056) | -15.01 | -1.061 | 0.6 | 1394 |
| White | -0.127 | ( 0.142) | -11.79 | -0.172 | 2.3 | 1390 |
| Quality of Life Arrests | 66.87 | ( 53.69) | 4.51 | 52.60 | 3120 | 1283 |
| Black | 8.07 | ( 8.42) | 4.01 | 190.91 | 424 | 1280 |
| White | 62.20 | ( 46.68) | 5.17 | 58.25 | 2530 | 1281 |
| Index Arrests | 18.91 | ( 18.50) | 2.94 | 14.81 | 1357 | 1270 |
| Black | 0.62 | ( 3.10) | 0.55 | 14.40 | 238 | 1261 |
| White | 19.28 | ( 16.44) | 3.85 | 18.01 | 1057 | 1268 |
| Index Crimes | -91.77 | ( 79.56) | -3.29 | -71.87 | 6015 | 1414 |
| (2) **% Black Residents - Interquartile Range** | | | | | | |
| *(50th Percentile= 11.96%, First Stage F-Test = 22.33)* | | | | | | |
| Homicide Victims | -0.102*** | ( 0.010) | -3.57 | -0.004 | 253.6 | 3205 |
| Black | -0.049*** | ( 0.004) | -3.45 | -0.010 | 127.5 | 3188 |
| White | -0.044*** | ( 0.001) | -4.52 | -0.005 | 79.2 | 3186 |
| Quality of Life Arrests | 19.39*** | ( 5.50) | 1.53 | 1.15 | 73583 | 2859 |
| Black | 7.06*** | ( 1.78) | 1.17 | 2.24 | 35056 | 2854 |
| White | 12.52*** | ( 3.71) | 1.93 | 1.06 | 37772 | 2854 |
| Index Arrests | -2.04*** | ( 0.28) | -0.67 | -0.12 | 17667 | 2843 |
| Black | -1.35*** | ( 0.13) | -0.95 | -0.43 | 8297 | 2839 |
| White | -0.77*** | ( 0.16) | -0.49 | -0.07 | 9012 | 2839 |
| Index Crimes | -23.48*** | ( 1.89) | -2.04 | -1.01 | 101305 | 3269 |
| (3) **% Black Residents - Top Quartile** | | | | | | |
| *(75th Percentile= 27.26%, First Stage F-Test = 18.26)* | | | | | | |
| Homicide Victims | -0.098*** | ( 0.036) | -1.21 | -0.016 | 145.6 | 1453 |
| Black | -0.024 | ( 0.024) | -0.43 | -0.010 | 101.8 | 1449 |
| White | -0.043*** | ( 0.007) | -3.77 | -0.022 | 20.4 | 1439 |
| Quality of Life Arrests | 32.99** | ( 14.59) | 2.86 | 4.93 | 22018 | 1245 |
| Black | 19.23* | ( 10.93) | 2.80 | 7.29 | 13079 | 1245 |
| White | 15.31*** | ( 4.27) | 3.24 | 4.16 | 8937 | 1228 |
| Index Arrests | 7.99*** | ( 1.71) | 2.37 | 1.20 | 6395 | 1260 |
| Black | 5.34*** | ( 1.42) | 2.36 | 2.03 | 4304 | 1250 |
| White | 2.52*** | ( 0.37) | 2.35 | 0.68 | 2040 | 1248 |
| Index Crimes | 13.79 | ( 9.44) | 0.51 | 2.21 | 48483 | 1476 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Quartiles of cities by racial composition are created using Black population share in 1980. The sample covers 1990-2018. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. The instrument is the number of eligible hires awarded through a COPS Hiring grant. Models include covariates in Table 1; and also controls for non-hiring grant award size and whether a city applied for a hiring or non-hiring grant (lagged). "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). F.B.I. UCR data on arrests does not include sub-categories for Hispanic residents; as a result, white population share includes Hispanic residents for these outcomes in calculating the "$\beta$/Pop." measure. All estimates pass a Bonferroni multiple hypothesis correction of 20, except for "Homicide Victims" (Specifications 1 and 3), "Quality of Life Arrests" (Specification 3), and "Black Quality of Life Arrests" (Specification 3).

# ONLINE APPENDIX

## A1 Identification Strategy

Our empirical strategy is motivated by the following least squares regression:

$$Y_{it} = \theta S_{it-1} + \gamma' X_{it} + \rho_i + \psi_{st} + \varepsilon_{it}$$

In this regression, $Y_{it}$ is a given outcome of interest measured in city $i$ in year $t$. In keeping with the extant literature, $S_{it-1}$ is the number of sworn police officers measured in the previous year (Levitt, 1996, 2002; Chalfin and McCrary, 2018). Recognizing that this regression may be compromised by either endogeneity or measurement errors in the right-hand side variable, we pursue two different instrumental variables strategies in order to obtain a plausibly consistent estimate of $\theta$. We describe each of the two strategies in this appendix.

### A1.1 Measurement Error Models

As Chalfin and McCrary (2018) show and as has been suggested indirectly by King et al. (2011), police force size in U.S. cities is measured with error in the available administrative data. We demonstrate this empirically using two measures of police manpower which are both available annually in a large number of U.S. cities. The first measure, which can be found in the Law Enforcement Officers Killed or Assaulted (LEOKA) data collected by the Federal Bureau of Investigation's Uniform Crime Reporting program is the mainstay of the empirical literature that studies police manpower or uses police manpower as a control variable. These data contain a point-in-time measure of the number of sworn police employees in each year, as of October 31st. A second measure of police manpower is available in the U.S. Census Annual Survey of Government Employment (ASG) which collects data on municipal employees. As with the UCR system, the ASG reports a point-in-time measure of police, reporting the number of sworn officers employed as of March 31st of a given year (for 1997-2018 the reference date is June 30th).

Following the approach of Chalfin and McCrary (2018), we begin by demonstrating that while the two available measures of police–one from the FBI's Uniform Crime Reports and the other from the U.S. Census' Annual Survey of Government Employment–align well when plotting the raw data, there are important differences between the two measures once city and state-by-year fixed effects and covariates are netted out. We present this analysis in Appendix Figure A1.

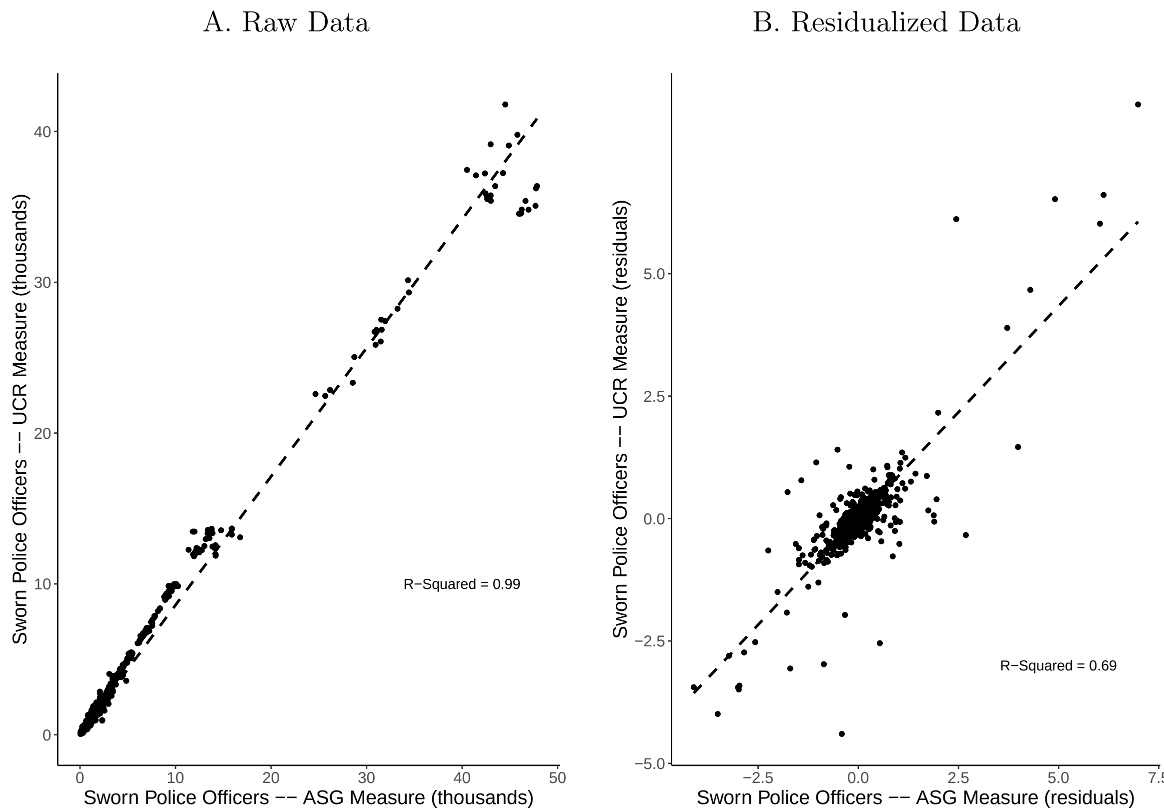
In the figure, Panel A presents a scatterplot of the raw measures; Panel B presents a scatterplot of the two measures, residualized using the covariates and fixed effects described in (1). The fact that the two measures are no longer as well aligned conditional on covariates provides evidence that there may be important errors in the official FBI UCR measure of police. It likewise implies that $\theta$, estimated using equation (1), may be biased as a result of measurement error.

In the presence of two potentially independent measures of the same quantity, the standard solution to the measurement error problem is to instrument one measure with the other, retaining variation that is common to both measures. As is shown by Fuller (1987), such an IV framework allows for a consistent estimate of the parameter of interest subject to the assumption that the measures are independent. To motivate this property of the classical measurement error model, suppose that the two observed series on police force size ($S_{it}$ and $Z_{it}$) are related to the true measure as:

$$S_{it} = S^*_{it} + u_{it} \tag{2}$$

$$Z_{it} = S^*_{it} + v_{it} \tag{3}$$

Figure A1: Two Measures of Police Force Size




Note: Panel A plots the UCR measure of police force size ($y$-axis) against the U.S. Census measure of police force size ($x$-axis). In Panel B, both measures are residualized to account for city and state-by-year fixed effects and covariates.

Further suppose that the outcome of interest, $Y_{it}$, is related to police force size as:

$$Y_{it} = \theta S^*_{it} + \gamma^{'} X_{it} + \varepsilon_{it} \tag{4}$$

Here, $S_{it}$ is the UCR measure of police in a given city and year, $Z_{it}$ is the ASG measure of police, $S^*_{it}$ is the "true" number of sworn police officers or the "signal" and $X_{it}$ are other covariates measured without error. For notational simplicity, we are omitting the fixed effects terms. The error terms, $u_{it}$ and $v_{it}$, are mean zero measurement errors that are mutually uncorrelated and are likewise uncorrelated with $\varepsilon_{it}$, $S^*_{it}$ and $X_{it}$ and $\varepsilon_{it}$.

A famous result from the econometric literature on measurement errors (see, for example, Wooldridge (2002), Section 4.4.2) relates the probability limit of the least squares regression estimate of $\theta$, under the assumptions of the classical measurement error model:

$$plim_{n\rightarrow\infty}\ \hat{\theta}_{OLS} = \theta\ \times\ \frac{\sigma^2_*(1-R^2)}{\sigma^2_*(1-R^2)+\sigma^2_u} \tag{5}$$

In (5), $\sigma^2_u$ is the variance of the error term in (2), and $R^2$ is the population $R$-squared from a regression of the signal, $S^*_{it}$, on $X_{it}$. This formula includes two important ideas. First, since $\sigma^2_u > 0$, a least squares estimate of $\theta$ will be too small in magnitude. Second, while it is a staple of empirical work to confirm that a regression estimate is robust to the inclusion of various control variables, equation (5) indicates that the cure of additional covariates may be worse than the disease of omitted variables bias. Adding more controls increases the $R^2$, exacerbating any attenuation bias.

Next, assume that $X_{it}$ is measured without error and that $S_{it}$ and $Z_{it}$ are residualized to remove shared variation with $X_{it}$. In that case, under the classical measurement error model, the probability limit on the coefficient on $Z_{it}$ in a regression of $S_{it}$ on $Z_{it}$ and $X_{it}$ is given by:

$$\frac{cov(\tilde{S},\tilde{Z})}{var(\tilde{Z})} = \frac{cov(\tilde{S^*}+\tilde{u},\tilde{S^*}+\tilde{v})}{var(\tilde{Z})} = \frac{var(\tilde{S^*})}{var(\tilde{Z})} \equiv \pi \qquad (6)$$

This implies that the ratio of the least squares estimate of the police elasticity of crime, relative to the estimate of $\pi$, is consistent for $\theta$, suggesting a role for an instrument.

Table A1: Test of the Equality of Forward and Reflected IV Estimates

| | (1) Forward | (2) Reflected | (3) $p$-value |
|---|---|---|---|
| Homicide victims | -0.0583 | -0.0642 | 0.35 |
| Black | -0.0261 | -0.0274 | 0.74 |
| White | -0.0159 | -0.0107 | 0.02 |
| Low-Level Arrests | 6.9197 | 6.1052 | 0.48 |
| Black | 2.0466 | 0.9354 | 0.15 |
| White | 4.9489 | 5.4349 | 0.35 |
| Index Crime Arrests | -0.9722 | -0.8804 | 0.81 |
| Black | -0.6881 | -0.6470 | 0.88 |
| White | -0.4466 | -0.4028 | 0.73 |
| Intermediate Arrests | 3.8724 | 3.9789 | 0.84 |
| Black | 1.8733 | 1.9808 | 0.69 |
| White | 1.7096 | 1.7007 | 0.98 |
| Clearance Rate | 0.0006 | 0.0000 | 0.56 |
| Black | 0.0008 | 0.0004 | 0.72 |
| White | -0.0007 | -0.0013 | 0.63 |
| Index crimes | -17.9668 | -20.6628 | 0.15 |

Note: Table reports coefficients from the "forward" and "reflected" IV regressions in which a given measure of police force size is instrumented using an alternative measure of police force size. In the forward specification, the UCR measure of police is the endogenous regressor and the U.S. Census measure of police is the instrument. The roles are reversed in the reflected specification. In the third column, we report the $p$-value on a test of the equality of the forward and reflected coefficients.

Finally, we need to consider the extent to which the assumptions of the classical measurement error model hold in practice. As noted by Chalfin and McCrary (2018), the classical measurement error assumes that $S$ and $Z$ are independent and mean zero but does not prescribe a precise role for $S$ and $Z$ in the instrumental variables setup. That is, under the classical measurement error model, it is *a priori* unclear which measure should play the role of the instrumental variable and which measure should play the role of the endogenous covariate in the IV setup. More formally, $\frac{cov(Z,Y)}{cov(Z,S)}$ will, in expectation, equal $\frac{cov(S,Y)}{cov(S,Z)}$. This insight suggests that an omnibus test of the classical measurement error model is available by empirically testing the equality of $\theta$ from an IV regression in which $S$ is instrumented using $Z$ and $\theta$ from an IV regression in which $Z$ is instrumented using $S$. To the extent that these estimates are significantly different from one another, at least one of the assumptions of the classical measurement error must fail to hold–see Chalfin and McCrary (2018) for a detailed motivation of this feature of the classical measurement error model. We can test this proposition formally by stacking the IV orthogonality conditions for the "'forward" and "reflected" IV models in a broader set of moments:

$$g_i(\beta) = \begin{pmatrix} Z_{it}(Y_{it} - \theta_1 S_{it} - \gamma_1^* X_{it} \\ X_{it}(Y_{it} - \theta_1 S_{it} - \gamma_1^* X_{it} \\ S_{it}(Y_{it} - \theta_2 Z_{it} - \gamma_2^* X_{it} \\ X_{it}(Y_{it} - \theta_2 Z_{it} - \gamma_2^* X_{it} \end{pmatrix} \tag{7}$$

We then test the pooling restriction that $\theta_1 = \theta_2$. The results of this exercise are available in Appendix Table A1 which, for each of our primary outcomes, reports the forward and reflected IV estimates as well as the $p$-value on the equality between the coefficients.[15]

With respect to our most central outcome — homicide victimization by race — there is little evidence against the classical measurement error model as the forward and reflected IV estimates are extraordinarily similar. With only a single exception among 16 tests, we fail to reject the null hypothesis that $\theta_1 = \theta_2$. As such, the IV estimates presented in Table 2 in which we instrument for the UCR measure of police manpower using the U.S. Census measure are expected to be consistent subject to selection assumptions.

[15]This test is available as Hansen's $J$-test of overidentifying restrictions. In practice, this test is also available by stacking the equations and estimating the interaction term between the instrument and the sample.

## A1.2 COPS Eligible Hires Instrument

### A1.2.1 Background on COPS Grants

The Community Oriented Policing Services (COPS) office of the Department of Justice was established under the Violent Crime Control Act of 1994 with the goal of distributing funding for local police departments to improve operations and increase police hiring. Approximately half of COPS funding has been distributed through hiring grants, which have retained the same basic features over time. These three year grants require that police departments not use this funding to supplant funds for existing officers and that departments match a portion of the funds distributed.[16] Non-hiring grants have supported investments police technology, targeted crime initiatives, and community policing programs.

Appendix Figure A2.A displays the number of hiring and non-hiring grants distributed in each year within our sample of large police departments in the U.S. Hiring grants have not been evenly distributed over time; funding declined in the early 2000s amid concerns that the funds were being used to supplant police department budgets for existing hires. However, following the financial crisis in 2008, funding for this program was increased as a way of providing stimulus funds to local governments and to avoid large cuts to police forces. Appendix Figure A2.B shows that funding for hiring grant programs has exceeded funding for non-hiring grants in each year, with a large $600 million spike in 2009.

Each hiring grant designates a number of "eligible hires." Appendix Figure A2.C shows the total eligible hires granted in each year within our sample of large cities. These grants are capable of providing meaningful shocks to the size of police departments, as the average department in our sample has 740 officers (5830 officers when weighted by population) and the average hiring grant awards 23.5 officers (143 when weighted by population).

Law enforcement agencies apply for grants by submitting short narrative applications that outline plans for using funds. Applications are then reviewed by the COPS office and awarded according to fiscal need, application narrative and other office funding constraints. In later years of the grant program, COPS scored applications and weighted scores based on fiscal need (30-75%), local crime conditions (20-35%), and community policing objectives (15-50%). The COPS office faces the additional allocation constraint that at least 0.5% of funds must go to each state and 50% of funding must go to departments serving cities with fewer than 150,000 residents during each grant cycle. While local crime conditions are a small factor in the allocation process, prior work has shown that conditional on fixed effects and city-level covariates, grant awards do not appear to be endogenous to changes in crime rates (Evans and Owens, 2007; Weisburst, 2019b).

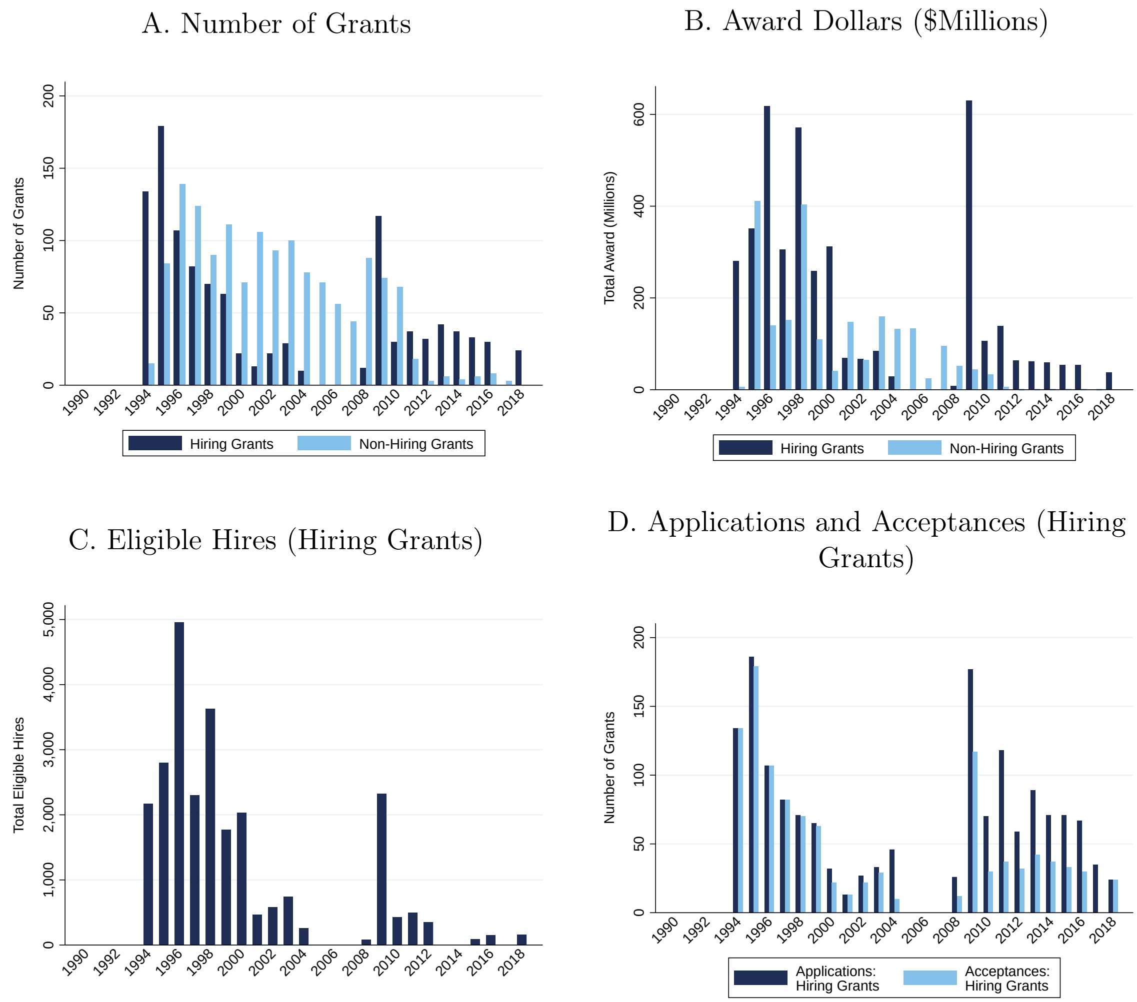
This paper is also able to exploit variation in grant applications that are rejected in the estimation model. Appendix Figure A2.D shows the number of grant applications and acceptances in each year of the COPS program within our sample. Prior to 2000, nearly all applications for hiring grants were awarded. However, after 2000, these grants became more competitive and demand for hiring grants exceeded the number of grants awarded.

### A1.2.2 Discussion of Model

The main features of the model are provided in Section 1; this section provides additional detail on the model specification and robustness. The general model used in this paper is:

[16] Prior to 2009, hiring grants provided up to 75% funding per officer or a max of $75,000 per officer over 3 years. In 2009, funding rules were changed to provide up to 100% funding per officer or a max of $125,000 per officer over 3 years.

Figure A2: COPS Grants Over Time




Note: The figures above summarize the DOJ COPS grant variation between 1990-2018 for this sample of cities. Panel A plots the number of hiring grants and other non-hiring COPS grants distributed in each year. Panel B plots the award dollars distributed each year under these two types of grants. Panel C plots the number of eligible hires designated by hiring grants in each year. Panel D plots the number of grant applications and acceptances in each year of the sample.

$$Y_{it} = \theta S_{it-1} + \gamma^{'} X_{it} + \rho_i + \psi_{st} + \varepsilon_{it}$$
$$S_{it-1} = \pi Z_{it-1} + \phi^{'} X_{it} + \rho_i + \psi_{st} + \mu_{it}$$

where $Y_{it}$ is the outcome of interest, $S_{it-1}$ is the UCR measure of police employment, and $Z_{it-1}$ is the COPS instrument. This model includes U.S. Census covariates in $X_{it}$ (included in Table 1), police department fixed effects $\rho_i$, and state by year fixed effects $\psi_{st}$. More specifically, the COPS Eligible Hires IV specification is as follows:

$$
\begin{aligned}
Y_{it} =& \theta Police_{it-1} + \gamma_1 AwardNonHiring_{it-1} \\
& + \gamma_2 ApplyHiring_{it-1} + \gamma_3 ApplyNonHiring_{it-1} \\
& + \gamma^{'} X_{it} + \rho_i + \psi_{st} + \varepsilon_{it} \\
Police_{it-1} =& \pi COPSEligible_{it-1} + \phi_1 AwardNonHiring_{it-1} \\
& + \phi_2 ApplyHiring_{it-1} + \phi_3 ApplyNonHiring_{it-1} \\
& + \phi_x^{'} X_{it} + \rho_i + \psi_{st} + \mu_{it}
\end{aligned}
$$

There are three additional grant controls in these models. First, the model controls for the size of any non-hiring grant awards, which may fund technology improvements or targeted crime initiatives. [17] Second, the model includes indicators for whether an agency applied for hiring or non-hiring grants in a particular year. This variable captures changes in police employment and crime outcomes associated with grant applications, rather than awards, and controls for the possible outcome that departments increase (or decrease) hiring when they are interested in obtaining COPS grant funds but these funds are not awarded. The resulting model has the identification assumption that conditional on the decision to apply for a hiring grant, the number of officers designated by an awarded COPS hiring grant does not depend on changes in crime within a city. These application controls increase precision, though as discussed below, the models are robust to excluding them.

The model draws heavily on the existing literature on the COPS program. The models used in Evans and Owens (2007); Owens (2013) are identical to the model above, when the application controls are not included. Weisburst (2019b) explicitly controls for grant applications and uses an excluded instrument of indicators for grant awards, where both application and award variables are defined over a grant award period rather than in the first year the grant was distributed (lagged), as in the above model.

We include several variants of this model as robustness checks in Appendix Table A2. In specification (2), we assign grant eligible hires, awards, and applications according to the full time period of a grant from the first year of the award to the year when the funding ends, a feature of the design in Weisburst (2019b). The estimates using this approach are larger in magnitude but qualitatively consistent with the preferred estimates. In specifications (3)-(5), we consider different sub-groups of the sample defined by police department participation in the COPS grant programs. The results are robust to restricting to cities that applied for a hiring grant (3), received a hiring grant (4), or cities that both had grant applications that were accepted and rejected (5) at different points in the study sample period. Lastly, in specification (6), the results are robust to excluding controls for time-varying grant applications.

[17] The dollar value of hiring grants is excluded as this quantity is nearly perfectly collinear with the number of officers eligible for hiring for a grant, or $COPSEligible$.

Table A2: Additional Robustness Specifications, COPS IV

| B. COPS Eligible Hires IV | (1) Homicide Victims | (2) Black Homicide Victims | (3) White Homicide Victims | (4) Quality of Life Arrests | (5) Black Quality of Life Arrests | (6) White Quality of Life Arrests |
|---|---|---|---|---|---|---|
| (1) Baseline Model | -0.1023*** | -0.0500*** | -0.0441*** | 21.879*** | 8.099*** | 13.948*** |
| *(First Stage F-Test = 16.13)* | (0.0098) | (0.0044) | (0.0009) | ( 5.000) | ( 1.604) | ( 3.424) |
| Y-Mean | 223.31 | 129.98 | 59.16 | 49908.0 | 24807.0 | 24674.4 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |
| (2) Grants Split Across Grant Years | -0.1909*** | -0.0855*** | -0.0737*** | 46.107*** | 16.803*** | 29.290*** |
| *(First Stage F-Test = 64.07)* | (0.0189) | (0.0102) | (0.0109) | ( 5.104) | ( 1.479) | ( 4.840) |
| Y-Mean | 223.31 | 129.98 | 59.16 | 49908.0 | 24807.0 | 24674.4 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |
| (3) Cities that Applied for Grants | -0.1023*** | -0.0500*** | -0.0441*** | 21.880*** | 8.100*** | 13.949*** |
| *(First Stage F-Test = 16.13)* | (0.0098) | (0.0044) | (0.0009) | ( 4.999) | ( 1.604) | ( 3.423) |
| Y-Mean | 223.62 | 130.16 | 59.24 | 50006.3 | 24856.8 | 24722.5 |
| N | 6503 | 6475 | 6462 | 5805 | 5797 | 5784 |
| (4) Cities with Accepted Grant | -0.1021*** | -0.0499*** | -0.0441*** | 21.777*** | 8.069*** | 13.876*** |
| *(First Stage F-Test = 16.73)* | (0.0093) | (0.0042) | (0.0009) | ( 4.928) | ( 1.589) | ( 3.370) |
| Y-Mean | 225.77 | 131.41 | 59.81 | 50571.1 | 25142.8 | 24998.4 |
| N | 6331 | 6304 | 6293 | 5636 | 5628 | 5615 |
| (5) Cities with Accepted & Rejected Grants | -0.1202*** | -0.0589*** | -0.0457*** | 27.642*** | 9.751*** | 18.032*** |
| *(First Stage F-Test = 4.99)* | (0.0249) | (0.0115) | (0.0016) | ( 7.155) | ( 1.918) | ( 5.280) |
| Y-Mean | 237.47 | 130.43 | 65.77 | 60307.1 | 29672.2 | 30154.9 |
| N | 4712 | 4690 | 4685 | 4269 | 4263 | 4253 |
| (6) Drop Application Controls | -0.1032*** | -0.0504*** | -0.0451*** | 21.642*** | 7.946*** | 13.837*** |
| *(First Stage F-Test = 14.92)* | (0.0097) | (0.0045) | (0.0011) | ( 4.889) | ( 1.556) | ( 3.362) |
| Y-Mean | 223.31 | 129.98 | 59.16 | 49908.0 | 24807.0 | 24674.4 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Baseline specifications correspond to models in Table 2 and Table 2.

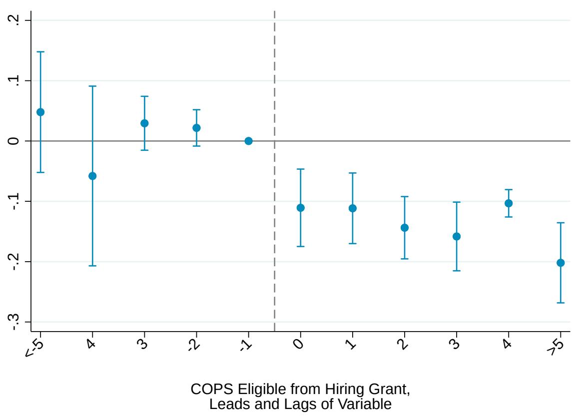
### A1.2.3 Reduced Form Results Over Time

As an additional check of the COPS instrument, we present the reduced form results of the model over time. This exercise directly relates the number of COPS eligible hires to our primary outcomes in the years preceding a grant award. To do this, we construct lead variables of the Cops Eligible Hires IV for the four preceding periods (t=-4...-2, -1 omitted) and lag variables of the IV (t=0...4) as well as bookend variables that sum the leads and lags for periods -5 and before and +5 and later. Note that this framework uses the IV of Eligible Hires which is not an indicator for a grant but the number of officers designated by a grant. This structure flexibly permits multiple treatments over time, as a department that has two grant awards separated by a period of years may have positive values for both leads and lags in the same observation that reflect these multiple treatments.

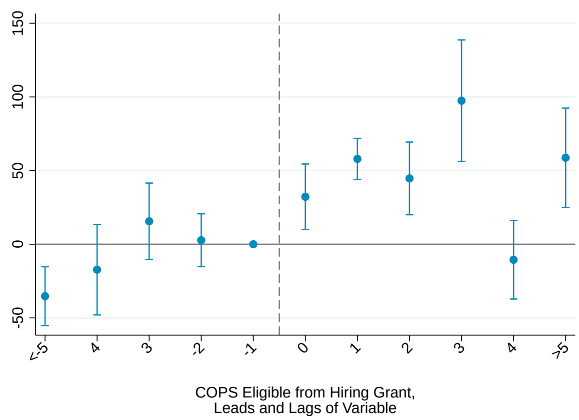
Appendix Figure A3.A and A3.B shows these result for homicides and quality of life arrests. Prior to a COPS hiring grant, there is no trend in homicides, suggesting that the distribution of grants is exogenous to these outcomes. Coinciding with the grant awards there is a negative shift in the number of homicide victims that is persistent over time. Similarly, these outcomes do not show a pre-trend and show a consistent increase in this arrest category after the grant receipt.

Figure A3: Reduced Form Estimates Over Time, COPS Eligible Hires

A. Homicide Victims




B. Quality of Life Arrests




Note: Standard errors are clustered at the city-level. The sample covers treatment variation from 1990-2018. Each graph plots the reduced form relationship between the number of eligible hires designated by COPS hiring grants and an outcome over time (IV). The graphs plot lags and leads of the IV, where the -5 and +5 categories are summed values of remaining periods, and the first lead (t=-1) is omitted. Controls include corresponding lags and leads of other grant variables: whether a city applied for a hiring or non-hiring grant, and the award size of non-hiring grant awards.

# A2 Supplementary Results

In this appendix we present a series of supplementary results which compliment the analyses presented in the main body of the paper.

## A2.1 Ordinary Least Squares Estimates

We begin by presenting least squares estimates of the effect of police manpower on each of our main outcomes estimated using equation (1). The results are presented in Table A3. In keeping with prior literature which studies police manpower, least squares estimates are negative but are smaller in magnitude than IV estimates using the COPS hiring instrument. With respect to the measurement error models, given the that the first stage coefficient is not far from 1, the OLS estimates are fairly similar in magnitude, but remain smaller in magnitude. We note that as our models are estimated in levels, the strength of the first stage coefficient is closer to 1 than in Chalfin and McCrary (2018) which estimates models using growth rates.

## A2.2 Robustness

Our estimates indicate that each police officer hired saves between 0.06 and 0.1 lives, depending upon the approach to identification. Approximately half of those saved are Black victims and between 25-50% are white though, in per capita terms, the effects are approximately twice as large for Black than for white civilians. In this appendix, we subject these results to greater scrutiny by re-estimating the models conditioning on a number of more granular fixed effects as well as using several different functional forms. These estimates are presented in Tables A4 and A5 which tests the robustness of the measurement error-corrected models as well as the models which use the COPS instrument.

In each table, we begin by presenting estimates from our baseline model referenced in Table 2. Next, we re-estimate our models without using population weights. These estimates conform closely with the baseline estimates that are weighted by each city's 1980 population. Next, we present "reflected" estimates in which we switch the role of the UCR and the U.S. Census measures of police manpower or, in the case of the COPS instrument, substitute the U.S. Census ASG measure of police for the UCR measure. These coefficients provide an alternative estimate of the effect of police manpower given that the role of each variable is ambiguous under the assumptions of the classical measurement error model. In the case of the COPS instrument, the estimates also provide some assurance that the estimates reported in the main body of the paper are not the result of specification searching. In all cases, the estimates are extremely similar.

In keeping with much of the literature, in our baseline model, we estimate the effect of police manpower on race-specific homicide victimization using the first lag of the police variable. In model (4), we re-specify the model using a contemporaneous measure of police manpower. Once again, estimates are very similar. In models (5) and (6), instead of conditioning on interacted state-by-year fixed effects we condition instead on either on population group-by-year fixed effects, dividing our cities into the following population groups 50-100k, 100-200k, >250k residents in 1980 (5) or homicide group-by-year fixed effects which use quartiles of the homicide rate in 1980 (6). In each case, estimates are nearly identical to those reported in Table 2. In model (7), we estimate the model with additional controls for municipal education spending to adjust for spending allocation decisions in cities; the results show that the returns to police manpower are similar when holding total municipal spending and education spending fixed.

In model (8) we present estimates in which we do not condition on covariates. For the measurement error models, these estimates are larger in magnitude which is consistent with the idea that the inclusion of covariates helps to capture time-varying omitted factors which are correlated with police hiring and outcomes. For the models which use the COPS instrument, the homicide estimates are smaller, though the sign of the estimates is consistent with that in our baseline models.

Next, we consider a log-log specification which yields a direct estimate of the elasticity of each outcome with respect to police force size, where outcomes are defined as $\log(y+1)$ to account for zeros (9). Because there are sometimes zero homicides in a given year for a given subgroup of victims, we utilize the inverse hyperbolic sine transformation (Ramirez et al., 1994) in (10). For the measurement error corrected models, we see that the elasticity of overall homicides with respect to police manpower is approximately $-0.5$, which is smaller than the elasticity calculated from our levels models of $-1.4$-3. It is worth mentioning that our levels models yield incredibly similar estimates for population weighted and unweighted models implying that the number of lives saved is a constant function of the change in police employment in a city. Because these constant changes in homicide occur relative to very different base rates of homicide (and police employment), we do not expect a percentage change in police employment to produce a uniform percentage decrease in homicide in our sample. It is therefore unsurprising that the elasticities from the log-log models differ from the elasticities that are implied by our baseline models. At the same time, using the ASG instrument, we note that our log-log models show estimates are substantively similar to those reported in most of the prior literature including Evans and Owens (2007) and Chalfin and McCrary (2018). Using the COPS instrument, there is no first stage when the model is specified in log-log form in this set of cities; as such the estimates cannot be interpreted. This lack of a first stage is likely due to the small set of cities in this sample, as we are restricted to using large cities to merge to Census police employment and expenditure data which defines our baseline set of covariates. This sample differs from prior work on COPS that typically uses a larger set of cities with a lower population threshold (Evans and Owens, 2007; Mello, 2019; Weisburst, 2019b).

We investigate the potential role of reporting. There are generally four reasons reported arrests could increase.

1. There is an increase in criminality. 2. There is an increase officer behavior 3. The hierarchical structure of the UCR. 4. There is a change in police reporting.

The first point is not consistent with the large decreases in homicides we observe. The second is leading primary hypothesis. The third is unlikely because the increases in low level (and other non-index arrests) dwarf the magnitude of the decreases in index arrests and homicides. We provide more evidence against the fourth in Table A6.

While our primary estimates provide robust evidence *reported* arrests for low level crimes increase, these are based on police reports. Thus a natural questions is whether police reporting change. First it is worth noting in all models we control for state by year FE, so any policy which varies within state across years (but is shared with departments) are accounted for with that control. We focus on large departments which generally have most consistent reporting regimes. Moreover, we include uncategorized arrests in our definition of low level arrests, so our approach accounts for any discretionary behavior that could be picked up there.

However, it could still be that as resources become more plentiful, departments record better records. To address this, in Table A6, we reestimate the main models for low level arrests. In panel 1 we provide our main estimates for comparison. In the next panel, we present estimates for the same models, expect now dropping all observations in which there were zero observations in a crime. Essentially the results are unchanged. In the final two panels we explore whether the extensive margin crime reporting changes for departments for arrests subgroups. We find generally the estimated relationships are small, suggestive there are not large increases in reporting due to increases in police reporting.

Next, we present estimates in which we do not remove outliers (11) and in which we use a balanced panel retaining only panels with complete data (12); estimates are not sensitive to either of these choices. Also, for the ASG models, we present estimates for the 1990-2018 sample period which corresponds with the sample period in the COPS models (13). Estimates for homicides are very similar between the two IV strategies when the models are executed using the same data. For "quality of life" arrests, the estimates are considerably larger in the COPS models indicating either that there is some remaining simultaneity bias in the measurement error corrected models or that the instruments identify different local average treatment effects.

Finally, we consider the sensitivity of our estimates to highly leveraged cities. Given that estimates are similar with and without the use of population weights, highly leveraged cities are unlikely. We confirm this empirically in Appendix Figure A4 which re-estimates our primary outcomes removing one city at a time and plots the distribution of estimated treatment effects for homicide (Panels A and B) and "quality of life" arrests (Panels C and D).

### A2.3 Treatment Effect Heterogeneity

In this section, we explore several different dimensions of treatment effect heterogeneity. We begin by the sensitivity of our estimates to the inclusion of individuals of Hispanic ethnicity in our Black and white homicide counts. Next, we consider the heterogeneous effects of police force size on homicides and various types of arrests by age and gender as well as by race.

#### A2.3.1 Disaggregated Race Categories

Our main analyses consider the impact of police force size on homicides with non-Hispanic white and non-Hispanic Black victims. In this section, we consider an alternative conceptualization in which individuals of Hispanic ethnicity are folded into the Black and white categories. We also separately estimate the effect of police force size on homicides with Hispanic victims. Estimates are presented in Appendix Table A7 and Appendix Table A8. There is not a large difference between estimates for non-Hispanic Black victims and overall Black victims since there are relatively few Black victims of Hispanic origin in the data. With respect to Hispanic victims, each police officer abates between 0.006 and 0.015 homicides with Hispanic victims depending on which IV estimate is used.

#### A2.3.2 Homicide Victimization by Race, Sex and Age

Next, we consider the effect of police on homicides focusing on more granular demographic subgroups, segmenting the population into sixteen age-race-gender bins. We present this analysis in Appendix Figure A5. In the figure, we present estimates separately by race for eight different gender-age groups defined by the intersection of four age groups ($< 14$, 15-24, 25-44 and $> 45$) and male and female gender. The analysis shows that public investments in police manpower are considerably more effective at abating male homicides than female homicides. While roughly 80% of the homicides in our data have a male victim, the homicide reductions arising from a larger police force are even more concentrated among men than the raw data suggest. Effects are large in magnitude and statistically significant for both Black and white males between the ages of 15-44.

#### A2.3.3 Arrest Outcomes by Offense Type

We present several supplementary results for our analysis of arrests. For each aggregate category (index crimes, "quality of life" crimes and other crimes), we provide estimates of the effect of police manpower on arrests of each type. For index crimes, we also provide estimates of the effect of police manpower on crimes known to law enforcement. In each table, we present the coefficient from a regression of the number of arrests of each crime type in city $i$ in year $t$ on police force size, net of fixed effects and covariates. We also transform the coefficient into an elasticity and a per capita estimate and report the average number of arrests in order to provide a sense for the density of each arrest type in the data.

**Index Arrests** We provide additional detail on the effect of police manpower on index crimes known to law enforcement and index crime arrests in Table A9. The most common index crimes are theft and burglary. Overall, violent crimes (homicide, rape, robbery and aggravated assault) constitute just over 20 percent of index crimes. Index crime arrests follow a similar pattern.

Consistent with the extant literature, for both of our identification strategies, there is strong evidence that a larger police force leads to a reduction in index crimes. On an annual basis, each police officer hired is estimated to abate between approximately 0.07-0.1 homicides, 3-4 robberies, 4-5 burglaries, 5-7 thefts and 4-6 motor vehicle thefts. In elasticity terms, estimates are largest for murder, robbery, burglary, and motor vehicle theft, a finding that is consistent with the majority of prior literature (Chalfin and McCrary, 2017). With respect to arrests, larger police forces lead to significantly *fewer* arrests for robbery and motor vehicle theft, two common street crimes. In the COPS model, there is also evidence that large police forces make fewer arrests for homicide and burglary. Since a larger police force leads to reductions in both crime and arrests, this suggests that the primary driver of manpower-led crime reductions is deterrence rather than incapacitation (Owens, 2013), a finding which narrows the scope for police hiring to contribute to mass incarceration.

Appendix Figure A6 explores heterogeneity in the arrest estimates by race. The figures show that the level changes in low-level arrests are, for the most part, evenly split across Black and white civilians though there is evidence that robbery arrests decline with police force size to a greater degree for Black versus white civilians. As in the aggregate results, similar level effects for Black and white civilians imply disproportionately large decreases in index crime arrests for Black civilians.

**"Quality of Life" Arrests** We provide additional detail on the effect of police manpower on "quality of life" arrests focusing on specific arrest types in Appendix Table A10. Leaving aside uncategorized arrests, the most common quality of life arrests are drug possession, disorderly conduct and liquor law violations. Using both of our identification strategies, we see that the marginal "quality of life" arrests that are made when a city expands the size of its police force are predominantly for liquor law violations and drug possession and, to a lesser extent, disorderly conduct. The coefficients on liquor violations imply that such arrests are incredibly sensitive to police force size with elasticities of 8-14 depending upon the model.

Appendix Figure A7 explores heterogeneity in the arrest estimates by age and race. Effects are similar in magnitude for Black and white civilians. As Black civilians constitute just under one quarter of our sample, this implies that they disproportionately bear the burden of such arrests.

**Other Arrests** We also present results for other arrests which are classified as neither index nor "quality of life" crimes. Such crimes include simple assaults, the sale of illegal drugs, driving under the influence (DUI), fraud and weapons charges among other offense types. Here we report evidence that larger police forces make more arrests for simple assault, fraud, forgery and sex offenses (other than rape) and fewer arrests for weapons possession and stolen property. Appendix Figure A8 explore heterogeneity in the arrest estimates by age and race.

### A2.4 Fatal Encounters Between Police Officers and Suspects

The current emphasis of public discourse on racial differences in the use of lethal force by law enforcement, particularly as it pertains to Black men (Knox et al., 2020; Knox and Mummolo, 2020), raises a natural question regarding the extent to which the protective effects of an expanded police force might be "outweighed" by the number of lives taken by the police. In this section, we estimate the effect of police force size on fatal encounters between police officers and civilians of different races and motivate a simple bounding exercise that is intended to shed further light on the degree to which the taking of lives by police might erode their protective effects.

We begin by estimating the effect of police force size on fatal encounters with police in Table A12. By necessity, we study a shorter time period (2010-2018) given the absence of reliable national data on fatal police shootings in earlier years. This restricted sample creates important challenges to our COPS identification strategy due to an insufficiently powered first stage, and as a result, we instead focus on the measurement error correction models. Overall, the results are similar for gunshot deaths and all causes of death, though the gunshot death category is likely more precisely estimated than other causes of death in the data.

Next, while scholars have noted a number of serious limitations with respect to documentation of police killings in the Supplementary Homicide Reports (Barber et al., 2016), we nevertheless report estimates using these data for the sake of completeness. We begin by assessing the extent to which the Fatal Encounters data and the SHR data move together during the 2010-2018 period for which we have reliable Fatal Encounters data. To do so we regress the Fatal Encounters measure of police killings on police killings in the SHR net of covariates and fixed effects. Despite evidence that police killings are under-counted in the SHR, these results indicate a close correspondence between the two measures. Indeed, for overall police killings as well as police killings of Black and white suspects, $t$-ratios on the SHR measure are between 6 and 10. Given the positive and significant correlation between the two measures during the 2010-2018 period, we use the SHR data to estimate the effect of police manpower on police killings for both the 2010-2018 period and the full 1980-2018 sample period. For the 2010-2018 period, estimates using the SHR data are quantitatively and qualitatively similar to those estimated using the Fatal Encounters data. For the 1980-2018 sample period, estimates are statistically significant and *negative*, though these estimates are more likely to be compromised by data quality issues in this longer sample period.

The point estimate on fatal shootings suggests that each police officer hired leads to an increase of 0.0005 in the number of civilians shot by police, though this estimate is not statistically significant. Turning to our race-specific results, the point estimate for white civilians is negative ($-0.006$) but is likewise not significant at conventional levels. For Black civilians, the estimate (0.0019) is positive and statistically significant ($p < 0.05$), though the significance of the result is sensitive to the time window employed in the analysis (Figure A9). This result is echoed in the analysis using the SHR measure of police killings over the same period. Further, the $p$-value on a test of equality for the Black and white coefficients in our preferred specification is $< 0.01$ and remains significant when alternative time windows or measures of fatalities are used. Collectively, the analysis suggests that police hiring has different implications for fatal encounters between police and Black versus white suspects.

Though diminished precision in our analysis of police shootings means that we are unable to make strong claims about the precise relationship between police manpower and fatal encounters, we can perform a simple but potentially informative bounding exercise using the 95% confidence interval around our estimates of the effects of police manpower on homicides and police shootings. This exercise is necessarily speculative and we note that police force size is only one of many elements that may contribute to a department's use of deadly force. Moreover, the analysis is based on the normative assumption that the life of a homicide victim and the life of an individual shot by a police officer receives the same social weight; we recognize that this social weighting assumption is restrictive as it could certainly be the case that lives taken by actors of the state could be more costly than homicides perpetrated by civilians.

The upper bounds of the 95% confidence intervals of our estimates imply that an additional officer hired results in 0.05 fewer homicides and 0.0019 additional civilians fatally shot by police. These conservative bounds imply that for every life the marginal police officer takes in a fatal encounter, he/she abates at least 17.2 homicides. For white civilians, the upper boundaries of the confidence intervals indicates that there is, at most, one additional fatal shooting of a white civilian for every 129 police officers hired. For black civilians, the estimate is 6.3, indicating that perhaps as many as 16% of abated homicides are outweighed by fatal shootings (Appendix Table A13).

Two lessons are apparent from this exercise. First, it is unlikely that expanding the number of police has resulted in a net increase in the number of lives lost for either Black or white civilians. It is critical to note though that this is an extraordinarily low accountability standard for the police. Second, while the addition of police manpower disproportionately saves Black lives, larger police forces may, in fact, generate more fatal shootings of Black civilians. As a result, a meaningful share of Black lives saved by police may be outweighed by lives taken by police, a proposition which is especially likely to hold for cities with large Black populations.

### A2.5 Deaths and Injuries of Police Officers

In addition to estimating the effect of police force size on police shootings, we also estimate the effect of police force size on violence *against* police officers. These results are presented in Appendix Table A14. We observe that each officer hired leads to between 0.14 and 0.23 *fewer* officer injuries. This result is counter-intuitive in the sense that, other things equal, the risk of adverse events rises with the size of a city's police force. Instead, the evidence suggests that this mechanical "exposure" effect is dominated by the protective effect of greater manpower and may increase the share of officers who patrol in teams or the speed which officers are able to assist a fellow officer in distress. We do not find any robust effects of law enforcement on officer deaths but these are difficult to study given that they are rare events.

Table A3: OLS Model Results

| **OLS Model** | Coeff. | S.E. | Elasticity | $\beta$/Pop. | Mean | N |
|---|---|---|---|---|---|---|
| **Homicides** | | | | | | |
| Victims | -0.051*** | ( 0.004) | -1.25 | -0.00 | 249.0 | 8582 |
| Black | -0.022*** | ( 0.002) | -0.95 | -0.01 | 140.4 | 8552 |
| White | -0.009*** | ( 0.001) | -0.75 | -0.00 | 65.5 | 8531 |
| Clearance Rate | 0.000 | ( 0.001) | 0.00 | - | 65.2 | 7699 |
| Black | 0.000 | ( 0.001) | 0.03 | - | 62.5 | 6089 |
| White | -0.001 | ( 0.001) | -0.10 | - | 69.4 | 7070 |
| **Arrests** | | | | | | |
| Quality of Life | 5.96*** | ( 0.72) | 0.45 | 0.44 | 60121 | 7824 |
| Black | 1.04* | ( 0.53) | 0.15 | 0.32 | 30843 | 7788 |
| White | 5.16*** | ( 0.24) | 0.82 | 0.56 | 28758 | 7799 |
| Index | -0.81*** | ( 0.24) | -0.23 | -0.06 | 16340 | 7817 |
| Black | -0.60*** | ( 0.18) | -0.31 | -0.18 | 8931 | 7775 |
| White | -0.37*** | ( 0.08) | -0.24 | -0.04 | 7200 | 7792 |
| **Index Crimes** | -16.50*** | ( 0.86) | -1.03 | -0.99 | 96791 | 8675 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. All models are weighted by population of each city in 1980 and cover the period 1981-2018. Models have differing numbers of observations due to data availability and the outlier cleaning procedure for outcomes described in Appendix A3. OLS models directly relate UCR police employment to outcomes. All models include covariates in Table 1. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). F.B.I. UCR data on arrests does not include sub-categories for Hispanic residents; as a result, white population share includes Hispanic residents for these outcomes in calculating the "$\beta$/Pop." measure. All estimates pass a Bonferroni multiple hypothesis correction of 20, except for the coefficient on "Quality of Life Arrests, Black."

Table A4: Robustness Specifications, ASG Employment IV

| A. ASG Employment IV | (1) Homicide Victims | (2) Black Homicide Victims | (3) White Homicide Victims | (4) Quality of Life Arrests | (5) Black Quality of Life Arrests | (6) White Quality of Life Arrests |
|---|---|---|---|---|---|---|
| (1) Baseline Model | -0.0583*** | -0.0261*** | -0.0159*** | 7.317*** | 2.278*** | 5.101*** |
| *(First Stage F-Test = 559.17)* | (0.0043) | (0.0026) | (0.0018) | ( 0.882) | ( 0.529) | ( 0.477) |
| Y-Mean | 249.01 | 140.32 | 65.54 | 60243.6 | 30896.3 | 28827.3 |
| N | 8554 | 8524 | 8503 | 7804 | 7768 | 7779 |
| (2) Not Weighted by Population | -0.0506*** | -0.0229*** | -0.0111*** | 8.757*** | 3.157** | 5.684*** |
| *(First Stage F-Test = 45.32)* | (0.0118) | (0.0069) | (0.0041) | ( 1.754) | ( 1.418) | ( 0.524) |
| Y-Mean | 39.20 | 22.91 | 9.94 | 8483.8 | 3796.1 | 4565.0 |
| N | 8554 | 8524 | 8503 | 7804 | 7768 | 7779 |
| (3) ASG as Endogenous X, UCR as IV | -0.0642*** | -0.0274*** | -0.0107*** | 6.413*** | 1.111* | 5.558*** |
| *(First Stage F-Test = 5016.29)* | (0.0045) | (0.0028) | (0.0014) | ( 0.763) | ( 0.570) | ( 0.254) |
| Y-Mean | 249.01 | 140.32 | 65.54 | 60243.6 | 30896.3 | 28827.3 |
| N | 8554 | 8524 | 8503 | 7804 | 7768 | 7779 |
| (4) Police Employment not Lagged | -0.0635*** | -0.0303*** | -0.0209*** | 10.726*** | 3.976*** | 6.877*** |
| *(First Stage F-Test = 550.33)* | (0.0048) | (0.0029) | (0.0035) | ( 0.858) | ( 0.617) | ( 0.317) |
| Y-Mean | 255.93 | 143.51 | 67.94 | 61510.7 | 31364.9 | 29596.9 |
| N | 8568 | 8532 | 8522 | 7831 | 7789 | 7809 |
| (5) Population Group by Year FE | -0.0563*** | -0.0248*** | -0.0157*** | 6.850*** | 2.077*** | 4.851*** |
| *(First Stage F-Test = 524.25)* | (0.0043) | (0.0027) | (0.0018) | ( 0.920) | ( 0.531) | ( 0.533) |
| Y-Mean | 249.01 | 140.32 | 65.54 | 60243.6 | 30896.3 | 28827.3 |
| N | 8554 | 8524 | 8503 | 7804 | 7768 | 7779 |
| (6) Homicide Group by Year FE | -0.0541*** | -0.0235*** | -0.0157*** | 6.892*** | 2.159*** | 4.793*** |
| *(First Stage F-Test = 621.39)* | (0.0045) | (0.0028) | (0.0021) | ( 0.948) | ( 0.508) | ( 0.576) |
| Y-Mean | 249.01 | 140.32 | 65.54 | 60243.6 | 30896.3 | 28827.3 |
| N | 8554 | 8524 | 8503 | 7804 | 7768 | 7779 |
| (7) Control for Education Spending | -0.0556*** | -0.0256*** | -0.0168*** | 6.875*** | 2.151*** | 4.785*** |
| *(First Stage F-Test = 550.96)* | (0.0045) | (0.0028) | (0.0021) | ( 0.964) | ( 0.511) | ( 0.587) |
| Y-Mean | 249.87 | 140.80 | 65.76 | 60472.4 | 31016.4 | 28935.2 |
| N | 8448 | 8419 | 8397 | 7704 | 7668 | 7679 |

Table A4: Robustness Specifications, ASG Employment IV (Continued)

| A. ASG Employment IV | (1) Homicide Victims | (2) Black Homicide Victims | (3) White Homicide Victims | (4) Quality of Life Arrests | (5) Black Quality of Life Arrests | (6) White Quality of Life Arrests |
|---|---|---|---|---|---|---|
| (8) Excluding Covariates | -0.1084*** | -0.0492*** | -0.0317*** | 0.638** | -0.534** | 1.213*** |
| *(First Stage F-Test = 4258.64)* | (0.0039) | (0.0019) | (0.0006) | ( 0.253) | ( 0.267) | ( 0.116) |
| Y-Mean | 248.45 | 140.01 | 65.38 | 60099.8 | 30823.4 | 28758.3 |
| N | 8603 | 8573 | 8552 | 7848 | 7812 | 7822 |
| (9) Log Model (Variable+1) | -0.5359** | -0.7677*** | -0.3841 | 0.3993* | 0.5049** | 0.3826* |
| *(First Stage F-Test = 177.23)* | (0.2530) | (0.2920) | (0.2374) | (0.2077) | (0.2303) | (0.2226) |
| Y-Mean | 4.14 | 3.45 | 2.87 | 9.51 | 8.46 | 8.81 |
| N | 8552 | 8522 | 8501 | 7802 | 7766 | 7777 |
| (10) Inverse Hyperbolic Sine | -0.5224* | -0.7643** | -0.3367 | 0.3875* | 0.4666** | 0.3753* |
| *(First Stage F-Test = 176.98)* | (0.2731) | (0.3275) | (0.2483) | (0.2062) | (0.2324) | (0.2217) |
| Y-Mean | 4.75 | 4.00 | 3.41 | 10.20 | 9.14 | 9.50 |
| N | 8554 | 8524 | 8503 | 7804 | 7768 | 7779 |
| (11) Raw Data | -0.0594*** | -0.0268*** | -0.0162*** | 7.017*** | 2.084*** | 5.000*** |
| *(First Stage F-Test = 65.94)* | (0.0041) | (0.0026) | (0.0018) | ( 0.848) | ( 0.494) | ( 0.478) |
| Y-Mean | 247.77 | 139.74 | 64.88 | 60079.1 | 30821.9 | 28739.9 |
| N | 8591 | 8561 | 8540 | 7834 | 7798 | 7808 |
| (12) Balanced Panel | -0.0599*** | -0.0268*** | -0.0192*** | 8.335* | 4.398* | -0.334 |
| *(First Stage F-Test = 1907.06)* | (0.0040) | (0.0025) | (0.0059) | ( 4.878) | ( 2.562) | ( 1.767) |
| Y-Mean | 257.33 | 142.45 | 26.18 | 21708.0 | 9051.5 | 11912.2 |
| N | 6951 | 6386 | 6157 | 4687 | 4345 | 4536 |
| (13) COPS Timeframe, 1990-2018 | -0.0894*** | -0.0453*** | -0.0409*** | 5.343*** | 0.899 | 4.670*** |
| *(First Stage F-Test = 806.26)* | (0.0045) | (0.0023) | (0.0016) | ( 1.424) | ( 0.762) | ( 0.620) |
| Y-Mean | 223.25 | 129.78 | 59.17 | 50034.1 | 24854.8 | 24751.9 |
| N | 6504 | 6476 | 6463 | 5819 | 5811 | 5798 |

*p<0.1, **p<0.05, ***p<0.01.

Note: Standard errors are clustered at the city-level. Baseline specifications correspond to models in Table 2. Model (2) removes population weights, (3) replaces the endogenous X as the U.S. Census police employment record, (4) estimates the model using a police employment measure (and IV) that are not lagged, (5) includes Population Bin (50K-100K, 100K-250K, >250K in 1980) by Year by Month Fixed Effects, (6) includes City Homicide Quartile (1980) by Year by Month Fixed Effects, (7) controls for Education Spending at the city-year level, (8) removes covariates in Table 1 from the model, (9) transforms the model to a log-log specifications where variables are transformed as $y' = log(y+1)$, (10) uses an inverse hyperbolic sine transformation $y = log(y + \sqrt{y^2+1})$, (11) does not remove outlier observations identified in data cleaning, (12) restricts the sample to the balanced panel, and (13) restricts to the sample period of the COPS IV specification, 1990-2018.

Table A5: Robustness Specifications, COPS Eligible Hires IV

| **B. COPS Eligible Hires IV** | (1) Homicide Victims | (2) Black Homicide Victims | (3) White Homicide Victims | (4) Quality of Life Arrests | (5) Black Quality of Life Arrests | (6) White Quality of Life Arrests |
|---|---|---|---|---|---|---|
| (1) Baseline Model | -0.1023*** | -0.0500*** | -0.0441*** | 21.879*** | 8.099*** | 13.948*** |
| *(First Stage F-Test = 16.13)* | (0.0098) | (0.0044) | (0.0009) | ( 5.000) | ( 1.604) | ( 3.424) |
| Y-Mean | 223.31 | 129.98 | 59.16 | 49908.0 | 24807.0 | 24674.4 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |
| (2) Not Weighted by Population | -0.0939*** | -0.0461*** | -0.0424*** | 21.179*** | 8.379*** | 12.927*** |
| *(First Stage F-Test = 10.12)* | (0.0057) | (0.0030) | (0.0024) | ( 5.529) | ( 2.060) | ( 3.509) |
| Y-Mean | 37.48 | 22.55 | 8.87 | 8119.6 | 3644.0 | 4337.2 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |
| (3) ASG as Endogenous X, UCR as IV | -0.1090*** | -0.0533*** | -0.0456*** | 20.465*** | 7.581*** | 13.041*** |
| *(First Stage F-Test = 17.04)* | (0.0098) | (0.0043) | (0.0008) | ( 5.037) | ( 1.632) | ( 3.433) |
| Y-Mean | 223.18 | 129.74 | 59.15 | 50017.6 | 24846.4 | 24745.9 |
| N | 6510 | 6482 | 6469 | 5824 | 5816 | 5802 |
| (4) Police Employment not Lagged | -0.1039*** | -0.0569*** | -0.0516*** | 11.650** | 3.932* | 7.451*** |
| *(First Stage F-Test = 17.29)* | (0.0039) | (0.0037) | (0.0016) | ( 4.627) | ( 2.311) | ( 2.302) |
| Y-Mean | 226.35 | 131.40 | 60.72 | 51136.5 | 25433.3 | 25235.6 |
| N | 6318 | 6292 | 6282 | 5662 | 5657 | 5644 |
| (5) Population Group by Year FE | -0.0965*** | -0.0465*** | -0.0442*** | 21.885*** | 8.027*** | 14.031*** |
| *(First Stage F-Test = 16.10)* | (0.0075) | (0.0034) | (0.0010) | ( 5.455) | ( 1.812) | ( 3.669) |
| Y-Mean | 223.31 | 129.98 | 59.16 | 49908.0 | 24807.0 | 24674.4 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |
| (6) Homicide Group by Year FE | -0.0911*** | -0.0435*** | -0.0444*** | 24.258*** | 9.280*** | 15.144*** |
| *(First Stage F-Test = 17.48)* | (0.0054) | (0.0026) | (0.0013) | ( 5.761) | ( 2.009) | ( 3.774) |
| Y-Mean | 223.31 | 129.98 | 59.16 | 49908.0 | 24807.0 | 24674.4 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |
| (7) Control for Education Spending | -0.0924*** | -0.0438*** | -0.0452*** | 24.242*** | 9.267*** | 15.141*** |
| *(First Stage F-Test = 17.56)* | (0.0053) | (0.0026) | (0.0015) | ( 5.812) | ( 2.033) | ( 3.802) |
| Y-Mean | 224.31 | 130.56 | 59.42 | 50157.7 | 24934.7 | 24795.7 |
| N | 6425 | 6398 | 6384 | 5740 | 5732 | 5719 |

Table A5: Robustness Specifications, COPS Eligible Hires IV (Continued)

| **B. COPS Eligible Hires IV** | (1) Homicide Victims | (2) Black Homicide Victims | (3) White Homicide Victims | (4) Quality of Life Arrests | (5) Black Quality of Life Arrests | (6) White Quality of Life Arrests |
|---|---|---|---|---|---|---|
| (8) Excluding Covariates | -0.0174*** | -0.0124*** | -0.0088*** | 24.687*** | 9.912*** | 14.829*** |
| *(First Stage F-Test = 10.27)* | (0.0066) | (0.0014) | (0.0011) | ( 6.636) | ( 2.185) | ( 4.428) |
| Y-Mean | 222.78 | 129.68 | 59.01 | 49779.7 | 24744.6 | 24610.5 |
| N | 6571 | 6543 | 6530 | 5875 | 5867 | 5853 |
| (9) Log Model (Variable+1) | 3.2687 | 1.8721 | 5.0370 | 7.3753 | 7.3990 | 8.7346 |
| *(First Stage F-Test = 3.82)* | (2.2771) | (2.0776) | (3.0909) | (5.2097) | (5.1982) | (5.9580) |
| Y-Mean | 4.09 | 3.44 | 2.75 | 9.41 | 8.40 | 8.69 |
| N | 6529 | 6501 | 6488 | 5837 | 5829 | 5816 |
| (10) Inverse Hyperbolic Sine | 3.8236 | 2.1946 | 5.0905 | 6.8524 | 6.9950 | 8.1472 |
| *(First Stage F-Test = 4.45)* | (2.5043) | (2.3211) | (3.1539) | (4.5679) | (4.6455) | (5.2531) |
| Y-Mean | 4.70 | 3.99 | 3.29 | 10.11 | 9.09 | 9.38 |
| N | 6531 | 6503 | 6490 | 5839 | 5831 | 5818 |
| (11) Raw Data | -0.1023*** | -0.0500*** | -0.0441*** | 21.887*** | 8.103*** | 13.951*** |
| *(First Stage F-Test = 16.03)* | (0.0097) | (0.0044) | (0.0009) | ( 5.008) | ( 1.608) | ( 3.428) |
| Y-Mean | 223.24 | 129.94 | 59.14 | 49891.6 | 24798.6 | 24668.4 |
| N | 6537 | 6509 | 6496 | 5844 | 5836 | 5822 |
| (12) Balanced Panel | -0.1041*** | -0.0506*** | -0.0306*** | 48.697*** | 9.569*** | 39.080*** |
| *(First Stage F-Test = 15.18)* | (0.0103) | (0.0045) | (0.0059) | ( 3.528) | ( 1.210) | ( 2.584) |
| Y-Mean | 224.69 | 127.22 | 21.36 | 20994.3 | 8693.0 | 11261.9 |
| N | 5321 | 4918 | 4887 | 3755 | 3726 | 3583 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Baseline specifications correspond to models in Table 2. Model (2) removes population weights, (3) replaces the endogenous X as the U.S. Census police employment record, (4) estimates the model using a police employment measure (and IV) that are not lagged, (5) includes Population Bin (50K-100K, 100K-250K, >250K in 1980) by Year by Month Fixed Effects, (6) includes City Homicide Quartile (1980) by Year by Month Fixed Effects, (7) controls for Education Spending at the city-year level, (8) removes covariates in Table 1 from the model, (9) transforms the model to a log-log specifications where variables are transformed as $y' = log(y + 1)$, (10) uses an inverse hyperbolic sine transformation $y = log(y + \sqrt{y^2 + 1})$, (11) does not remove outlier observations identified in data cleaning, and (12) restricts the sample to the balanced panel.

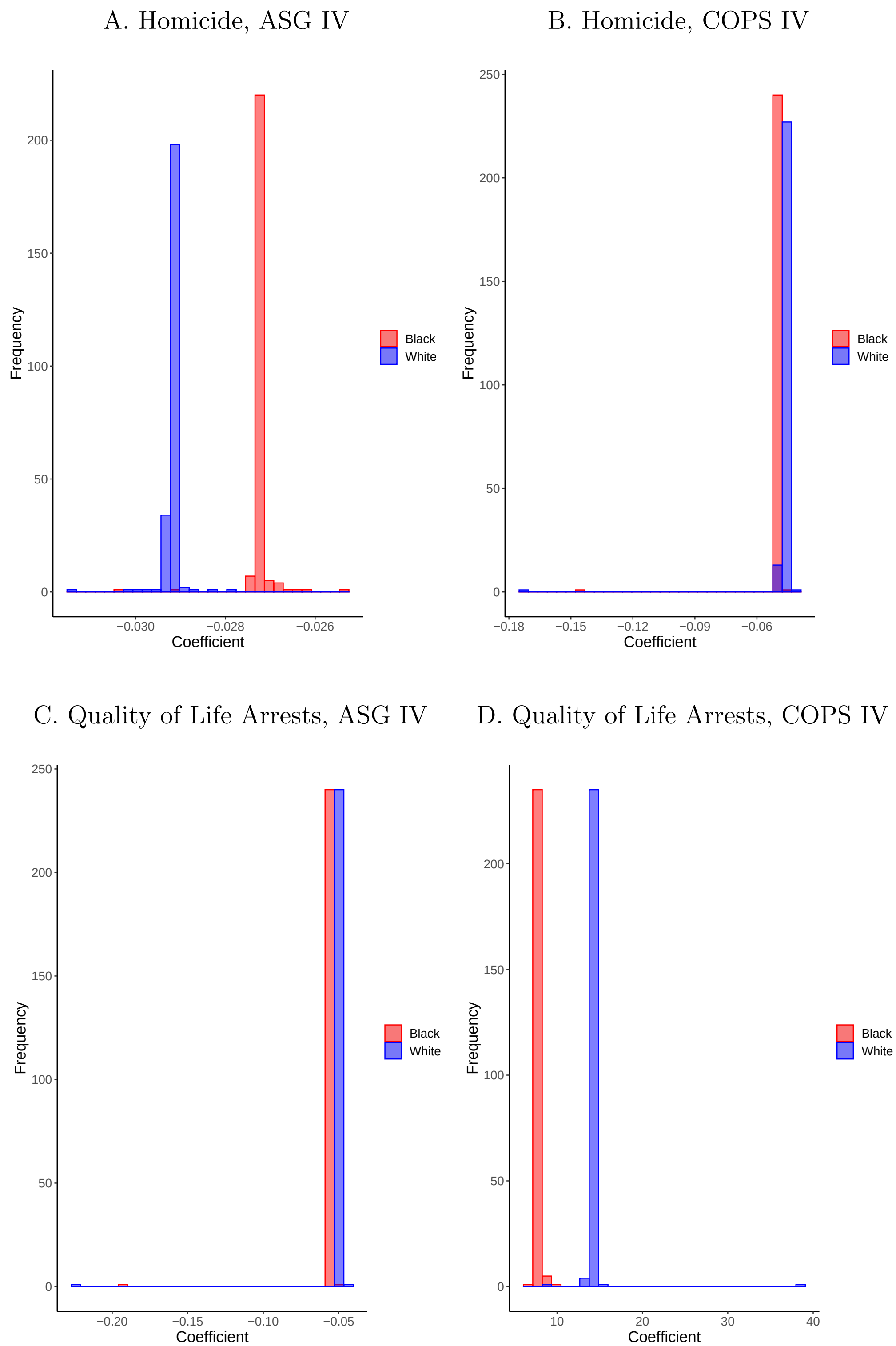
Table A6: Reporting of Quality of Life Arrests

| | A. ASG IV | | | B. COPS IV | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | | Black | White | | Black | White |
| | Quality of Life Arrests | Quality of Life Arrests | Quality of Life Arrests | Quality of Life Arrests | Quality of Life Arrests | Quality of Life Arrests |
| (1) Baseline Model | 7.317*** | 2.278*** | 5.101*** | 21.879*** | 8.099*** | 13.948*** |
| | ( 0.882) | ( 0.529) | ( 0.477) | ( 5.000) | ( 1.604) | ( 3.424) |
| Y-Mean | 60243.6 | 30896.3 | 28827.3 | 49908.0 | 24807.0 | 24674.4 |
| N | 7804 | 7768 | 7779 | 5839 | 5831 | 5818 |
| (2) Drop Zero Values | 7.312*** | 2.273*** | 5.099*** | 21.875*** | 8.096*** | 13.944*** |
| | ( 0.881) | ( 0.527) | ( 0.477) | ( 5.003) | ( 1.607) | ( 3.422) |
| Y-Mean | 60315.9 | 30996.6 | 28871.8 | 49989.7 | 24885.0 | 24721.5 |
| N | 7793 | 7715 | 7765 | 5828 | 5797 | 5805 |
| (3) Any Reporting (Total) | -0.000000 | 0.000000 | -0.000000 | -0.000000 | -0.000000 | -0.000001 |
| | (0.000000) | (0.000001) | (0.000000) | (0.000000) | (0.000001) | (0.000001) |
| Y-Mean | 0.999 | 0.997 | 0.999 | 0.998 | 0.997 | 0.998 |
| N | 7804 | 7768 | 7779 | 5839 | 5831 | 5818 |
| (4) Report Arrests in All Sub-Categories | -0.000044*** | -0.000051*** | -0.000042*** | -0.000031 | -0.000011 | -0.000016 |
| | (0.000014) | (0.000013) | (0.000012) | (0.000020) | (0.000020) | (0.000022) |
| Y-Mean | 0.284 | 0.226 | 0.238 | 0.276 | 0.219 | 0.223 |
| N | 7804 | 7768 | 7779 | 5839 | 5831 | 5818 |

*p<0.1, **p<0.05, ***p<0.01.

Note: Standard errors are clustered at the city-level. Baseline specifications correspond to models in Table 2. Model (2) replaces all zero values for aggregated Quality of Life arrests as missing. Model (3) test the binary outcome of reporting any positive value for aggregated Quality of Life arrests. Model (4) tests the binary outcome of whether all sub-categories of Quality of Life arrests have positive (non-zero) values, excluding "Uncategorized Arrests," which may serve as a residual category, and "Suspicious Person Arrests" which has zero values for a majority of city-years in the data.

Figure A4: Distribution of Estimates Excluding One City at a Time




Note: Standard errors are clustered at the city-level. Figures present histograms of the primary specifications (with identical controls and sample periods) where each estimate drops a different single city from the sample. All models are weighted by population.

Table A7: Results Dis-aggregated by Race Subgroups, ASG Employment IV

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | Non-Hispanic | | Non-Hispanic | | Other |
| **A. ASG Employment IV** | Black | Black | White | White | Hispanic | Race |
| (1) Homicide Victims | -0.0272*** | -0.0261*** | -0.0292*** | -0.0159*** | -0.0149*** | -0.0015*** |
| | (0.0027) | (0.0026) | (0.0018) | (0.0018) | (0.0009) | (0.0001) |
| Y-Mean | 142.19 | 140.32 | 100.71 | 65.54 | 37.57 | 6.09 |
| N | 8521 | 8524 | 8513 | 8503 | 8470 | 8475 |
| (2) Clearance Rates | 0.0007 | 0.0008 | -0.0004 | -0.0007 | 0.0012 | -0.0007 |
| | (0.0007) | (0.0007) | (0.0009) | (0.0008) | (0.0024) | (0.0013) |
| Y-Mean | 62.60 | 62.56 | 67.73 | 69.46 | 65.73 | 60.12 |
| N | 6072 | 6067 | 7315 | 7046 | 2393 | 2370 |
| (3) Index Arrests | -0.688*** | | -0.447*** | | | 0.161*** |
| | ( 0.202) | | ( 0.091) | | | ( 0.017) |
| Y-Mean | 8927.8 | | 7212.2 | | | 244.7 |
| N | 7755 | | 7772 | | | 7723 |
| (4) Quality of Life Arrests | 2.278*** | | 5.101*** | | | -0.198*** |
| | ( 0.529) | | ( 0.477) | | | ( 0.036) |
| Y-Mean | 30896.3 | | 28827.3 | | | 709.9 |
| N | 7768 | | 7779 | | | 7729 |

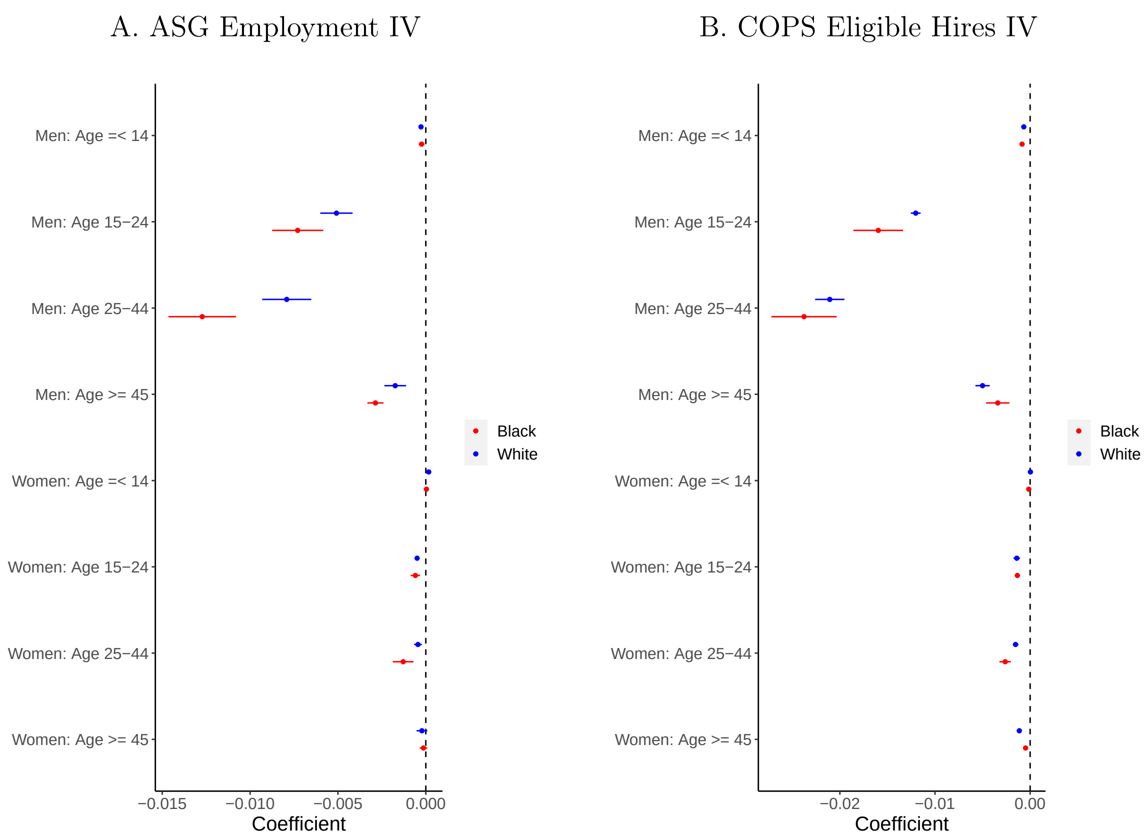
*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Results show outcomes by race group, using the most granular categories available for each outcome data source. FBI UCR arrest records do not include information on Hispanic ethnicity and the Fatal Encounters data includes Hispanic as a distinct race rather than ethnicity group. Baseline specifications correspond to models in Table 2.

Table A8: Results Dis-aggregated by Race Subgroups, COPS Employment IV

| **B. COPS Eligible Hires IV** | (1) Black | (2) Non-Hispanic Black | (3) White | (4) Non-Hispanic White | (5) Hispanic | (6) Other Race |
|---|---|---|---|---|---|---|
| (1) Homicide Victims | -0.0504*** | -0.0500*** | -0.0476*** | -0.0441*** | -0.0055 | -0.0039*** |
| | (0.0044) | (0.0044) | (0.0051) | (0.0009) | (0.0043) | (0.0005) |
| Y-Mean | 131.69 | 129.98 | 85.67 | 59.16 | 28.71 | 5.76 |
| N | 6501 | 6503 | 6495 | 6490 | 6476 | 6467 |
| (2) Clearance Rates | 0.0014 | 0.0014 | 0.0007 | 0.0003 | 0.0210** | -0.0023 |
| | (0.0012) | (0.0012) | (0.0017) | (0.0019) | (0.0084) | (0.0041) |
| Y-Mean | 56.81 | 56.76 | 64.05 | 66.41 | 60.02 | 57.80 |
| N | 4602 | 4600 | 5456 | 5224 | 1734 | 1945 |
| (3) Index Arrests | -1.109*** | | -0.535*** | | | 0.081*** |
| | ( 0.185) | | ( 0.155) | | | ( 0.007) |
| Y-Mean | 7005.5 | | 6135.1 | | | 250.9 |
| N | 5810 | | 5813 | | | 5795 |
| (4) Quality of Life Arrests | 8.099*** | | 13.948*** | | | -0.177*** |
| | ( 1.604) | | ( 3.424) | | | ( 0.039) |
| Y-Mean | 24807.0 | | 24674.4 | | | 565.9 |
| N | 5831 | | 5818 | | | 5804 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Results show outcomes by race group, using the most granular categories available for each outcome data source. FBI UCR arrest records do not include information on Hispanic ethnicity and the Fatal Encounters data includes Hispanic as a distinct race rather than ethnicity group. Baseline specifications correspond to models in Table 2.

Figure A5: Effects of Police Force Size on Homicide: Age, Sex, and Race




Note: Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Figure A covers 1981-2018 ; Figure B covers 1990-2018. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. In Figure A, the instrument is police employment from the U.S. Census; in Figure B the instrument is the number of eligible hires awarded through a COPS Hiring grant. Models include covariates in Table 1; Figure B also controls for non-hiring grant award size and whether a city applied for a hiring or non-hiring grant (lagged).

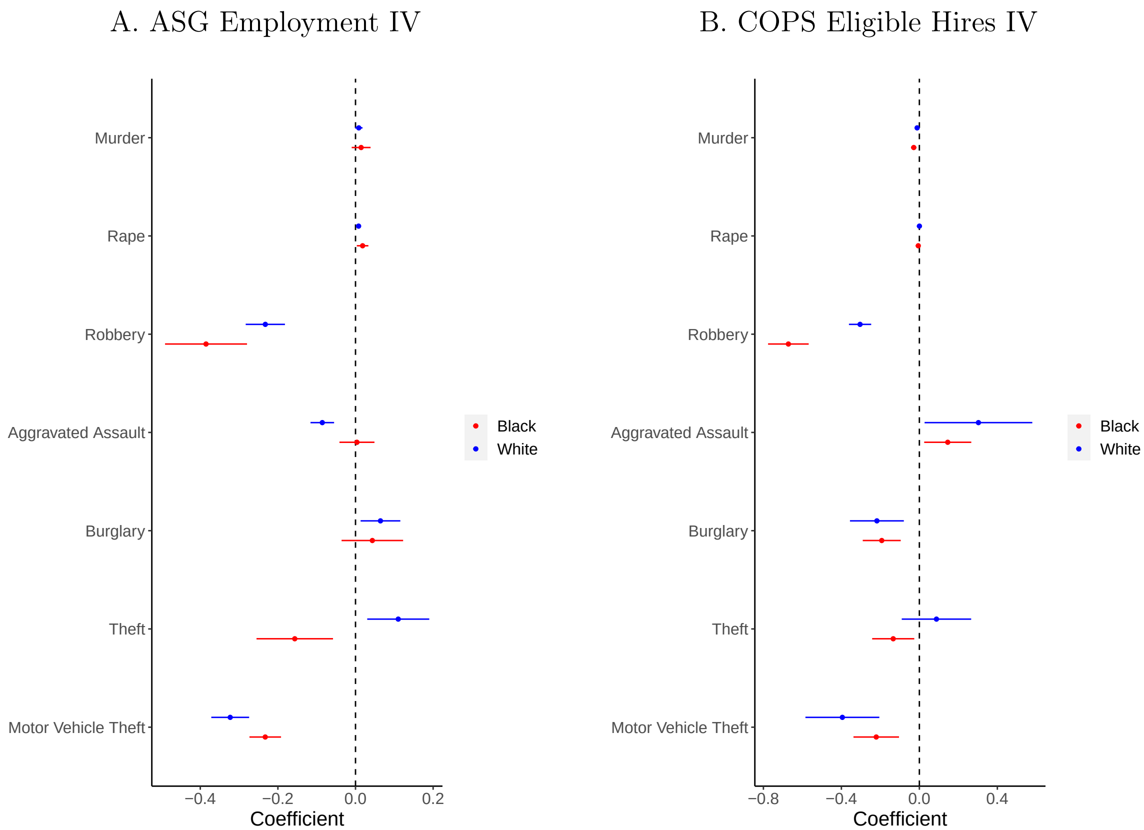
Table A9: Results for Index Crimes and Arrests by Sub-Type

| **A. ASG Employment IV** | Coeff. | S.E. | Elasticity | $\beta$/Population | Mean | N |
|---|---|---|---|---|---|---|
| **Index Crimes** | | | | | | |
| Murder/Manslaughter | -0.069*** | ( 0.004) | -1.59 | -0.004 | 254.1 | 8559 |
| Rape | -0.054*** | ( 0.015) | -0.50 | -0.003 | 633.8 | 8561 |
| Robbery | -3.150*** | ( 0.157) | -1.91 | -0.189 | 10018.6 | 8565 |
| Aggravated Assault | -0.554*** | ( 0.098) | -0.34 | -0.033 | 9997.1 | 8595 |
| Burglary | -4.599*** | ( 0.457) | -1.61 | -0.275 | 17299.9 | 8560 |
| Theft | -5.504*** | ( 0.588) | -0.74 | -0.330 | 45487.9 | 8552 |
| Motor Vehicle Theft | -4.004*** | ( 0.371) | -1.56 | -0.261 | 14138.6 | 8592 |
| **Index Crime Arrests** | | | | | | |
| Murder/Manslaughter | 0.025 | ( 0.018) | 0.56 | 0.002 | 205.3 | 7797 |
| Rape | 0.028*** | ( 0.009) | 0.56 | 0.002 | 232.3 | 7803 |
| Robbery | -0.607*** | ( 0.086) | -1.05 | -0.045 | 2638.4 | 7798 |
| Aggravated Assault | -0.029 | ( 0.034) | -0.04 | -0.002 | 3527.2 | 7828 |
| Burglary | 0.125* | ( 0.072) | 0.29 | 0.009 | 1967.2 | 7794 |
| Theft | 0.023 | ( 0.082) | 0.02 | 0.002 | 6293.0 | 7794 |
| Motor Vehicle Theft | -0.550*** | ( 0.037) | -1.70 | -0.041 | 1478.8 | 7807 |

| **B. COPS Eligible Hires IV** | Coeff. | S.E. | Elasticity | $\beta$/Population | Mean | N |
|---|---|---|---|---|---|---|
| **Index Crimes** | | | | | | |
| Murder/Manslaughter | -0.106*** | ( 0.009) | -2.96 | -0.006 | 221.2 | 6546 |
| Rape | -0.093*** | ( 0.023) | -1.04 | -0.006 | 559.9 | 6554 |
| Robbery | -4.138*** | ( 0.309) | -3.20 | -0.243 | 8305.6 | 6560 |
| Aggravated Assault | -0.851*** | ( 0.267) | -0.57 | -0.050 | 9627.5 | 6585 |
| Burglary | -4.884*** | ( 0.476) | -2.44 | -0.286 | 12899.2 | 6553 |
| Theft | -7.170*** | ( 0.645) | -1.14 | -0.420 | 40592.1 | 6541 |
| Motor Vehicle Theft | -6.443*** | ( 0.575) | -3.10 | -0.421 | 11801.9 | 6577 |
| **Index Crime Arrests** | | | | | | |
| Murder/Manslaughter | -0.041*** | ( 0.009) | -1.14 | -0.003 | 158.1 | 5840 |
| Rape | -0.005 | ( 0.007) | -0.12 | -0.000 | 177.2 | 5842 |
| Robbery | -0.985*** | ( 0.054) | -2.02 | -0.078 | 2139.6 | 5838 |
| Aggravated Assault | 0.487** | ( 0.222) | 0.64 | 0.039 | 3307.4 | 5880 |
| Burglary | -0.405*** | ( 0.131) | -1.28 | -0.032 | 1393.2 | 5828 |
| Theft | 0.003 | ( 0.150) | 0.00 | 0.000 | 5023.4 | 5826 |
| Motor Vehicle Theft | -0.613*** | ( 0.167) | -2.34 | -0.049 | 1146.3 | 5848 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Models correspond to primary specifications for both strategies and are weighted by population of each city in 1980. Panel A covers 1981-2018; Panel B covers 1990-2018. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. The instrument is police employment recorded in the U.S. Census. Models include covariates in Table 1. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). All estimates pass a Bonferroni multiple hypothesis correction of 20, except for the coefficients on "Arrest: Burglary" in Panel A, and "Arrest: Aggravated Assault" in Panel B.

Figure A6: Effects of Police Force Size on Index Arrests by Race



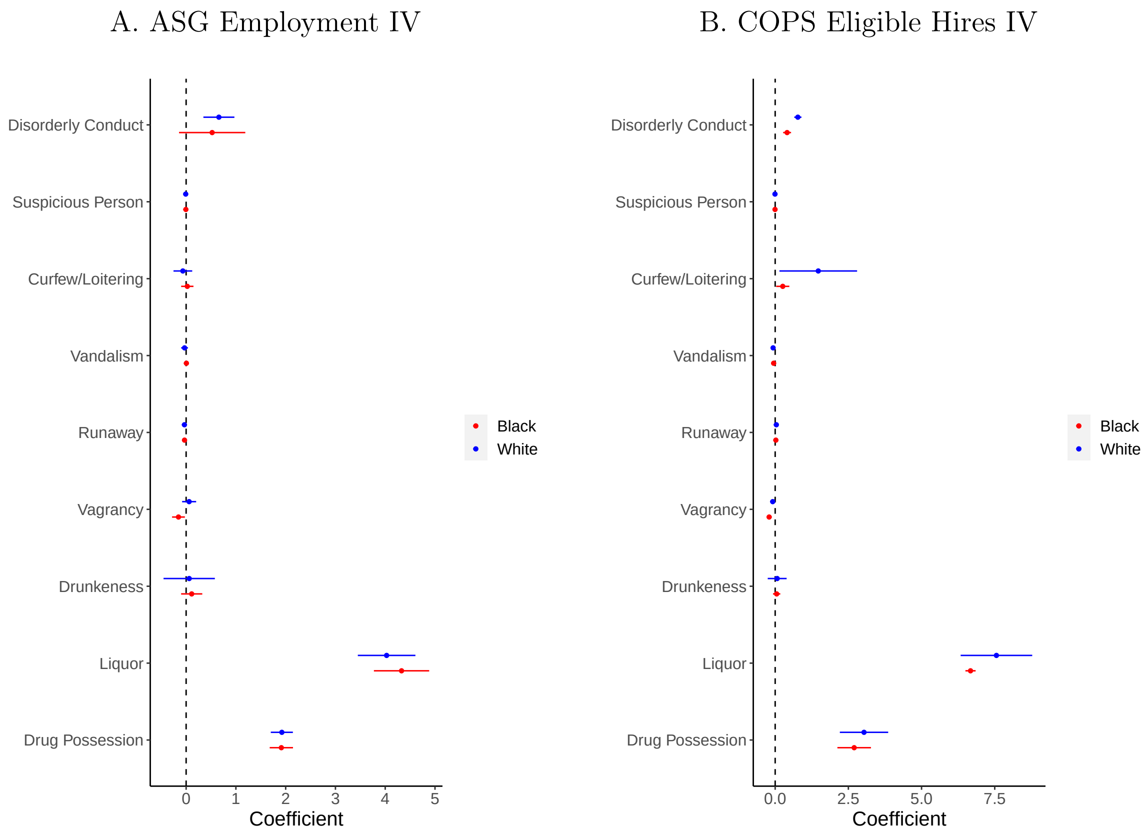

Note: Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Figure A covers 1981-2018; Figure B covers 1990-2018. Arrest categories correspond to Appendix Table A9. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. In Figure A, the instrument is police employment from the U.S. Census; in Figure B the instrument is the number of eligible hires awarded through a COPS Hiring grant. Models include covariates in Table 1; Figure B also controls for non-hiring grant award size and whether a city applied for a hiring or non-hiring grant (lagged).

Table A10: Results by Quality of Life Arrest Sub-Type

| **A. ASG Employment IV** | Coeff. | S.E. | Elasticity | $\beta$/Population | Mean | N |
|---|---|---|---|---|---|---|
| **Quality of Life Arrests** | | | | | | |
| Disorderly Conduct | 1.199*** | ( 0.351) | 0.83 | 0.089 | 6588.9 | 7788 |
| Suspicious Person | -0.011 | ( 0.015) | -1.86 | -0.001 | 28.3 | 7800 |
| Curfew/Loitering | -0.036 | ( 0.111) | -0.16 | -0.003 | 1052.2 | 7791 |
| Vandalism | -0.011 | ( 0.030) | -0.04 | -0.001 | 1452.9 | 7801 |
| Vagrancy | -0.085 | ( 0.096) | -0.63 | -0.006 | 615.7 | 7799 |
| Gambling | 0.332*** | ( 0.028) | 2.40 | 0.025 | 630.9 | 7791 |
| Drunkenness | 0.178 | ( 0.252) | 0.43 | 0.013 | 1869.1 | 7794 |
| Liquor | 8.354*** | ( 0.436) | 7.87 | 0.620 | 4822.1 | 7791 |
| Drug Possession | 3.860*** | ( 0.153) | 2.41 | 0.286 | 7294.4 | 7811 |
| Uncategorized Arrests | -6.532*** | ( 0.767) | -0.83 | -0.484 | 35887.3 | 7818 |
| **B. COPS Eligible Hires IV** | Coeff. | S.E. | Elasticity | $\beta$/Population | Mean | N |
| **Quality of Life Arrests** | | | | | | |
| Disorderly Conduct | 1.196*** | ( 0.141) | 1.20 | 0.095 | 4390.6 | 5831 |
| Suspicious Person | -0.015 | ( 0.023) | -2.80 | -0.001 | 23.8 | 5838 |
| Curfew/Loitering | 1.726** | ( 0.853) | 6.79 | 0.136 | 1115.5 | 5844 |
| Vandalism | -0.109* | ( 0.063) | -0.38 | -0.009 | 1260.9 | 5840 |
| Vagrancy | -0.290*** | ( 0.081) | -2.83 | -0.023 | 448.4 | 5844 |
| Gambling | 0.280*** | ( 0.018) | 2.71 | 0.022 | 455.1 | 5825 |
| Drunkenness | 0.139 | ( 0.244) | 0.41 | 0.011 | 1479.8 | 5831 |
| Liquor | 14.216*** | ( 0.765) | 11.79 | 1.132 | 5230.4 | 5834 |
| Drug Possession | 5.893*** | ( 0.815) | 3.55 | 0.467 | 7259.1 | 5880 |
| Uncategorized Arrests | -1.075 | ( 2.772) | -0.17 | -0.085 | 28131.5 | 5872 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Models correspond to primary specifications for both strategies and are weighted by population of each city in 1980. Panel A covers 1981-2018; Panel B covers 1990-2018. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. The instrument is police employment recorded in the U.S. Census. Models include covariates in Table 1. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). All estimates pass a Bonferroni multiple hypothesis correction of 20, except for the coefficients on "Quality of Life: Curfew/Loitering" and "Quality of Life: Vandalism" in Panel B.

Figure A7: Effects of Police Force Size on Quality of Life Arrests by Race



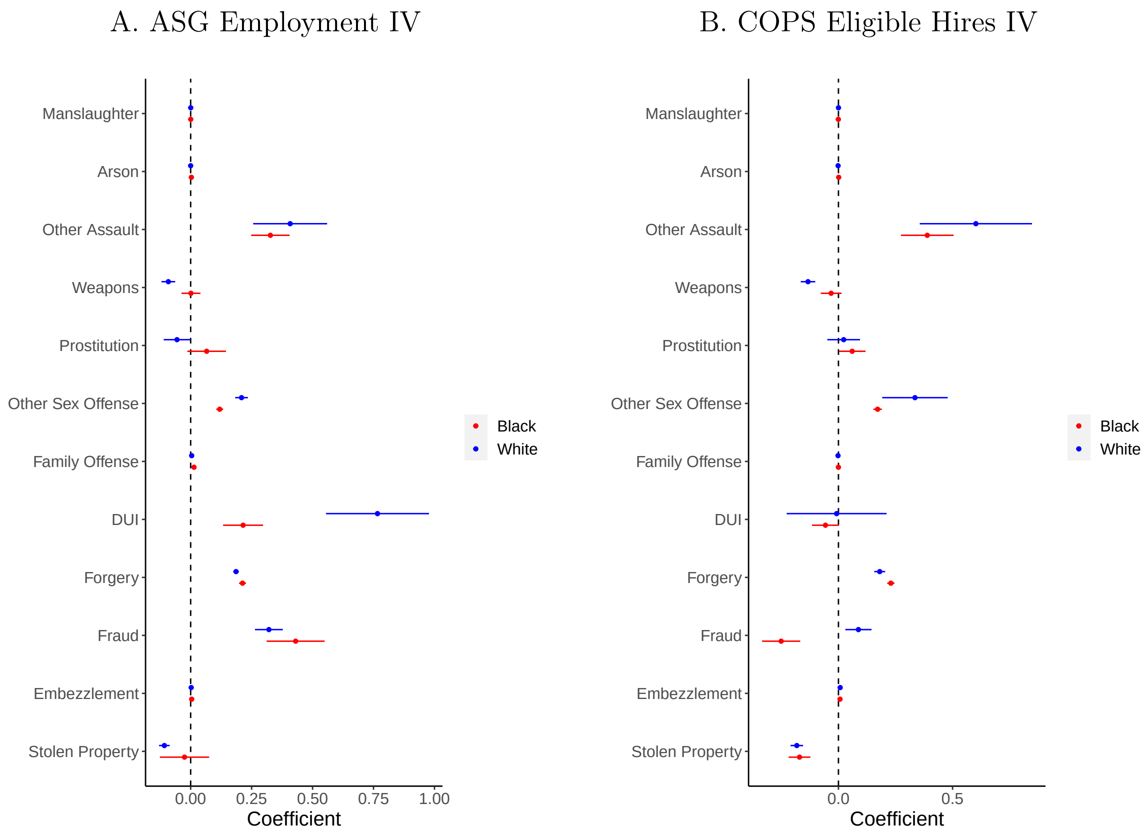

Note: Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Figure A covers 1981-2018; Figure B covers 1990-2018. Arrest categories correspond to Appendix Table A10. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. In Figure A, the instrument is police employment from the U.S. Census; in Figure B the instrument is the number of eligible hires awarded through a COPS Hiring grant. Models include covariates in Table 1; Figure B also controls for non-hiring grant award size and whether a city applied for a hiring or non-hiring grant (lagged).

Table A11: Results by Non-Index Arrest Sub-Type

| **A. ASG Employment IV** | Coeff. | S.E. | Elasticity | $\beta$/Population | Mean | N |
|---|---|---|---|---|---|---|
| **Non-Index Arrests** | | | | | | |
| Negligent Manslaughter | 0.001* | ( 0.001) | 0.62 | 0.000 | 7.3 | 7794 |
| Arson | 0.004 | ( 0.003) | 0.25 | 0.000 | 66.7 | 7795 |
| Other Assault | 0.806*** | ( 0.115) | 0.73 | 0.060 | 4997.6 | 7826 |
| Family Offense | 0.018*** | ( 0.005) | 0.81 | 0.001 | 102.7 | 7792 |
| Weapons | -0.085*** | ( 0.032) | -0.24 | -0.006 | 1631.7 | 7805 |
| Prostitution | 0.045 | ( 0.071) | 0.11 | 0.003 | 1889.3 | 7792 |
| Other Sex Offense | 0.342*** | ( 0.015) | 2.57 | 0.025 | 608.9 | 7795 |
| Runaway | -0.066*** | ( 0.025) | -0.93 | -0.005 | 323.7 | 7799 |
| DUI | 1.038*** | ( 0.157) | 1.54 | 0.077 | 3091.1 | 7794 |
| Drug Sale | 0.276* | ( 0.149) | 0.30 | 0.021 | 4186.5 | 7810 |
| Forgery | 0.432*** | ( 0.013) | 3.93 | 0.032 | 501.7 | 7795 |
| Fraud | 0.800*** | ( 0.099) | 1.49 | 0.059 | 2447.0 | 7806 |
| Embezzlement | 0.007*** | ( 0.003) | 0.77 | 0.001 | 44.2 | 7790 |
| Stolen Property | -0.124* | ( 0.065) | -0.68 | -0.009 | 833.1 | 7801 |
| **B. COPS Eligible Hires IV** | Coeff. | S.E. | Elasticity | $\beta$/Population | Mean | N |
| **Non-Index Arrests** | | | | | | |
| Negligent Manslaughter | 0.000 | ( 0.000) | 0.02 | 0.000 | 6.0 | 5838 |
| Arson | 0.001 | ( 0.002) | 0.08 | 0.000 | 49.2 | 5834 |
| Other Assault | 1.063*** | ( 0.179) | 0.95 | 0.084 | 4902.3 | 5887 |
| Family Offense | -0.003 | ( 0.008) | -0.11 | -0.000 | 99.5 | 5855 |
| Weapons | -0.168*** | ( 0.034) | -0.52 | -0.013 | 1410.2 | 5845 |
| Prostitution | 0.099** | ( 0.049) | 0.33 | 0.008 | 1318.8 | 5842 |
| Other Sex Offense | 0.519*** | ( 0.090) | 4.09 | 0.041 | 558.7 | 5825 |
| Runaway | 0.063 | ( 0.040) | 1.21 | 0.005 | 227.9 | 5837 |
| DUI | -0.064 | ( 0.145) | -0.11 | -0.005 | 2509.4 | 5827 |
| Drug Sale | 0.184 | ( 0.141) | 0.20 | 0.015 | 3986.9 | 5869 |
| Forgery | 0.437*** | ( 0.016) | 3.75 | 0.035 | 511.4 | 5832 |
| Fraud | -0.113 | ( 0.070) | -0.21 | -0.009 | 2298.9 | 5852 |
| Embezzlement | 0.015*** | ( 0.003) | 1.57 | 0.001 | 40.7 | 5854 |
| Stolen Property | -0.350*** | ( 0.038) | -2.49 | -0.028 | 615.0 | 5838 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Models correspond to primary specifications for both strategies and are weighted by population of each city in 1980. Panel A covers 1981-2018; Panel B covers 1990-2018. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. The instrument is police employment recorded in the U.S. Census. Models include covariates in Table 1. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). All estimates pass a Bonferroni multiple hypothesis correction of 20, except for the coefficients on "Non-Index Arrest: Negligent Manslaughter," "Non-Index Arrest: Weapons," "Non-Index Arrest: Runaway," "Non-Index Arrest: Drug Sale," "Non-Index Arrest: Stolen Property" in Panel A and "Non-Index Arrest: Prostitution" in Panel B.

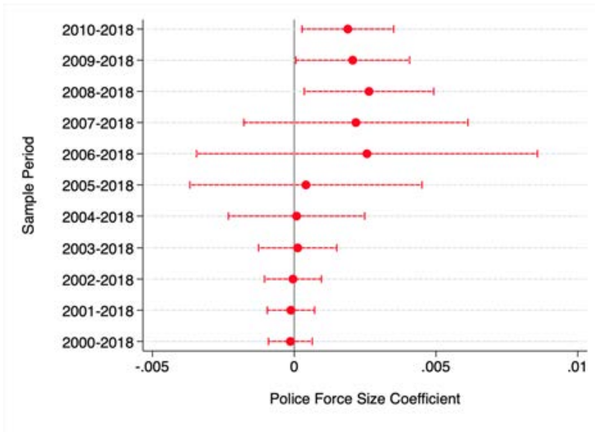
Figure A8: Effects of Police Force Size on Non-Index Arrests by Race



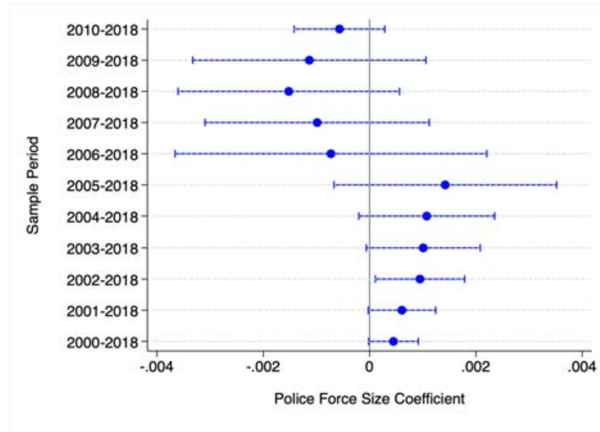

Note: Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Figure A covers 1981-2018; Figure B covers 1990-2018. Arrest categories correspond to Appendix Table A11. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. In Figure A, the instrument is police employment from the U.S. Census; in Figure B the instrument is the number of eligible hires awarded through a COPS Hiring grant. Models include covariates in Table 1; Figure B also controls for non-hiring grant award size and whether a city applied for a hiring or non-hiring grant (lagged).

Figure A9: Varying Sample Window of Fatal Encounters Data: Civilians Shot by Police, ASG IV

A. Black Civilians




B. White Civilians




Note: Standard errors are clustered at the city-level. Each row presents the estimates adding an additional year of Fatal Encounters data. Our preferred estimates are 2010-2018, as years prior to 2013 were compiled retrospectively and are less likely to be comprehensive and precise. Models correspond to those in Table 2. Significance of estimates do not pass a Bonferroni multiple hypothesis correction of 20.

Table A12: Civilians Killed by Police, ASG IV

| **A. ASG Employment IV** | Coeff. | S.E. | Elasticity | $\beta$/Population | Mean | N |
|---|---|---|---|---|---|---|
| **Fatal Encounters (2010-2018)** | | | | | | |
| *(F-test= 19.55)* | | | | | | |
| Cause of Death: Gunshot | 0.0005 | (0.0012) | 0.77 | 0.00003 | 3.920 | 2015 |
| Black | 0.0019** | (0.0008) | 7.06 | 0.00047 | 1.749 | 2002 |
| White | -0.0006 | (0.0004) | -4.34 | -0.00009 | 0.853 | 1999 |
| Cause of Death: Vehicle | -0.0008* | (0.0004) | -5.12 | -0.00005 | 1.047 | 2016 |
| Black | -0.0005 | (0.0004) | -6.19 | -0.00012 | 0.519 | 2005 |
| White | 0.0001 | (0.0001) | 2.30 | 0.00001 | 0.183 | 2005 |
| Cause of Death: Other | 0.0000 | (0.0004) | 0.29 | 0.00000 | 0.556 | 2022 |
| Black | 0.0001 | (0.0003) | 2.40 | 0.00002 | 0.247 | 2009 |
| White | 0.0003 | (0.0002) | 17.63 | 0.00004 | 0.099 | 2004 |
| Total Civilians Killed | -0.0001 | (0.0013) | -0.12 | -0.00001 | 5.549 | 2016 |
| Black | 0.0014* | (0.0008) | 3.59 | 0.00035 | 2.547 | 2003 |
| White | -0.0003 | (0.0005) | -1.44 | -0.00004 | 1.183 | 1999 |
| **SHR Records (2010-2018)** | | | | | | |
| *(F-test= 19.55)* | | | | | | |
| Total Civilians Killed | 0.0015 | (0.0013) | 3.08 | 0.00008 | 3.474 | 1736 |
| Black | 0.0026** | (0.0010) | 11.63 | 0.00060 | 1.546 | 1720 |
| White | -0.0007 | (0.0006) | -2.70 | -0.00010 | 1.708 | 1720 |
| **SHR Records (1980-2018)** | | | | | | |
| *(F-test=559.16)* | | | | | | |
| Total Civilians Killed | -0.0006*** | (0.0001) | -0.72 | -0.00003 | 5.045 | 7673 |
| Black | -0.0002* | (0.0001) | -0.48 | -0.00005 | 2.467 | 7619 |
| White | -0.0004*** | (0.0001) | -1.00 | -0.00006 | 2.429 | 7620 |
| **Adjusted Correlation** | | | | | | |
| **Y=Fatals, X=SHR (2010-2018)** | | | | | | |
| Total Civilians Killed, *(F-test=152.28)* | 0.817*** | ( 0.066) | - | - | 7059.0 | 1727 |
| Black, *(F-test= 88.26)* | 0.768*** | ( 0.082) | - | - | 7084.4 | 1703 |
| White, *(F-test= 21.78)* | 0.282*** | ( 0.060) | - | - | 7085.1 | 1702 |

Note: *p<0.1, **p<0.05, ***p<0.01.
Standard errors are clustered at the city-level. Models are weighted by population of each city in 1980. Models correspond to primary specification Table 2.A. The “Cause of Death: Gunshot” estimates from the Fatal Encounters series correspond to our preferred estimates of civilians killed by police, as these models are most likely to be accurately reported. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. Supplementary Homicide Report (SHR) Records of civilians killed by police are coded using the cause of death variable in the F.B.I. SHR data series. The first stage panel at the bottom of the table regresses the Fatal Encounters measure of total civilian deaths on the SHR measure during the same time period (2010-2018), conditional on demographic covariates, state by year fixed effects, and agency fixed effects. F-tests in parentheses refer to the first stage F-test associated with a regression. The outcome estimates presented do not pass a Bonferroni multiple hypothesis adjustment of 20, except for the coefficient in “Total Civilians Killed” and “Total Civilians Killed, White” in the SHR Records (1980-2018) series.

Table A13: Police Force Size and Net Mortality

| | **Homicides** | | | **Civilians Shot by Police** | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta$ | se($\beta$) | Upper Limit | $\beta$ | se($\beta$) | Upper Limit | Lives Saved Per Life Taken (Upper Bound of 95% CI) | Share of Abated Homicides Offset by Police Shootings |
| Overall | -0.0583 | 0.0043 | -0.0499 | 0.0005 | 0.0012 | 0.0029 | 17.2 | 5.8% |
| Black | -0.0272 | 0.0027 | -0.0219 | 0.0019 | 0.0008 | 0.0035 | 6.3 | 15.9% |
| White | -0.0292 | 0.0018 | -0.0257 | -0.0006 | 0.0004 | 0.0002 | 128.5 | 0.8% |

Note: Table presents estimated coefficients and standard errors for homicides and civilians shot by the police from our main regressions, using the ASG instrument, which are reported in Table 2A. We also report the upper boundary of the 95% confidence intervals for each estimate. We obtain a conservative estimate of the number of lives saved by police for each life taken in a fatal encounter by dividing the upper limit of the confidence interval for homicides by the upper limit of the confidence interval for fatal encounters. Estimates are presented separately by civilian race.

Table A14: Police Force Size and Officer Deaths and Injuries

| **A. ASG IV** | Coeff. | S.E. | Elasticity | Mean | N |
|---|---|---|---|---|---|
| Officer Felonious Deaths | 0.0000 | (0.0000) | 0.42 | 0.224 | 8554 |
| Officers Assault Injuries | -0.1358*** | (0.0108) | -2.83 | 291.4 | 8563 |
| **B. COPS IV** | Coeff. | S.E. | Elasticity | Mean | N |
| Officer Felonious Deaths | -0.0001*** | (0.0001) | -4.92 | 0.158 | 6566 |
| Officers Assault Injuries | -0.2259*** | (0.0061) | -7.14 | 203.6 | 6555 |

*p<0.1, **p<0.05, ***p<0.01.
Note: Standard errors are clustered at the city-level. Models correspond to primary specifications for both strategies and are weighted by population of each city in 1980. Panel A covers 1981-2018; Panel B covers 1990-2018. Officer deaths includes only felonious deaths of officers; and officer injuries include injuries caused by assaults on the job. Models have differing observations due to data availability and the outlier cleaning procedure described in Appendix A3. The endogenous measure of police employment is recorded in the UCR LEOKA files. All estimates pass a Bonferroni multiple hypothesis correction.

# A3 Data Appendix

## A3.1 Data and Procedures

This project compiles data from a number of different public data sources. Below is a description of each data set and the procedures used to clean the data.

**FBI Uniform Crime Report, Law Enforcement Officers Killed or Assaulted (UCR LEOKA)** The principal measure of police manpower used in this paper comes from the FBI's Law Enforcement Officers Killed or Assaulted (LEOKA) series, which has been collected annually since 1960. This data set compiles information on officers that are killed or assaulted in the field as well as total officer employment each year. We access the LEOKA data using Jacob Kaplan's concatenated LEOKA data available from ICPSR (Kaplan, 2019b). These data are used to create the primary police employment measure that is the main focus of the analysis. We define police employment as full time sworn officer employment. We measure officer deaths as deaths that occur as a result of a civilian felony. We measure officer assaults as assaults by civilians that resulted in officer injuries. This dataset covers the period between 1981-2018.

**Annual Survey of Governments, Public Employment and Payroll (ASG Census)** This U.S. Census survey collects data on employment in local governments and is the source of data for the measurement error instrument, or Annual Survey of Governments (ASG) IV. The ASG is an annual survey of municipal employment and payrolls that has been administered by the Bureau of Labor Statistics and reported to the U.S. Census annually since 1952. The ASG data provide annual payroll data for a large number of municipal functions including elementary and secondary education, judicial functions, public health and hospitals, streets and highways, sewerage and police and fire protection, among others.[18] The survey generally provides information on the number of full-time, part-time and full-time equivalent sworn and civilian employees for each function and for each municipal government.

The instrument is a measure of full time sworn police officer employment from this survey. As with the UCR system, the ASG reports a point-in-time measure of police, reporting the number of sworn officers employed as of March 31st of a given year (for 1997-2010 the reference date is June 30th). We linearly interpolate values for years when this data is missing in particular years, including 1996 and 2003, when no survey was collected for any city. This dataset covers the period 1981-2018.

**Department of Justice, Community Oriented Policing Services (COPS) Grants** Data on grants administered by the Department of Justice COPS office was obtained through a Freedom of Information Act (FOIA) request. These grants were established in 1994 through the Violent Crime Control Act (VCCA). Given the coverage period of the grants, the analysis using COPS grants spans the period of 1990-2018. The COPS data includes records of all grants awarded by the office as well records of all applications that were rejected by the office. Grants are divided into grants whose primary purpose is hiring police officers versus grants for other law enforcement needs (non-hiring grants), including investments in technology and targeted crime control. The dollar size

[18] This data surveys all local governments every 5 years and a sub sample of local governments including large cities (covering our sample of cities) every year.

of a grant is available for grants that were awarded and the number of eligible hires designated by a hiring grant is available for hiring grants that were awarded. This data is collapsed to contain records of new hiring and non-hiring grant applications and awards for each city-year in the data. Data covering award amounts are converted into 2018 constant dollars using the consumer price index as an inflator.

**FBI Uniform Crime Report, Supplementary Homicide Report (UCR SHR)** These data include records of homicides as reported to the FBI by police departments. The SHR has been available since 1976 and is the most comprehensive national source of information on the victims and, when available, the perpetrators of homicide (Loftin et al., 2015). We access the SHR data using Jacob Kaplan's concatenated Supplementary Homicide Reports files available from ICPSR (Kaplan, 2019a). We use these data to construct our primary outcomes of total number of homicides each year, as well as homicides by race, gender and age group. Unlike with the UCR Arrest data (below), the category of Hispanic or Latino is available in this dataset. These outcomes are replaced as zeros when missing (but are subject to the outlier cleaning described below). We exclude homicides where the civilian was killed by a police officer, as well as homicides where the person killed was engaging in a felony and killed by a private civilian and homicides that occur in institutional settings such as prisons. These data are also used to construct our measure of homicide clearance rates. We code a homicide as being "cleared" if demographic information for the suspect of the homicide is available in the SHR, which permits the construction of clearance rates separately by victim race. This data covers the period 1981-2018.

**FBI Uniform Crime Report, Arrest Data (UCR Arrest)** This data set includes records of arrests for different types of offenses as submitted by city agencies. We access these data using Jacob Kaplan's concatenated offenses known and clearances by arrest files available from ICPSR (Kaplan, 2019c). These data have been collected annually at the agency-level since 1974. The data includes records of total arrests, and arrests by the race of the civilian (e.g. Black or white), where the category of Hispanic or Latino is not available. We extract records of individual crime category arrests, total and by race, as well as construct larger group categories of arrests by type (see Appendix Tables A9, A10, and A11 for groupings). Before constructing these sums, we replace any negative arrest values as missing. In several cases, an individual crime category may be missing for a particular year or city, when this happens we treat this value as a zero in the sum. Our procedure that identifies outliers (see below) helps identify cases when this approach might create large fluctuations in the data over time. This data set covers the period 1981-2018.

**Fatal Encounters Data (Civilians Shot by Police)** We utilize the Fatal Encounters data to measure civilians shot by police, a data set that is collected and maintained by journalist D. Brian Burghart (Edwards et al., 2018; Goh, 2020). The Fatal Encounters data are collected via three methods: 1) public records requests made by journalists to law enforcement agencies, 2) directed internet searches by volunteers and paid researchers and 3) cross-referencing data with similar enterprises launched in recent years by *The Guardian* and *The Washington Post*. As noted by Goh (2020), the Fatal Encounters data carries two key advantages over other crowd-sourced data sets. First, the number of years for which information is available is greater than that of other well-known crowd-sourced data sets, given that many online data sets track police killings from only 2014 or later. Second, the there is a rigorous process to validate the data (Finch et al., 2019). In research by Ozkan et al. (2018), a comparison of records of fatal officer-involved shootings from the Dallas Police Department with crowd-sourced data sets, the Fatal Encounters data mostly closely tracked the official records. The site was established in 2013 and tracks data going back to 2000. We focus on the 2010-2018 period as early data have been found to be less reliable (Goh, 2020).

City-level outcomes are replaced as zeros when missing (but are subject to the outlier cleaning described below). We exclude events where the cause of death was suicide or the location was missing, and utilize the imputed race variable to determine race subgroups in this data. Fatal encounters includes Hispanic or Latino as a race option, unlike the UCR data sources, and this is coded as a category separate from white and Black in our analysis.

**Annual Survey of Governments, Local Government Finances (Census)** This U.S. Census survey collects data on local government finances, tax collection, and spending. With a few exceptions, the Census Bureau has conducted an Annual Survey of Government Finances in every year since 1902. Like the Annual Survey of Public Employment and Payroll, this survey covers all local governments every 5 years and a sub-sample of local governments (including large cities) every year (covering our sample). Like the data on employees and payroll, data on government expenditures are reported separately for a large number of municipal functions, including elementary and secondary education, judicial functions, public health and hospitals, streets and highways, sewerage, police and fire protection among others. For each function, expenditures are divided among three categories of spending: (1) current operations,(2) capital expenditures and (3) expenditures on construction. The data are reported annually in dollars and, as such, we convert all dollar figures into 2018 constant dollars using the consumer price index as an inflator.

We use this resource to gather data on total government expenditures, taxes, and revenue, which we include as controls in our preferred specifications. This data covers the period of 1981-2018. Similar to the Census covariates and employment variables, we linearly interpolate the expenditure variables when missing.

**U.S. Census and American Community Survey (Census)** We collect information from the U.S. Census on a vector of time-varying covariates upon which to condition in all subsequent models. The data we collect includes each city's population, the resident share in each age group (<14, 15-24, 25-44, >45), share male, share Black, white and Hispanic, the share of residents never married, the share of female headed households, the poverty rate, median household income, and the unemployment rate. Since 2000, we can obtain annual measures for each of these variables from the American Communities Survey; prior to 2000 we use the decennial Census and, following Levitt (1996) and Chalfin and McCrary (2018) among others, linearly interpolate between Census years.

## A3.2 Identifying Outliers

UCR crime data sets are voluntarily reported by police departments and are known for having issues with reporting and measurement. Further, mass homicide events, while rare, can create large volatile swings in homicide outcomes. We follow prior papers using UCR data that clean these outcomes for outliers (Evans and Owens, 2007; Mello, 2019; Weisburst, 2019b). Specifically, we separately regress the set of outcomes on a polynomial cubic time trend for each city and calculate the percent deviation of the actual value from the values predicted by this regression (the outcomes used for this exercise are the raw values plus one, given the large number of zeros in homicide data). The Civilians Shot by Police uses a polynomial squared time trend instead given its shorter panel.We then summarize the absolute value of these percent deviations within city population groups (of 50k-100k, 100k-250k and >250k residents in 1980) and replace the value as missing if it is greater than the 99th percentile of this distribution or 50%, whichever is larger. This procedure is used for all outcomes as well as the UCR measure of police employment, the Census expenditure variables and the Census ASG police employment instrument. We clean sub-groups of outcomes, such as arrest sub-types or race sub-groups using this procedure as a first step, but also replace these sub-groups as missing if the total associated with a sub-group is identified as an outlier.

In addition to using this general algorithm correction, we pay particular attention to correcting outliers in our largest city, New York. We manually impute the UCR police employment measure for 2003, which represents over 2,000 reduction in sworn police officers in New York in that year, that is recovered the following year (identified in Chalfin and McCrary (2018)).

### A3.3 Other Cleaning and Sample Restrictions

We merge our data sets together using the UCR police department identifier and the crosswalk to census identifiers. Our data set includes only the 242 large cities that regularly report to the Census Annual Survey of Local Government Finances and Annual Survey of Public Employment and Payroll. These cities all have populations that exceed 50,000 in 1980.

The final panel is not balanced. This can occur because of outliers that are replaced as missing (see above), or impartial panels in the source data sets. We use the imbalanced panel to capture as much information as possible in the estimation and to increase power.