# ONLINE APPENDIX

Police Force Size and Civilian Race

Aaron Chalfin

Benjamin Hansen

Emily K. Weisburt

Morgan C. Williams, Jr.

## A1 Identification Strategy

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

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

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). We pursue two different instrumental variables strategies in order to obtain a plausibly consistent estimate of $\beta$. 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 FBI Uniform Crime Report (UCR) 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 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 $\beta$, 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), this 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:

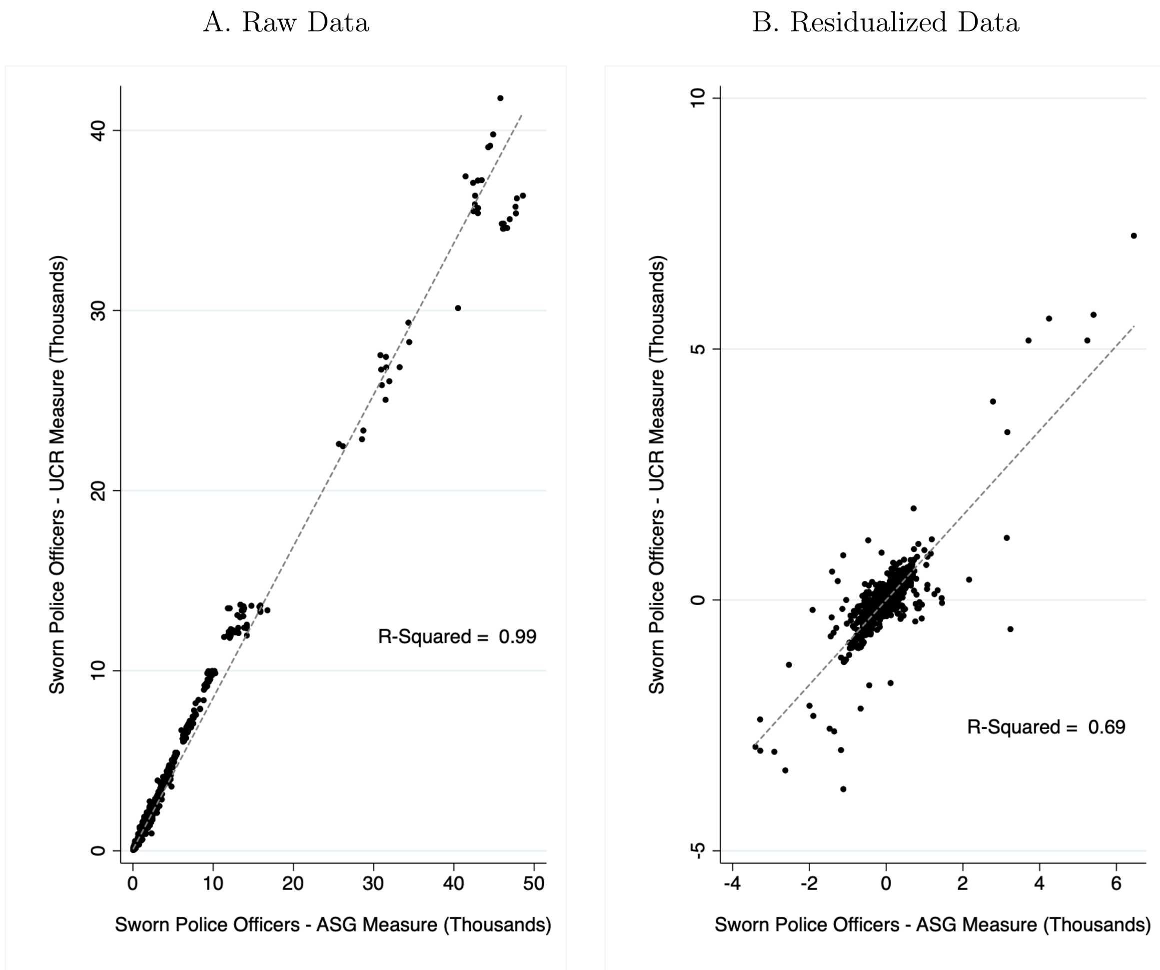
$$S_{it} = S^*_{it} + u_{it} \tag{2}$$

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

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

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

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.

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 $\beta$, under the assumptions of the classical measurement error model:

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

In (5), $\sigma_u^2$ 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 $\sigma_u^2 > 0$, a least squares estimate of $\beta$ 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 $\tilde{S}_{it}$ on $\tilde{Z}_{it}$ is given by:

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

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 $\beta$, suggesting a role for an instrument.

Table 1: Test of the Equality of Forward and Reflected IV Estimates

| | Forward Coeff. | S.E. | Reflected Coeff. | S.E. | Difference: P-Value |
|---|---|---|---|---|---|
| Homicide Victims | -0.058 | ( 0.004) | -0.064 | ( 0.005) | 0.370 |
| Black | -0.026 | ( 0.003) | -0.027 | ( 0.003) | 0.757 |
| White | -0.016 | ( 0.002) | -0.010 | ( 0.001) | 0.025 |
| Homicide Clearance Rate | 0.001 | ( 0.001) | 0.000 | ( 0.001) | 0.606 |
| Black | 0.001 | ( 0.001) | 0.000 | ( 0.001) | 0.750 |
| White | -0.001 | ( 0.001) | -0.001 | ( 0.001) | 0.666 |
| Quality of Life Arrests | 7.12 | ( 0.88) | 6.29 | ( 0.75) | 0.471 |
| Black | 2.15 | ( 0.51) | 1.03 | ( 0.57) | 0.143 |
| White | 5.03 | ( 0.50) | 5.51 | ( 0.25) | 0.388 |
| Index Arrests | -0.97 | ( 0.28) | -0.87 | ( 0.26) | 0.789 |
| Black | -0.68 | ( 0.20) | -0.63 | ( 0.19) | 0.860 |
| White | -0.45 | ( 0.09) | -0.40 | ( 0.09) | 0.716 |
| Index Crimes | -17.82 | ( 1.40) | -20.40 | ( 1.22) | 0.165 |

Note: Table reports coefficients from the "forward" and "reflected" IV regressions derived from equation (1) 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 to test the equality of $\beta$ from an IV regression in which $S$ is instrumented using $Z$ and $\beta$ 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:

$$g_i(\beta) = \begin{pmatrix} Z_{it}(Y_{it} - \beta_1 S_{it} - \gamma_1^* X_{it} \\ X_{it}(Y_{it} - \beta_1 S_{it} - \gamma_1^* X_{it} \\ S_{it}(Y_{it} - \beta_2 Z_{it} - \gamma_2^* X_{it} \\ X_{it}(Y_{it} - \beta_2 Z_{it} - \gamma_2^* X_{it} \end{pmatrix} \tag{7}$$

We then test the pooling restriction that $\beta_1 = \beta_2$. The results of this exercise are available in Appendix Table 1 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.[13]

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 $\beta_1 = \beta_2$. 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.

[13]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.[14] 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 funding for 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).

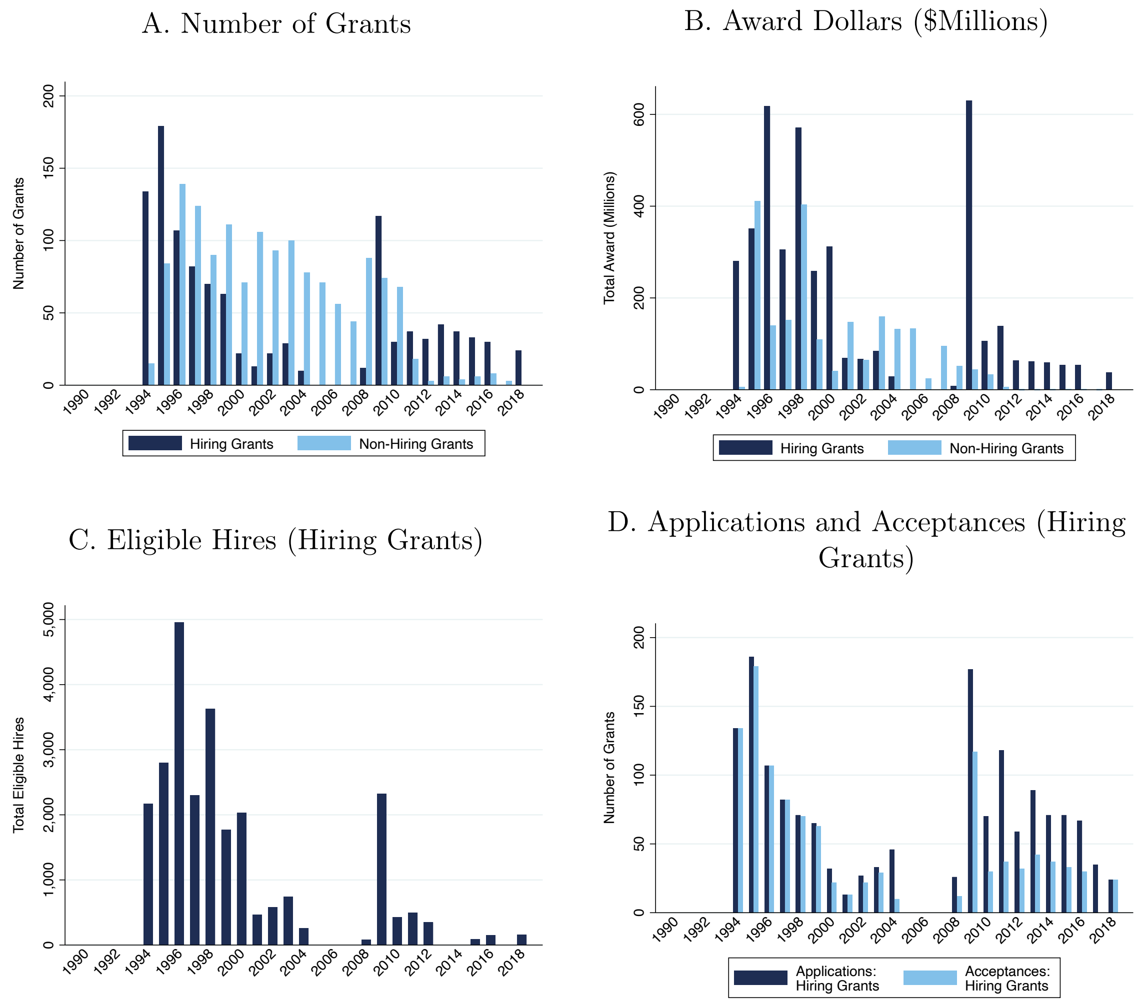
This paper is also able to exploit variation in grant applications that are rejected in the estimation model. Appendix Figure A2.D shows the number of grant applications and acceptances in each year of the COPS program within our sample. Prior to 2000, nearly all applications for hiring grants were awarded.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

[14] 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.

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

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:

$$
\begin{aligned}
Y_{it} =& \beta Police_{it-1} + \gamma_1 AwardNonHiring_{it-1} \\
& + \gamma_2 ApplyHiring_{it-1} + \gamma_3 ApplyNonHiring_{it-1} \\
& + \gamma^{'} X_{it} + \rho_i + \psi_{st} + \varepsilon_{it}
\end{aligned}
$$

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

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.[15] 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 2. 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.

[15] 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 2: 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** | -0.1030 | -0.0503 | -0.0442 | 22.013 | 8.169 | 14.015 |
| *(First Stage F-Test = 15.71)* | (0.0104) | (0.0047) | (0.0010) | ( 5.087) | ( 1.642) | ( 3.473) |
| *Race Difference: P-Value* | | | *0.205* | | | *0.128* |
| $\beta$/Pop | -0.006 | -0.012 | -0.008 | 1.74 | 2.80 | 1.66 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.102* |
| Y-Mean | 223.35 | 130.03 | 59.17 | 49908.0 | 24807.0 | 24674.4 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |
| **(2) Grants Split Across Grant Years** | -0.1871 | -0.0837 | -0.0725 | 45.777 | 16.668 | 29.100 |
| *(First Stage F-Test = 60.07)* | (0.0188) | (0.0103) | (0.0105) | ( 5.498) | ( 1.421) | ( 5.093) |
| *Race Difference: P-Value* | | | *0.445* | | | *0.019* |
| $\beta$/Pop | -0.011 | -0.021 | -0.012 | 3.62 | 5.71 | 3.44 |
| *Race Difference: P-Value* | | | *0.003* | | | *0.004* |
| Y-Mean | 223.35 | 130.03 | 59.17 | 49908.0 | 24807.0 | 24674.4 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |
| **(3) Cities that Applied for Grants** | -0.1030 | -0.0503 | -0.0442 | 22.016 | 8.171 | 14.016 |
| *(First Stage F-Test = 15.70)* | (0.0104) | (0.0047) | (0.0010) | ( 5.088) | ( 1.642) | ( 3.474) |
| *Race Difference: P-Value* | | | *0.204* | | | *0.128* |
| $\beta$/Pop | -0.006 | -0.012 | -0.008 | 1.74 | 2.79 | 1.65 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.102* |
| Y-Mean | 223.65 | 130.20 | 59.25 | 50006.3 | 24856.8 | 24722.6 |
| N | 6502 | 6473 | 6461 | 5805 | 5797 | 5784 |

Table 2: Additional Robustness Specifications, COPS 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 |
|---|---|---|---|---|---|---|
| **(4) Cities with Accepted Grant** | -0.1028 | -0.0503 | -0.0442 | 21.901 | 8.123 | 13.948 |
| *(First Stage F-Test = 16.19)* | (0.0101) | (0.0046) | (0.0010) | ( 5.018) | ( 1.619) | ( 3.430) |
| *Race Difference: P-Value* | | | *0.198* | | | *0.125* |
| $\beta$/Pop | -0.006 | -0.012 | -0.007 | 1.71 | 2.74 | 1.63 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.100* |
| Y-Mean | 225.81 | 131.46 | 59.82 | 50571.0 | 25142.8 | 24998.4 |
| N | 6330 | 6302 | 6292 | 5636 | 5628 | 5615 |
| **(5) Cities with Accepted & Rejected Grants** | -0.1219 | -0.0599 | -0.0460 | 27.969 | 9.890 | 18.225 |
| *(First Stage F-Test = 4.90)* | (0.0271) | (0.0128) | (0.0019) | ( 7.187) | ( 1.917) | ( 5.313) |
| *Race Difference: P-Value* | | | *0.281* | | | *0.140* |
| $\beta$/Pop | -0.006 | -0.013 | -0.007 | 1.94 | 3.08 | 1.87 |
| *Race Difference: P-Value* | | | *0.015* | | | *0.138* |
| Y-Mean | 237.52 | 130.49 | 65.78 | 60307.1 | 29672.2 | 30154.9 |
| N | 4711 | 4688 | 4684 | 4269 | 4263 | 4253 |
| **(6) Drop Application Controls** | -0.1039 | -0.0508 | -0.0452 | 21.807 | 8.031 | 13.917 |
| *(First Stage F-Test = 14.49)* | (0.0104) | (0.0049) | (0.0012) | ( 4.983) | ( 1.596) | ( 3.417) |
| *Race Difference: P-Value* | | | *0.270* | | | *0.119* |
| $\beta$/Pop | -0.006 | -0.012 | -0.008 | 1.73 | 2.75 | 1.65 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.104* |
| Y-Mean | 223.35 | 130.03 | 59.17 | 49908.0 | 24807.0 | 24674.4 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using the number of eligible hires awarded through a COPS Hiring grant. Baseline specifications correspond to models in Table 3. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). FBI 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. Standard errors are clustered at the city-level.

### A1.2.3 Reduced Form Results Over Time

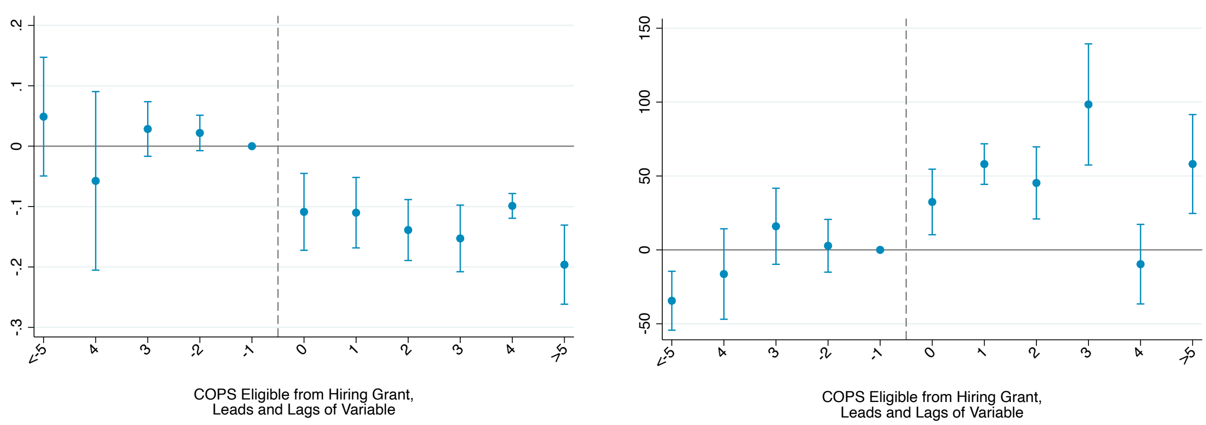
As an additional check of the COPS instrument, we present the reduced form results of the model over time. This exercise directly relates the number of COPS eligible hires to our primary outcomes in the years preceding a grant award. To do this, we construct lead variables of the Cops Eligible Hires IV for the four preceding periods (t=-4,-3,-2, -1 omitted) and lag variables of the IV (t=0,1,2,3,4) 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 3. The 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, but remain smaller in magnitude.[16]

### A2.2 Robustness

#### A2.2.1 Common Sample Results

We begin by re-estimating our results on the common sample of city-years for which all dependent variables are non-missing in the data. These estimates are presented in Table 4 and Table 5. In all cases, estimates are substantively similar to those reported in the main analysis. We utilize a sample that does not fully overlap in dependent variables for our baseline specification in order to use the maximal amount of information and increase power.

#### A2.2.2 Alternative Specifications

Next, in Tables 6 and 7 we further test the robustness of the results. First, we re-estimate our models without using population weights (2). 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 (3). 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 assurance that the estimates reported in the main body of the paper are robust to this alternative employment measure.

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 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 each case, estimates are nearly identical to those reported in Tables 2 and 3.

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

[16] 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.

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 ASG IV models, 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. 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).[17] Using the COPS IV, there is no first stage when the model is specified in log-log form in this set of cities. 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).

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.

#### A2.2.3 Police Force Size and Reporting of Arrests

Next we explicitly consider whether larger police forces could change patterns of reporting crime and arrests, particularly in terms of the large increases we observe in low-level or "quality of life" arrests. There are generally four reasons reported arrests could increase when police force size increases:

1. There is an increase in criminality when the police force expands.
2. There is an increase officer enforcement of offenses when the police force expands.
3. The UCR reports only the "highest offense" for any incident. This means that a reduction in serious charges that had previously been coupled with less serious charges could result in a mechanical increase in lower level offenses.
4. There is a change in police reporting of crime when the police force expands.

The first point is not consistent with the large decreases in homicides we observe. The second point is our leading primary hypothesis. The third hypothesis is unlikely because the increases in

[17] It is worth noting 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.

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 hypothesis in Table 8.

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 changes as a police force expands. In particular, we are concerned with the possibility that increased manpower allows departments to record minor offenses to the FBI that it previously did not disclose.

We do not believe that this hypothesis is driving the results for several reasons. 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) that is related to reporting protocols is accounted for with that control. Second, 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 reduction in this category that could be offset by better categorization of other arrests.

However, it could still be that as resources (and officers) become more plentiful, departments record better records. To address this, in Table 8, we re-estimate the main models for low-level arrests. In panel (1) we provide our main estimates for comparison. In the next panel (2), we present estimates for the same models, expect now dropping all observations in which there were zero observations in the low-level or "quality of life" category. Essentially the results are unchanged, suggesting that the extensive margin of reporting particular sub-offenses is not driving the results. In the final two panels we explore whether the extensive margin crime reporting changes for departments for arrests subgroups. In Panel (3), we measure the outcome of whether a police department reported any low-level arrests in a particular year. In Panel (4), we measure the outcome of whether a police department reported at least one arrest in all sub-categories of the low-level arrest group. We find generally the estimated relationships are small, suggestive there are not large increases in reporting due to increases in police reporting.

#### A2.2.4 Sensitivity of Results

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

#### 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 groupings for these outcomes 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 9 and Appendix Table 10. 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. Arrest outcomes cannot be split by Hispanic/Latinx ethnicity and these individuals are classified as white in this data set.

#### A2.3.2 Arrest Outcomes by Offense Type

**Index Crime and 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 11. 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.

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, 0.05-0.1 rapes, 0.6-0.8 aggravated assaults, 3-4 robberies, 4-5 burglaries, 5-7 thefts and 4-6 motor vehicle thefts. With respect to arrests, larger police forces lead to significantly fewer arrests for robbery and motor vehicle theft. 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).

Appendix Figure A5 explores heterogeneity in the arrest estimates by race, using the per capita estimates. Our aggregate finding that per capita declines in index arrests are larger for Black vs. white arrestees is driven by disparate race group effects for robbery, theft and motor vehicle theft arrests.

**"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 12. Aside from "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 increases of 0.3-0.5 arrests per 100,000 residents (or 6 to 8 total arrests) for every additional officer hired.

Appendix Figure 1 displays the race heterogeneity for each sub-offense using the per capita estimates. There are large and highly significant race disparities in liquor violation and drug possession arrests for both strategies, whereby arrests of Black individuals disproportionately increase in per capita terms. The opposite pattern is present for "uncategorized arrests," leading to total race differences in this aggregate category that are marginally significant using the COPS strategy and not significant for the ASG strategy.

**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, though the effects vary by strategy. Likewise, the race differences shown in Appendix Figure A6 also differ across strategies.

### A2.4 Deaths and Injuries of Police Officers

We also estimate the effect of police force size on violence against police officers. These results are presented in Appendix Table 14. 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 3: OLS Model Results

| | Coeff. | S.E. | $\beta$/Pop. | S.E. | Mean | N |
|---|---|---|---|---|---|---|
| **Homicides** | | | | | | |
| Victims | -0.051 | ( 0.004) | -0.003 | ( 0.000) | 249.0 | 8581 |
| Black | -0.022 | ( 0.002) | -0.005 | ( 0.001) | 140.5 | 8550 |
| White | -0.008 | ( 0.001) | -0.001 | ( 0.000) | 65.5 | 8530 |
| *Difference: P-Value* | *0.000* | | *0.000* | | | |
| Clearance Rate | 0.000 | ( 0.001) | - | - | 65.2 | 7698 |
| Black | 0.000 | ( 0.001) | - | - | 62.5 | 6087 |
| White | -0.001 | ( 0.001) | - | - | 69.4 | 7069 |
| *Difference: P-Value* | *0.264* | | | | | |
| **Arrests** | | | | | | |
| Quality of Life | 5.85 | ( 0.70) | 0.43 | ( 0.05) | 60121 | 7824 |
| Black | 0.96 | ( 0.53) | 0.30 | ( 0.16) | 30843 | 7788 |
| White | 5.12 | ( 0.24) | 0.56 | ( 0.03) | 28758 | 7799 |
| *Difference: P-Value* | *0.000* | | *0.113* | | | |
| Index | -0.80 | ( 0.24) | -0.06 | ( 0.02) | 16342 | 7816 |
| Black | -0.59 | ( 0.18) | -0.18 | ( 0.05) | 8933 | 7773 |
| White | -0.37 | ( 0.08) | -0.04 | ( 0.01) | 7202 | 7790 |
| *Difference: P-Value* | *0.276* | | *0.012* | | | |
| **Index Crimes** | -16.33 | ( 0.85) | -0.98 | ( 0.05) | 96791 | 8675 |

Note: Table reports estimates from equation (1) in which each outcome is regressed on the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports conditional on fixed effects and covariates. 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). FBI 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, except for the coefficient on "Quality of Life Arrests, Black."

Table 4: Common Outcome Sample: Marginal Impact of Police Employment
ASG Employment IV

| | A. ASG IV | | | | | |
|---|---|---|---|---|---|---|
| | Coeff. | S.E. | $\beta$/Pop. | S.E. | Mean | N |
| **First Stage** | | | | | | |
| Police Employment | 0.925 | ( 0.029) | - | - | 4600.4 | 7511 |
| *(F-Test = 1047.27)* | | | | | | |
| **Homicides** | | | | | | |
| Victims | -0.079 | ( 0.009) | -0.006 | ( 0.001) | 245.3 | 7511 |
| Black | -0.039 | ( 0.005) | -0.012 | ( 0.002) | 132.1 | 7511 |
| White | -0.042 | ( 0.006) | -0.008 | ( 0.001) | 66.9 | 7511 |
| *Difference: P-Value* | *0.695* | | *0.016* | | | |
| **Arrests** | | | | | | |
| Quality of Life | 6.83 | ( 0.85) | 0.50 | ( 0.06) | 61548 | 7511 |
| Black | 2.05 | ( 0.50) | 0.63 | ( 0.15) | 31431 | 7511 |
| White | 4.98 | ( 0.49) | 0.54 | ( 0.05) | 29401 | 7511 |
| *Difference: P-Value* | *0.000* | | *0.566* | | | |
| Index | -1.02 | ( 0.26) | -0.07 | ( 0.02) | 16554 | 7511 |
| Black | -0.71 | ( 0.19) | -0.22 | ( 0.06) | 8970 | 7511 |
| White | -0.47 | ( 0.09) | -0.05 | ( 0.01) | 7339 | 7511 |
| *Difference: P-Value* | *0.243* | | *0.004* | | | |
| **Index Crimes** | -19.87 | ( 2.52) | -1.45 | ( 0.18) | 95511 | 7511 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using an alternative measure of sworn officers from the U.S. Census. Standard errors are clustered at the city-level. This table replicates Table 2 on a common sample that includes data on all outcomes. Clearance outcomes are excluded from this exercise as data is more restricted for this outcome. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). FBI 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. Standard errors are clustered at the city-level. All estimates pass a Bonferroni multiple hypothesis correction.

Table 5: Common Outcome Sample: Marginal Impact of Police Employment
COPS Eligible Hires IV

| | B. COPS IV | | | | | |
|---|---|---|---|---|---|---|
| | Coeff. | S.E. | $\beta$/Pop. | S.E. | Mean | N |
| **First Stage** | | | | | | |
| Police Employment | 2.452 | ( 0.526) | - | - | 4412.3 | 5608 |
| *(F-Test = 21.76)* | | | | | | |
| **Homicides** | | | | | | |
| Victims | -0.117 | ( 0.007) | -0.009 | ( 0.001) | 208.3 | 5608 |
| Black | -0.058 | ( 0.003) | -0.020 | ( 0.001) | 114.2 | 5608 |
| White | -0.057 | ( 0.003) | -0.012 | ( 0.001) | 58.8 | 5608 |
| *Difference: P-Value* | *0.818* | | *0.000* | | | |
| **Arrests** | | | | | | |
| Quality of Life | 22.01 | ( 5.21) | 1.72 | ( 0.41) | 51016 | 5608 |
| Black | 8.17 | ( 1.70) | 2.80 | ( 0.58) | 25243 | 5608 |
| White | 14.01 | ( 3.50) | 1.64 | ( 0.41) | 25201 | 5608 |
| *Difference: P-Value* | *0.134* | | *0.103* | | | |
| Index | -1.63 | ( 0.38) | -0.13 | ( 0.03) | 13460 | 5608 |
| Black | -1.15 | ( 0.21) | -0.39 | ( 0.07) | 6964 | 5608 |
| White | -0.56 | ( 0.18) | -0.07 | ( 0.02) | 6243 | 5608 |
| *Difference: P-Value* | *0.038* | | *0.000* | | | |
| **Index Crimes** | -26.31 | ( 1.81) | -2.06 | ( 0.14) | 76620 | 5608 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using the number of eligible hires awarded through a COPS Hiring grant. This table replicates Table 3 on a common sample that includes data on all outcomes. Clearance outcomes are excluded from this exercise as data is more restricted for this outcome. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). FBI 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. Standard errors are clustered at the city-level. All estimates pass a Bonferroni multiple hypothesis correction.

Table 6: 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** | -0.0580 | -0.0258 | -0.0156 | 7.120 | 2.147 | 5.031 |
| *(First Stage F-Test = 553.38)* | (0.0043) | (0.0027) | (0.0018) | ( 0.876) | ( 0.511) | ( 0.500) |
| *Race Difference: P-Value* | | | *0.002* | | | *0.000* |
| $\beta$/Pop | -0.003 | -0.006 | -0.002 | 0.53 | 0.66 | 0.55 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.498* |
| Y-Mean | 249.05 | 140.36 | 65.55 | 60243.6 | 30896.3 | 28827.3 |
| N | 8553 | 8522 | 8502 | 7804 | 7768 | 7779 |
| **(2) Not Weighted by Population** | -0.0504 | -0.0228 | -0.0110 | 8.735 | 3.133 | 5.686 |
| *(First Stage F-Test = 45.41)* | (0.0118) | (0.0069) | (0.0041) | ( 1.743) | ( 1.412) | ( 0.526) |
| *Race Difference: P-Value* | | | *0.141* | | | *0.090* |
| $\beta$/Pop | -0.018 | -0.036 | -0.008 | 3.25 | 5.39 | 3.00 |
| *Race Difference: P-Value* | | | *0.013* | | | *0.325* |
| Y-Mean | 39.21 | 22.92 | 9.94 | 8483.7 | 3796.1 | 4565.0 |
| N | 8553 | 8522 | 8502 | 7804 | 7768 | 7779 |
| **(3) ASG as Endogenous X, UCR as IV** | -0.0636 | -0.0271 | -0.0105 | 6.290 | 1.027 | 5.514 |
| *(First Stage F-Test = 4997.01)* | (0.0046) | (0.0029) | (0.0014) | ( 0.749) | ( 0.568) | ( 0.254) |
| *Race Difference: P-Value* | | | *0.000* | | | *0.000* |
| $\beta$/Pop | -0.004 | -0.007 | -0.002 | 0.47 | 0.32 | 0.60 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.185* |
| Y-Mean | 249.05 | 140.36 | 65.55 | 60243.6 | 30896.3 | 28827.3 |
| N | 8553 | 8522 | 8502 | 7804 | 7768 | 7779 |

Table 6: 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 |
|---|---|---|---|---|---|---|
| **(4) Police Employment not Lagged** | -0.0631 | -0.0299 | -0.0203 | 10.472 | 3.816 | 6.780 |
| *(First Stage F-Test = 550.26)* | (0.0049) | (0.0029) | (0.0034) | ( 0.825) | ( 0.595) | ( 0.312) |
| *Race Difference: P-Value* | | | *0.035* | | | *0.000* |
| $\beta$/Pop | -0.004 | -0.007 | -0.003 | 0.77 | 1.16 | 0.73 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.021* |
| Y-Mean | 255.96 | 143.55 | 67.95 | 61510.7 | 31364.9 | 29596.9 |
| N | 8567 | 8530 | 8521 | 7831 | 7789 | 7809 |
| **(5) Population Group by Year FE** | -0.0562 | -0.0246 | -0.0154 | 6.732 | 1.990 | 4.818 |
| *(First Stage F-Test = 509.97)* | (0.0043) | (0.0028) | (0.0017) | ( 0.914) | ( 0.515) | ( 0.554) |
| *Race Difference: P-Value* | | | *0.005* | | | *0.000* |
| $\beta$/Pop | -0.003 | -0.006 | -0.002 | 0.50 | 0.62 | 0.53 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.606* |
| Y-Mean | 249.05 | 140.36 | 65.55 | 60243.6 | 30896.3 | 28827.3 |
| N | 8553 | 8522 | 8502 | 7804 | 7768 | 7779 |
| **(6) Homicide Group by Year FE** | -0.0532 | -0.0229 | -0.0148 | 6.887 | 2.101 | 4.855 |
| *(First Stage F-Test = 551.91)* | (0.0043) | (0.0028) | (0.0018) | ( 0.931) | ( 0.498) | ( 0.578) |
| *Race Difference: P-Value* | | | *0.015* | | | *0.000* |
| $\beta$/Pop | -0.003 | -0.006 | -0.002 | 0.51 | 0.65 | 0.53 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.478* |
| Y-Mean | 249.05 | 140.36 | 65.55 | 60243.6 | 30896.3 | 28827.3 |
| N | 8553 | 8522 | 8502 | 7804 | 7768 | 7779 |

Table 6: 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 |
|---|---|---|---|---|---|---|
| **(7) Control for Education Spending** | -0.0547 | -0.0250 | -0.0160 | 6.879 | 2.096 | 4.853 |
| *(First Stage F-Test = 488.81)* | (0.0043) | (0.0028) | (0.0018) | ( 0.945) | ( 0.502) | ( 0.585) |
| *Race Difference: P-Value* | | | *0.000* | | | *0.000* |
| $\beta$/Pop | -0.003 | -0.006 | -0.003 | 0.51 | 0.65 | 0.53 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.513* |
| Y-Mean | 249.90 | 140.84 | 65.77 | 60472.4 | 31016.4 | 28935.2 |
| N | 8447 | 8417 | 8396 | 7704 | 7668 | 7679 |
| **(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) |
| *Race Difference: P-Value* | | | *0.000* | | | *0.000* |
| $\beta$/Pop | -0.006 | -0.012 | -0.005 | 0.05 | -0.17 | 0.13 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.000* |
| Y-Mean | 248.48 | 140.05 | 65.39 | 60099.8 | 30823.4 | 28758.3 |
| N | 8602 | 8571 | 8551 | 7848 | 7812 | 7822 |
| **(9) Log Model (Variable+1)** | -0.5349 | -0.7687 | -0.3793 | 0.386 | 0.480 | 0.373 |
| *(First Stage F-Test = 180.94)* | (0.2530) | (0.2896) | (0.2294) | ( 0.209) | ( 0.228) | ( 0.222) |
| *Race Difference: P-Value* | | | *0.292* | | | *0.737* |
| Y-Mean | 4.14 | 3.46 | 2.87 | 9.5 | 8.5 | 8.8 |
| N | 8551 | 8520 | 8500 | 7802 | 7766 | 7777 |
| **(10) Inverse Hyperbolic Sine** | -0.5204 | -0.7649 | -0.3329 | 0.378 | 0.447 | 0.370 |
| *(First Stage F-Test = 180.86)* | (0.2737) | (0.3265) | (0.2422) | ( 0.207) | ( 0.230) | ( 0.222) |
| *Race Difference: P-Value* | | | *0.288* | | | *0.810* |
| Y-Mean | 4.75 | 4.00 | 3.41 | 10.2 | 9.1 | 9.5 |
| N | 8553 | 8522 | 8502 | 7804 | 7768 | 7779 |

Table 6: 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 |
|---|---|---|---|---|---|---|
| **(11) Raw Data** | -0.0543 | -0.0237 | -0.0141 | 8.495 | 3.089 | 5.422 |
| *(First Stage F-Test = 501.92)* | (0.0046) | (0.0028) | (0.0018) | ( 1.104) | ( 0.588) | ( 0.639) |
| *Race Difference: P-Value* | | | *0.004* | | | *0.007* |
| $\beta$/Pop | -0.003 | -0.006 | -0.002 | 0.62 | 0.93 | 0.58 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.065* |
| Y-Mean | 256.60 | 144.47 | 67.25 | 61581.2 | 31682.5 | 29375.7 |
| N | 7789 | 7759 | 7742 | 7139 | 7103 | 7113 |
| **(12) Balanced Panel** | -0.0597 | -0.0266 | -0.0193 | 8.005 | 4.265 | -0.491 |
| *(First Stage F-Test = 1833.49)* | (0.0039) | (0.0025) | (0.0058) | ( 4.763) | ( 2.519) | ( 1.660) |
| *Race Difference: P-Value* | | | *0.370* | | | *0.750* |
| $\beta$/Pop | -0.003 | -0.006 | -0.006 | 0.87 | 2.27 | -0.08 |
| *Race Difference: P-Value* | | | *0.123* | | | *0.828* |
| 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** | -0.0893 | -0.0453 | -0.0409 | 5.296 | 0.876 | 4.645 |
| *(First Stage F-Test = 797.67)* | (0.0044) | (0.0022) | (0.0016) | ( 1.438) | ( 0.769) | ( 0.625) |
| *Race Difference: P-Value* | | | *0.105* | | | *0.000* |
| $\beta$/Pop | -0.005 | -0.011 | -0.007 | 0.42 | 0.30 | 0.55 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.365* |
| Y-Mean | 223.29 | 129.82 | 59.18 | 50034.1 | 24854.8 | 24751.9 |
| N | 6503 | 6474 | 6462 | 5819 | 5811 | 5798 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using an alternative measure of sworn officers from the U.S. Census. 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. Standard errors are clustered at the city-level.

Table 7: 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** | -0.1030 | -0.0503 | -0.0442 | 22.013 | 8.169 | 14.015 |
| *(First Stage F-Test = 15.71)* | (0.0104) | (0.0047) | (0.0010) | ( 5.087) | ( 1.642) | ( 3.473) |
| *Race Difference: P-Value* | | | *0.205* | | | *0.128* |
| $\beta$/Pop | -0.006 | -0.012 | -0.008 | 1.74 | 2.80 | 1.66 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.102* |
| Y-Mean | 223.35 | 130.03 | 59.17 | 49908.0 | 24807.0 | 24674.4 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |
| **(2) Not Weighted by Population** | -0.0941 | -0.0462 | -0.0424 | 21.180 | 8.383 | 12.924 |
| *(First Stage F-Test = 10.14)* | (0.0058) | (0.0031) | (0.0023) | ( 5.535) | ( 2.068) | ( 3.505) |
| *Race Difference: P-Value* | | | *0.322* | | | *0.265* |
| $\beta$/Pop | -0.032 | -0.072 | -0.033 | 7.80 | 14.61 | 6.81 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.054* |
| Y-Mean | 37.48 | 22.55 | 8.87 | 8119.6 | 3644.0 | 4337.2 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |
| **(3) ASG as Endogenous X, UCR as IV** | -0.1097 | -0.0536 | -0.0458 | 20.615 | 7.656 | 13.120 |
| *(First Stage F-Test = 16.58)* | (0.0104) | (0.0047) | (0.0009) | ( 5.148) | ( 1.678) | ( 3.498) |
| *Race Difference: P-Value* | | | *0.101* | | | *0.159* |
| $\beta$/Pop | -0.006 | -0.013 | -0.008 | 1.63 | 2.63 | 1.55 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.129* |
| Y-Mean | 223.22 | 129.78 | 59.16 | 50017.6 | 24846.4 | 24745.9 |
| N | 6509 | 6480 | 6468 | 5824 | 5816 | 5802 |

Table 7: 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 |
|---|---|---|---|---|---|---|
| **(4) Police Employment not Lagged** | -0.1046 | -0.0573 | -0.0517 | 11.656 | 3.934 | 7.448 |
| *(First Stage F-Test = 16.97)* | (0.0044) | (0.0040) | (0.0016) | ( 4.711) | ( 2.354) | ( 2.340) |
| *Race Difference: P-Value* | | | *0.189* | | | *0.290* |
| $\beta$/Pop | -0.006 | -0.014 | -0.009 | 0.91 | 1.33 | 0.87 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.588* |
| Y-Mean | 226.39 | 131.45 | 60.73 | 51136.5 | 25433.3 | 25235.6 |
| N | 6317 | 6290 | 6281 | 5662 | 5657 | 5644 |
| | | | | | | |
| **(5) Population Group by Year FE** | -0.0970 | -0.0467 | -0.0443 | 21.983 | 8.079 | 14.078 |
| *(First Stage F-Test = 15.64)* | (0.0079) | (0.0036) | (0.0010) | ( 5.556) | ( 1.860) | ( 3.722) |
| *Race Difference: P-Value* | | | *0.520* | | | *0.150* |
| $\beta$/Pop | -0.006 | -0.011 | -0.008 | 1.74 | 2.77 | 1.66 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.155* |
| Y-Mean | 223.35 | 130.03 | 59.17 | 49908.0 | 24807.0 | 24674.4 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |
| | | | | | | |
| **(6) Homicide Group by Year FE** | -0.0911 | -0.0436 | -0.0441 | 23.979 | 9.146 | 15.007 |
| *(First Stage F-Test = 17.08)* | (0.0053) | (0.0026) | (0.0012) | ( 5.737) | ( 2.010) | ( 3.746) |
| *Race Difference: P-Value* | | | *0.856* | | | *0.168* |
| $\beta$/Pop | -0.005 | -0.011 | -0.008 | 1.90 | 3.13 | 1.77 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.098* |
| Y-Mean | 223.35 | 130.03 | 59.17 | 49908.0 | 24807.0 | 24674.4 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |

Table 7: 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 |
|---|---|---|---|---|---|---|
| **(7) Control for Education Spending** | -0.0924 | -0.0438 | -0.0450 | 23.975 | 9.141 | 15.009 |
| *(First Stage F-Test = 17.14)* | (0.0052) | (0.0025) | (0.0014) | ( 5.781) | ( 2.029) | ( 3.771) |
| *Race Difference: P-Value* | | | *0.234* | | | *0.128* |
| $\beta$/Pop | -0.005 | -0.011 | -0.008 | 1.89 | 3.11 | 1.77 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.102* |
| Y-Mean | 224.35 | 130.61 | 59.44 | 50157.7 | 24934.7 | 24795.7 |
| N | 6424 | 6396 | 6383 | 5740 | 5732 | 5719 |
| **(8) Excluding Covariates** | -0.0174 | -0.0124 | -0.0088 | 24.688 | 9.913 | 14.830 |
| *(First Stage F-Test = 10.27)* | (0.0066) | (0.0014) | (0.0011) | ( 6.637) | ( 2.186) | ( 4.428) |
| *Race Difference: P-Value* | | | *0.047* | | | *0.319* |
| $\beta$/Pop | -0.001 | -0.003 | -0.001 | 1.96 | 3.39 | 1.75 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.073* |
| Y-Mean | 222.82 | 129.73 | 59.02 | 49779.7 | 24744.6 | 24610.5 |
| N | 6570 | 6541 | 6529 | 5875 | 5867 | 5853 |
| **(9) Log Model (Variable+1)** | 2.9409 | 1.5147 | 4.5989 | 7.920 | 7.893 | 9.297 |
| *(First Stage F-Test = 4.15)* | (2.0710) | (1.9435) | (2.8371) | ( 5.659) | ( 5.658) | ( 6.485) |
| *Race Difference: P-Value* | | | *0.385* | | | *0.974* |
| Y-Mean | 4.09 | 3.44 | 2.75 | 9.4 | 8.4 | 8.7 |
| N | 6528 | 6499 | 6487 | 5837 | 5829 | 5816 |
| **(10) Inverse Hyperbolic Sine** | 3.4858 | 1.8241 | 4.6335 | 7.347 | 7.440 | 8.653 |
| *(First Stage F-Test = 4.89)* | (2.2836) | (2.1561) | (2.9028) | ( 4.895) | ( 4.983) | ( 5.640) |
| *Race Difference: P-Value* | | | *0.476* | | | *0.977* |
| Y-Mean | 4.70 | 3.99 | 3.29 | 10.1 | 9.1 | 9.4 |
| N | 6530 | 6501 | 6489 | 5839 | 5831 | 5818 |

Table 7: 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 |
|---|---|---|---|---|---|---|
| **(11) Raw Data** | -0.0984 | -0.0481 | -0.0421 | 22.020 | 8.172 | 14.018 |
| *(First Stage F-Test = 15.51)* | (0.0100) | (0.0046) | (0.0010) | ( 5.095) | ( 1.645) | ( 3.478) |
| *Race Difference: P-Value* | | | *0.197* | | | *0.129* |
| $\beta$/Pop | -0.006 | -0.012 | -0.007 | 1.74 | 2.80 | 1.66 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.102* |
| Y-Mean | 223.28 | 129.99 | 59.15 | 49891.6 | 24798.6 | 24668.4 |
| N | 6536 | 6507 | 6495 | 5844 | 5836 | 5822 |
| **(12) Balanced Panel** | -0.1049 | -0.0510 | -0.0313 | 47.698 | 9.482 | 38.787 |
| *(First Stage F-Test = 14.59)* | (0.0112) | (0.0049) | (0.0059) | ( 3.226) | ( 1.224) | ( 2.488) |
| *Race Difference: P-Value* | | | *0.000* | | | *0.000* |
| $\beta$/Pop | -0.006 | -0.012 | -0.010 | 5.13 | 5.32 | 5.98 |
| *Race Difference: P-Value* | | | *0.000* | | | *0.102* |
| Y-Mean | 224.69 | 127.22 | 21.36 | 20994.3 | 8693.0 | 11261.9 |
| N | 5321 | 4918 | 4887 | 3755 | 3726 | 3583 |

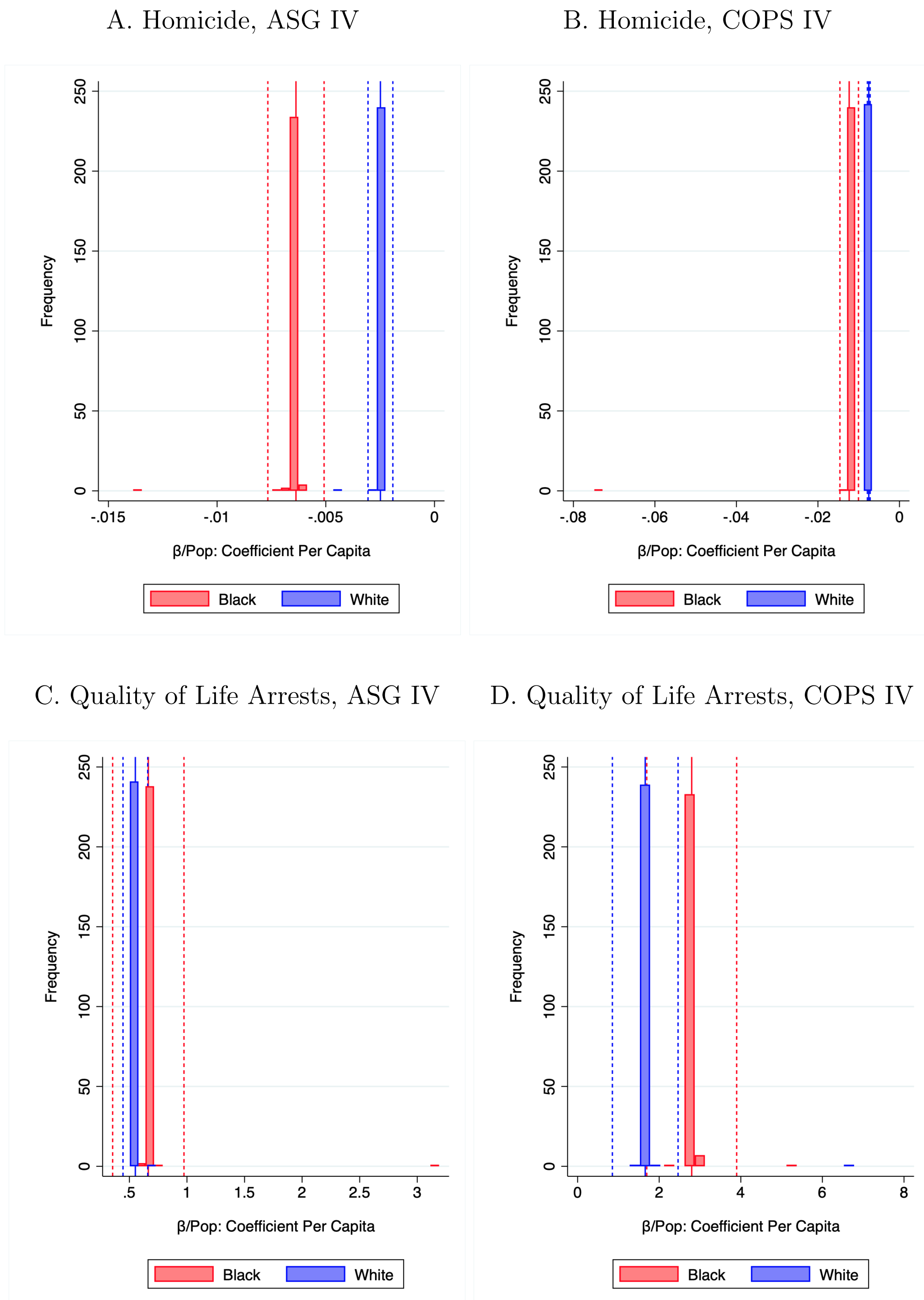
Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using the number of eligible hires awarded through a COPS Hiring grant. Baseline specifications correspond to models in Table 3. 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. Standard errors are clustered at the city-level.

Table 8: 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** | 7.120 | 2.147 | 5.031 | 22.013 | 8.169 | 14.015 |
| | ( 0.876) | ( 0.511) | ( 0.500) | ( 5.087) | ( 1.642) | ( 3.473) |
| *Race Difference: P-Value* | | | *0.000* | | | *0.128* |
| $\beta$/Pop | 0.53 | 0.66 | 0.55 | 1.74 | 2.80 | 1.66 |
| *Race Difference: P-Value* | | | *0.498* | | | *0.102* |
| 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** | 7.116 | 2.144 | 5.029 | 22.008 | 8.165 | 14.013 |
| | ( 0.875) | ( 0.509) | ( 0.500) | ( 5.091) | ( 1.645) | ( 3.474) |
| *Race Difference: P-Value* | | | *0.000* | | | *0.128* |
| $\beta$/Pop | 0.53 | 0.66 | 0.55 | 1.74 | 2.79 | 1.65 |
| *Race Difference: P-Value* | | | *0.500* | | | *0.103* |
| 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 | 0.000000 |
| | (0.000000) | (0.000001) | (0.000000) | (0.000000) | (0.000001) | (0.000000) |
| *Race Difference: P-Value* | | | *0.896* | | | *0.860* |
| Y-Mean | 0.999 | 0.997 | 0.998 | 0.998 | 0.997 | 0.998 |
| N | 7804 | 7768 | 7779 | 5839 | 5831 | 5818 |
| **(4) Report Arrests in All Sub-Categories** | -0.000044 | -0.000050 | -0.000042 | -0.000031 | -0.000011 | -0.000016 |
| | (0.000014) | (0.000012) | (0.000013) | (0.000021) | (0.000019) | (0.000023) |
| *Race Difference: P-Value* | | | *0.649* | | | *0.847* |
| Y-Mean | 0.284 | 0.226 | 0.238 | 0.276 | 0.219 | 0.223 |
| N | 7804 | 7768 | 7779 | 5839 | 5831 | 5818 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant. Baseline specifications correspond to models in Table 2 and Table 3. 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. Standard errors are clustered at the city-level.

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




Note: Figure reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant. Figures present histograms of the per capita effect estimates from 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 in 1980. "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). FBI 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. Standard errors are clustered at the city-level.

Table 9: Results Dis-aggregated by Race Subgroups, ASG Employment IV

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | | Non-Hispanic | | Non-Hispanic | |
| **A. ASG Employment IV** | Black | Black | White | White | Hispanic |
| (1) Homicide Victims | -0.0269 | -0.0258 | -0.0291 | -0.0156 | -0.0152 |
| | (0.0027) | (0.0027) | (0.0017) | (0.0018) | (0.0008) |
| $\beta$/Pop | -0.007 | -0.006 | -0.003 | -0.003 | -0.003 |
| Y-Mean | 142.19 | 140.36 | 100.74 | 65.55 | 37.58 |
| N | 8521 | 8522 | 8512 | 8502 | 8469 |
| (2) Clearance Rates | 0.0008 | 0.0008 | -0.0003 | -0.0005 | 0.0013 |
| | (0.0007) | (0.0007) | (0.0009) | (0.0008) | (0.0024) |
| Y-Mean | 62.61 | 62.56 | 67.74 | 69.47 | 65.73 |
| N | 6070 | 6065 | 7314 | 7045 | 2393 |
| (3) Index Arrests | -0.682 | | -0.451 | | |
| | ( 0.198) | | ( 0.093) | | |
| $\beta$/Pop | -0.21 | | -0.05 | | |
| Y-Mean | 8929.7 | | 7214.3 | | |
| N | 7753 | | 7770 | | |
| (4) Quality of Life Arrests | 2.147 | | 5.031 | | |
| | ( 0.511) | | ( 0.500) | | |
| $\beta$/Pop | 0.66 | | 0.55 | | |
| Y-Mean | 30896.3 | | 28827.3 | | |
| N | 7768 | | 7779 | | |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using an alternative measure of sworn officers from the U.S. Census. 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/Latinx ethnicity, and here the white subgroup includes Hispanic/Latinx individuals. Baseline specifications correspond to models in Table 2. Standard errors are clustered at the city-level.

Table 10: 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 |
|---|---|---|---|---|---|
| (1) Homicide Victims | -0.0507 | -0.0503 | -0.0480 | -0.0442 | -0.0057 |
| | (0.0047) | (0.0047) | (0.0054) | (0.0010) | (0.0045) |
| $\beta$/Pop | -0.012 | -0.012 | -0.004 | -0.007 | -0.001 |
| Y-Mean | 131.69 | 130.03 | 85.70 | 59.17 | 28.71 |
| N | 6501 | 6501 | 6494 | 6489 | 6475 |
| (2) Clearance Rates | 0.0014 | 0.0013 | 0.0006 | 0.0002 | 0.0212 |
| | (0.0012) | (0.0012) | (0.0016) | (0.0019) | (0.0085) |
| Y-Mean | 56.81 | 56.77 | 64.06 | 66.42 | 60.02 |
| N | 4600 | 4598 | 5455 | 5223 | 1734 |
| (3) Index Arrests | -1.137 | | -0.547 | | |
| | ( 0.212) | | ( 0.174) | | |
| $\beta$/Pop | -0.39 | | -0.07 | | |
| Y-Mean | 7007.2 | | 6137.3 | | |
| N | 5808 | | 5811 | | |
| (4) Quality of Life Arrests | 8.169 | | 14.015 | | |
| | ( 1.642) | | ( 3.473) | | |
| $\beta$/Pop | 2.80 | | 1.66 | | |
| Y-Mean | 24807.0 | | 24674.4 | | |
| N | 5831 | | 5818 | | |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using the number of eligible hires awarded through a COPS Hiring grant. 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/Latinx ethnicity, and here the white subgroup includes Hispanic/Latinx individuals. Baseline specifications correspond to models in Table 3. Standard errors are clustered at the city-level.

Table 11: Results for Index Crimes and Arrests by Sub-Type

| **A. ASG Employment IV** | Coeff. | S.E. | $\beta$/Population | S.E. | Mean | N |
|---|---|---|---|---|---|---|
| **Index Crimes** | | | | | | |
| Murder/Manslaughter | -0.068 | ( 0.004) | -0.004 | ( 0.000) | 254.1 | 8558 |
| Rape | -0.052 | ( 0.016) | -0.003 | ( 0.001) | 633.8 | 8561 |
| Robbery | -3.123 | ( 0.153) | -0.187 | ( 0.009) | 10018.6 | 8565 |
| Aggravated Assault | -0.594 | ( 0.094) | -0.036 | ( 0.006) | 9997.1 | 8595 |
| Burglary | -4.501 | ( 0.460) | -0.270 | ( 0.028) | 17299.9 | 8560 |
| Theft | -5.463 | ( 0.594) | -0.327 | ( 0.036) | 45487.9 | 8552 |
| Motor Vehicle Theft | -3.977 | ( 0.366) | -0.259 | ( 0.024) | 14138.6 | 8592 |
| **Index Crime Arrests** | | | | | | |
| Murder/Manslaughter | 0.027 | ( 0.018) | 0.002 | ( 0.001) | 205.3 | 7797 |
| Rape | 0.030 | ( 0.009) | 0.002 | ( 0.001) | 232.4 | 7801 |
| Robbery | -0.605 | ( 0.086) | -0.045 | ( 0.006) | 2639.1 | 7797 |
| Aggravated Assault | -0.055 | ( 0.035) | -0.004 | ( 0.003) | 3528.2 | 7827 |
| Burglary | 0.140 | ( 0.072) | 0.010 | ( 0.005) | 1967.5 | 7792 |
| Theft | 0.032 | ( 0.080) | 0.002 | ( 0.006) | 6293.0 | 7794 |
| Motor Vehicle Theft | -0.546 | ( 0.037) | -0.041 | ( 0.003) | 1478.8 | 7807 |

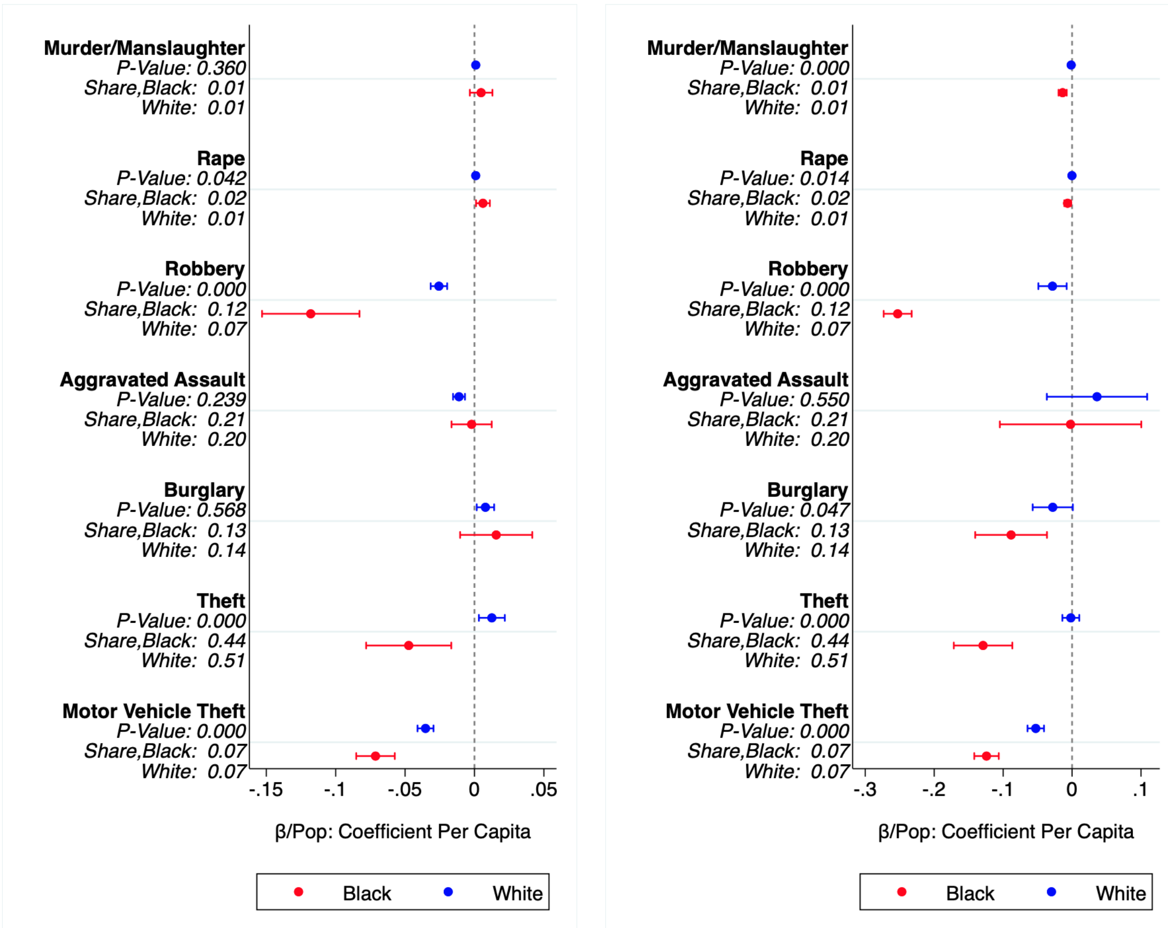
| **B. COPS Eligible Hires IV** | Coeff. | S.E. | $\beta$/Population | S.E. | Mean | N |
|---|---|---|---|---|---|---|
| **Index Crimes** | | | | | | |
| Murder/Manslaughter | -0.106 | ( 0.010) | -0.006 | ( 0.001) | 221.2 | 6546 |
| Rape | -0.094 | ( 0.023) | -0.006 | ( 0.001) | 559.9 | 6554 |
| Robbery | -4.168 | ( 0.347) | -0.244 | ( 0.020) | 8305.6 | 6560 |
| Aggravated Assault | -0.872 | ( 0.265) | -0.051 | ( 0.016) | 9627.5 | 6585 |
| Burglary | -4.914 | ( 0.520) | -0.288 | ( 0.030) | 12899.2 | 6553 |
| Theft | -7.227 | ( 0.689) | -0.423 | ( 0.040) | 40592.1 | 6541 |
| Motor Vehicle Theft | -6.479 | ( 0.609) | -0.424 | ( 0.040) | 11801.9 | 6577 |
| **Index Crime Arrests** | | | | | | |
| Murder/Manslaughter | -0.044 | ( 0.011) | -0.003 | ( 0.001) | 158.1 | 5840 |
| Rape | -0.007 | ( 0.009) | -0.001 | ( 0.001) | 177.3 | 5840 |
| Robbery | -0.994 | ( 0.061) | -0.078 | ( 0.005) | 2140.4 | 5837 |
| Aggravated Assault | 0.499 | ( 0.230) | 0.040 | ( 0.018) | 3308.7 | 5879 |
| Burglary | -0.416 | ( 0.144) | -0.033 | ( 0.011) | 1393.4 | 5826 |
| Theft | -0.013 | ( 0.168) | -0.001 | ( 0.013) | 5023.4 | 5826 |
| Motor Vehicle Theft | -0.622 | ( 0.179) | -0.049 | ( 0.014) | 1146.3 | 5848 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant.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). Standard errors are clustered at the city-level.

Figure A5: Effects of Police Force Size on Index Arrests by Race

A. ASG Employment IV

B. COPS Eligible Hires IV




Note: Figure reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant. Results correspond to per capita estimates. 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 11. 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). "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). FBI 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. Standard errors are clustered at the city-level. "Share, Black" and "Share, White" display the share of that arrest category within all index crime arrests.

Table 12: Results by Quality of Life Arrest Sub-Type

| **A. ASG Employment IV** | Coeff. | S.E. | $\beta$/Population | S.E. | Mean | N |
|---|---|---|---|---|---|---|
| **Quality of Life Arrests** | | | | | | |
| Disorderly Conduct | 1.178 | ( 0.343) | 0.087 | ( 0.025) | 6588.9 | 7788 |
| Suspicious Person | -0.011 | ( 0.015) | -0.001 | ( 0.001) | 28.3 | 7801 |
| Curfew/Loitering | -0.103 | ( 0.103) | -0.008 | ( 0.008) | 1052.5 | 7790 |
| Vandalism | -0.013 | ( 0.029) | -0.001 | ( 0.002) | 1452.9 | 7801 |
| Vagrancy | -0.082 | ( 0.098) | -0.006 | ( 0.007) | 616.0 | 7798 |
| Gambling | 0.334 | ( 0.028) | 0.025 | ( 0.002) | 630.9 | 7791 |
| Drunkenness | 0.184 | ( 0.257) | 0.014 | ( 0.019) | 1869.4 | 7793 |
| Liquor | 8.339 | ( 0.437) | 0.619 | ( 0.032) | 4822.9 | 7790 |
| Drug Possession | 3.829 | ( 0.154) | 0.284 | ( 0.011) | 7294.4 | 7811 |
| Uncategorized Arrests | -6.601 | ( 0.741) | -0.489 | ( 0.055) | 35887.3 | 7818 |

| **B. COPS Eligible Hires IV** | Coeff. | S.E. | $\beta$/Population | S.E. | Mean | N |
|---|---|---|---|---|---|---|
| **Quality of Life Arrests** | | | | | | |
| Disorderly Conduct | 1.182 | ( 0.152) | 0.093 | ( 0.012) | 4390.6 | 5831 |
| Suspicious Person | -0.014 | ( 0.023) | -0.001 | ( 0.002) | 23.8 | 5839 |
| Curfew/Loitering | 1.775 | ( 0.914) | 0.140 | ( 0.072) | 1115.8 | 5843 |
| Vandalism | -0.111 | ( 0.065) | -0.009 | ( 0.005) | 1260.9 | 5840 |
| Vagrancy | -0.285 | ( 0.085) | -0.023 | ( 0.007) | 448.7 | 5843 |
| Gambling | 0.278 | ( 0.016) | 0.022 | ( 0.001) | 455.1 | 5825 |
| Drunkenness | 0.136 | ( 0.245) | 0.011 | ( 0.019) | 1480.2 | 5830 |
| Liquor | 14.243 | ( 0.785) | 1.134 | ( 0.063) | 5231.4 | 5833 |
| Drug Possession | 5.934 | ( 0.853) | 0.470 | ( 0.068) | 7259.1 | 5880 |
| Uncategorized Arrests | -1.052 | ( 2.764) | -0.083 | ( 0.219) | 28131.5 | 5872 |

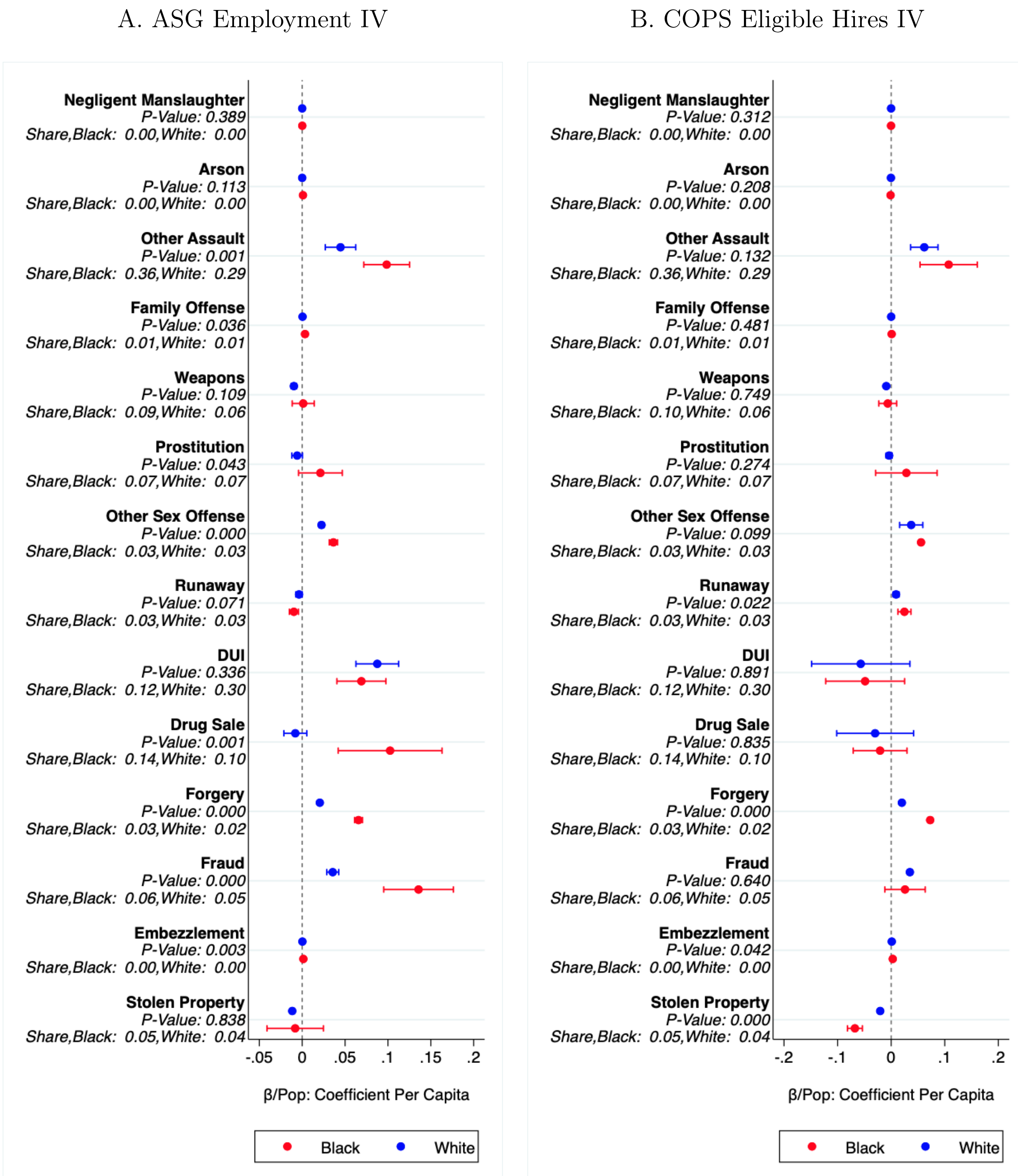
Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant. 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). Standard errors are clustered at the city-level.

Table 13: Results by Non-Index Arrest Sub-Type

| A. ASG Employment IV | Coeff. | S.E. | $\beta$/Population | S.E. | Mean | N |
|---|---|---|---|---|---|---|
| **Non-Index Arrests** | | | | | | |
| Negligent Manslaughter | 0.001 | ( 0.001) | 0.000 | ( 0.000) | 7.3 | 7792 |
| Arson | 0.004 | ( 0.003) | 0.000 | ( 0.000) | 66.7 | 7794 |
| Other Assault | 0.797 | ( 0.114) | 0.059 | ( 0.008) | 4997.6 | 7826 |
| Family Offense | 0.015 | ( 0.006) | 0.001 | ( 0.000) | 102.8 | 7791 |
| Weapons | -0.079 | ( 0.032) | -0.006 | ( 0.002) | 1631.7 | 7805 |
| Prostitution | 0.051 | ( 0.068) | 0.004 | ( 0.005) | 1889.3 | 7792 |
| Other Sex Offense | 0.338 | ( 0.015) | 0.025 | ( 0.001) | 609.1 | 7793 |
| Runaway | -0.065 | ( 0.024) | -0.005 | ( 0.002) | 323.7 | 7799 |
| DUI | 1.083 | ( 0.161) | 0.080 | ( 0.012) | 3091.6 | 7793 |
| Drug Sale | 0.269 | ( 0.154) | 0.020 | ( 0.011) | 4187.5 | 7809 |
| Forgery | 0.434 | ( 0.013) | 0.032 | ( 0.001) | 501.7 | 7795 |
| Fraud | 0.812 | ( 0.101) | 0.060 | ( 0.008) | 2448.9 | 7805 |
| Embezzlement | 0.008 | ( 0.003) | 0.001 | ( 0.000) | 44.2 | 7790 |
| Stolen Property | -0.123 | ( 0.064) | -0.009 | ( 0.005) | 832.9 | 7802 |
| **B. COPS Eligible Hires IV** | Coeff. | S.E. | $\beta$/Population | S.E. | Mean | N |
| **Non-Index Arrests** | | | | | | |
| Negligent Manslaughter | 0.000 | ( 0.000) | 0.000 | ( 0.000) | 6.0 | 5836 |
| Arson | 0.001 | ( 0.002) | 0.000 | ( 0.000) | 49.3 | 5833 |
| Other Assault | 1.079 | ( 0.185) | 0.086 | ( 0.015) | 4902.3 | 5887 |
| Family Offense | 0.000 | ( 0.008) | 0.000 | ( 0.001) | 99.5 | 5854 |
| Weapons | -0.175 | ( 0.038) | -0.014 | ( 0.003) | 1410.2 | 5845 |
| Prostitution | 0.092 | ( 0.052) | 0.007 | ( 0.004) | 1318.8 | 5842 |
| Other Sex Offense | 0.520 | ( 0.090) | 0.041 | ( 0.007) | 559.0 | 5823 |
| Runaway | 0.064 | ( 0.040) | 0.005 | ( 0.003) | 227.9 | 5837 |
| DUI | -0.092 | ( 0.174) | -0.007 | ( 0.014) | 2510.0 | 5826 |
| Drug Sale | 0.178 | ( 0.149) | 0.014 | ( 0.012) | 3988.2 | 5868 |
| Forgery | 0.435 | ( 0.018) | 0.034 | ( 0.001) | 511.4 | 5832 |
| Fraud | -0.117 | ( 0.071) | -0.009 | ( 0.006) | 2298.9 | 5852 |
| Embezzlement | 0.015 | ( 0.003) | 0.001 | ( 0.000) | 40.7 | 5854 |
| Stolen Property | -0.354 | ( 0.040) | -0.028 | ( 0.003) | 614.8 | 5839 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant. 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). Standard errors are clustered at the city-level.

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




Note: Figure reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant. Results correspond to per capita estimates. 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 13. 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). "$\beta$/Pop." divides the coefficient by population (units of 100,000 residents). FBI 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. Standard errors are clustered at the city-level. "Share, Black" and "Share, White" display the share of that arrest category within all non-index arrests.

Table 14: Police Force Size and Officer Deaths and Injuries

| **A. ASG IV** | Coeff. | S.E. | Mean | N |
|---|---|---|---|---|
| Officer Felonious Deaths | 0.0000 | (0.0000) | 0.224 | 8554 |
| Officers Assault Injuries | -0.1365 | (0.0110) | 291.4 | 8563 |

| **B. COPS IV** | Coeff. | S.E. | Mean | N |
|---|---|---|---|---|
| Officer Felonious Deaths | -0.0001 | (0.0001) | 0.158 | 6566 |
| Officers Assault Injuries | -0.2258 | (0.0058) | 203.6 | 6555 |

Note: Table reports estimates from equation (1) in which the once-lagged number of sworn police officers in a city derived from the FBI's Uniform Crime Reports is instrumented using either an alternative measure of sworn officers from the U.S. Census or the number of eligible hires awarded through a COPS Hiring grant. 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. Standard errors are clustered at the city-level.

# 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 Surveys of Governments, Annual Survey of Public Employment and Payroll (ASG Employment, 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.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. 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 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 11, 12, and 13 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.

**Annual Surveys of Governments, Annual Survey of State and Local Government Finances (ASG, 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, 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.