NBER WORKING PAPER SERIES

THE EFFECT OF POLICE ON CRIME:
NEW EVIDENCE FROM U.S. CITIES, 1960-2010

Aaron Chalfin
Justin McCrary

Working Paper 18815
http://www.nber.org/papers/w18815

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2013

For helpful comments and suggestions, we thank Orley Ashenfelter, Emily Bruce, David Card, Raj Chetty, Bob Cooter, John DiNardo, John Eck, Hans Johnson, Louis Kaplow, Mark Kleiman, Tomislav Kovandzic, Prasad Krishnamurthy, Thomas Lemieux, John MacDonald, Je˙ Miron, Denis Nekipelov, Alex Piquero, Jim Powell, Kevin Quinn, Steve Raphael, Jesse Rothstein, Daniel Richman, Seth Sanders, David Sklansky, Kathy Spier, Eric Talley, John Zedlewski, and Frank Zimring, but particularly Aaron Edlin, who discovered a mistake in a preliminary draft, and Emily Owens and Gary Solon, who both read a later draft particularly closely and provided incisive criticisms. We also thank seminar participants from the University of British Columbia, the University of Oregon, the University of California, Berkeley, Harvard University, Brown University, the University of Rochester, the Public Policy Institute of California, the NBER Summer Institute, the University of Texas at Dallas, the University of Cincinnati and the University of South Florida. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w18815.ack

The Effect of Police on Crime: New Evidence from U.S. Cities, 1960-2010
Aaron Chalfin and Justin McCrary
NBER Working Paper No. 18815
February 2013, Revised March 2013
JEL No. H76,J18,K42

ABSTRACT

We argue that the key impediment to accurate measurement of the effect of police on crime is not necessarily simultaneity bias, but bias due to mismeasurement of police. Using a new panel data set on crime in medium to large U.S. cities over 1960- 2010, we obtain measurement error corrected estimates of the police elasticity of the cost-weighted sum of crimes of roughly -0.5. The estimates confirm a controversial finding from the previous literature that police reduce violent crime more so than property crime.

Aaron Chalfin
Goldman School of Public Policy
2607 Hearst Ave
University of California, Berkeley
Berkeley, CA 94720
achalfin@berkeley.edu

Justin McCrary
School of Law
University of California, Berkeley
586 Simon Hall
Berkeley, CA 94720-7200
and NBER
jmccrary@law.berkeley.edu

I. Introduction

One of the most intuitive predictions of deterrence theory is that an increase in a typical offender's chance of being caught decreases crime. This prediction is a core part of Becker's (1968) account of deterrence theory and is also present in historical articulations of deterrence theory, such as Beccaria (1764) and Bentham (1789). The prediction is no less important in more recent treatments, such as the models discussed in Lochner (2004), Burdett, Lagos and Wright (2004), and Lee and McCrary (2009), among others.[1]

On the empirical side, one of the larger literatures in crime focuses on the effect of police on crime, where police are viewed as a primary factor influencing the chance of apprehension facing a potential offender.[2] This literature is ably summarized by Cameron (1988), Nagin (1998), Eck and Maguire (2000), Skogan and Frydl (2004), and Levitt and Miles (2006), all of whom provide extensive references.

Papers in this literature employ a wide variety of econometric approaches. Early empirical papers such as Ehrlich (1972) and Wilson and Boland (1978) focused on the cross-sectional association between police and crime. More recently, concern over the potential endogeneity of policing levels has led to a predominance of papers using panel data techniques such as first-differencing and, more recently, quasi-experimental techniques such as instrumental variables (IV) and differences-in-differences. Prominent panel data papers include Cornwell and Trumbull (1994), Marvell and Moody (1996), Witt, Clarke and Fielding (1999), Fajnzylber, Lederman and Loayza (2002), and Baltagi (2006). Some of the leading examples of quasi-experimental papers are Levitt (1997), Di Tella and Schargrodsky (2004), Klick and Tabarrok (2005), Evans and Owens (2007), and Machin and Marie (2011).

Despite their extraordinary creativity, the quasi-experimental approaches pursued in the literature are typically limited in terms of their inferences by difficulties with precision. For example, a typical finding from this literature is that the police elasticity is larger in magnitude for violent crime than for property crime. This finding is often viewed skeptically however, as there is a common belief that violent crimes such as murder or rape are more apt to be crimes of passion than property crimes such as motor vehicle theft. However, the standard errors on the violent and property crime estimates from the previous literature have been large enough that it is unclear whether the difference in the point estimates is distinguishable from zero. Indeed, for many of the papers in the literature, estimated police elasticities for specific crimes are only

[1] Polinsky and Shavell (2000) provide a review of the theoretical deterrence literature that emerged since Becker (1968), with a particular focus on the normative implications of the theory for the organization of law enforcement strategies.

[2] A related literature considers the efficacy of adoption of "best practices" in policing. Declines in crime have been linked to the adoption of "hot spots" policing (Sherman and Rogan 1995, Sherman and Weisburd 1995, Braga 2001, Braga 2005, Weisburd 2005, Braga and Bond 2008, Berk and MacDonald 2010), "problem-oriented" policing (Braga, Weisburd, Waring, Mazerolle, Spelman and Gajewski 1999, Braga, Kennedy, Waring and Piehl 2001, Weisburd, Telep, Hinckle and Eck 2010) and a variety of similarly proactive approaches. In this paper, we address the effect of additional manpower, under the assumption that police departments operate according to "business-as-usual" practices. As a result, the estimates we report are likely an underestimate with respect to what is possible if additional officers are hired and utilized optimally.

statistically distinct from zero if additional pooling restrictions are imposed (e.g., equal effect sizes for all violent crime categories). Overall, the imprecision of the estimates from the quasi-experimental literature has led to substantial ambiguity regarding the substance of its findings.

Approaches based on natural variation lead to notably more precise estimates than do quasi-experimental approaches, but the former may be apt to bias because of confounding. This suggests there is merit in assessing the extent of confounding. We present evidence that confounding may be less of an issue than previously believed. In particular, using a new panel data set on crime, police, and a host of covariates for 242 large U.S. cities over the period 1960-2010, we demonstrate empirically that, conditional on standard controls, year-over-year changes in police have generally weak associations with the confounders mentioned in the literature, such as demographic factors, the local economy, city budgets, social disorganization, and recent changes in crime. This new dataset covers more cities than have been used and more years than have been used in most (but not all) of the previous literature.

The weakness of the correlations between police and confounders suggests that estimates of the effect of police on crime using natural variation in police may be only slightly biased, despite the *a priori* concerns raised in the literature. A potential problem, however, with using natural variation is that any measurement error in police could lead to bias of a different nature—measurement error bias. The "iron law of econometrics" is that, in a regression, the coefficient on a predicting variable will be too small in magnitude if it is measured with error, with the bias increasing in the amount of measurement error (Hausman 2001). Most natural experiment approaches, such as IV, do not suffer from the same bias (see, for example, Bound, Brown and Mathiowetz (2001)), at least under the hypotheses of the classical measurement error model. Measurement error bias thus has the potential to explain the larger magnitude of the estimates from the quasi-experimental literature, as compared to the traditional literature using natural variation, which has not addressed the issue of measurement errors in police. We show that there is a surprisingly high degree of measurement error in the basic dataset on police used in the U.S. literature, the Uniform Crime Reports (UCR).[3] Estimates from the older panel data literature that failed to account for measurement error bias were likely too small by a factor of 5.

The core of our paper is a series of measurement error corrected estimates of the effect of police on crime using natural variation in year-over-year changes in police at the city level in the U.S. in recent decades. Our estimated police elasticities are substantively large and, taken at face value, suggest that the social value of an additional dollar spent on police in 2010 is approximately $1.60. We introduce a conceptual framework articulating precise conditions under which such a cost-benefit test justifies hiring additional police. The

[3]The degree to which estimates of the total number of police nationally are compromised by measurement errors in the UCR data has been noted by Eck and Maguire (2000). However, they do not discuss the potential for measurement errors at the city level to bias estimates of the police elasticity derived from panel data.

results we introduce along these lines parallel the "sufficient statistic" results discussed in some of the recent public finance literature (e.g., Chetty 2009).

In addition to being significant in substantive terms, our estimated police elasticities are significant in statistical terms. The precision of our estimates allows us to confirm the common and somewhat surprising finding from the previous literature, alluded to above, that police have more of an influence on violent crime than on property crime.[4] However, prior literature has not been able to reject the null hypothesis that the violent crime elasticity is equal to the property crime elasticity, due to imprecise estimates. Our analysis is the first to demonstrate that this apparent finding is unlikely to be due to chance.

Essential to our empirical approach is the existence of two independent measures of police. We combine the standard UCR data on the number of police with data on the number of police from the Annual Survey of Government (ASG). Under the assumptions of the classical measurement error model, described below, IV using one measure as an instrument for the other is a consistent estimator for the results of least squares, were there to be no measurement error. The assumptions of the classical measurement error model are strong, but partially testable. We present the results of a battery of tests of the hypotheses of the classical measurement error model, finding little evidence in our data against them. The tests we utilize would appear to be new to the literature.

Since we focus on natural variation in policing, it is, of course, possible that our estimates are subject to simultaneity bias. It is typical in this literature to difference the data, thus removing between-city variation, and to control for national crime trends using year effects. As the quasi-experimental literature has emphasized, however, this approach may be compromised by confounders associated with growth rates in police and growth rates in crime. A particular concern is that changes in regional macroeconomic conditions, shocks to regional crime markets, or changes in state-level criminal justice policies may act as important confounders, thus biasing the results from standard panel data approaches. The omission of time-varying state-level policy variables is especially concerning as the adoption of a "get tough on crime" attitude among a state's lawmakers might plausibly lead to both increases in police through increased block grants and passage of more punitive state sentencing policies. Such an attitude might be associated with harsher sentencing along both the intensive and extensive margin, changes in a state's capital punishment regime, decreases in the generosity of the state's welfare system or changes in the provision of other public services to low-income individuals.

We seek to address these potential sources of bias with the inclusion of state-by-year effects, an innovation

[4] The cross-crime pattern of the police elasticity estimates could reflect relative deterrence effects, relative incapacitation effects, or non-classical measurement error. The deterrence effect of police is that some crimes will not occur, because a person notes the increase in police presence and thereby is deterred from committing the offense. The incapacitation effect of police is that some crimes will not occur because additional police will result in arrests, pre-trial detention, and jail time for those who offend (McCrary 2009). The non-classical measurement error hypothesis we have in mind is that increases in police might increase reporting of crimes to police. See Levitt (1998) for discussion.

that has not, to date, been utilized in the literature. These state-by-year effects add roughly 1,500 parameters to each set of IV estimates and control for unobserved heterogeneity in city-level crime rates that is constant within the state. Inclusion of these variables increases the R^2 in crime regressions to nearly 60 percent for most crime categories. This is a remarkably high degree of explanatory power for a panel data model specified in growth rates. To the extent that omitted variables bias remains, we note that the previous literature has emphasized that simultaneity bias would lead regression estimates to be positively biased, i.e., to understate the magnitude of the police elasticity of crime (e.g., Nagin 1978, 1998). This reasoning would suggest that our estimates are conservative in magnitude.

The police elasticity of crime is obviously an important component of any public policy discussion regarding the wisdom of changes in police staffing. However, public investments in policing may crowd out private investments in precaution, making a social welfare evaluation of police more involved than it would at first appear. In Section II, we articulate precise conditions under which the police elasticity of crime can be used as a basis for social welfare analysis when private precautions are a first-order consideration. Our framework is related to recent work in public finance emphasizing the central role of policy elasticities in social welfare analysis (e.g., Chetty 2009).

After the social welfare analysis of Section II, Section III shows police hiring is only weakly related to the usual suspected confounders and discusses institutional aspects of police hiring that limit the scope for confounding. This section also provides some comments regarding interpretation. Next, in Section IV, we present direct evidence on the degree of measurement error in survey and administrative data on the number of police. We then outline our econometric methodology in Section V, discuss our primary data in Section VI, and report estimated police elasticities of crime in Section VII. In Section VIII, we compare our results to those from the previous literature. Section IX connects the social welfare analysis of Section II with the empirical findings of Sectio VII; produces a list of the 30 most overpoliced and 30 most underpoliced cities in our sample; and discusses the robustness of our policy conclusions to incapacitation effects of police. Finally, Section X concludes.

II. Conceptual Framework

Our paper provides an empirical examination of the magnitude of the police elasticity of crime. A natural question is whether the elasticity estimates we present are large or small. We now introduce a conceptual framework designed to adress this issue.[5] The framework will provide conditions under which comparing a police elasticity of crime to the ratio of taxes for supporting public policing to the expected cost of crime is a valid basis for welfare analysis (cf., Saez 2001, Chetty 2006, 2009). That is, this section answers the question: Supposing policing passes

[5] Our analysis holds fixed the punishment schedule facing offenders and asks only how to optimally set the probability of apprehension. This can be thought of as a social welfare analysis focused on the choice of policing facing a city having little influence on state sentencing policy.

a cost-benefit test, under what types of conditions is this sufficient to justify hiring additional police officers?

Here is the basic framework we consider. Suppose society consists of n individuals with linear utility over wealth. Each individual i faces a probability of victimization that depends on own precautions, X_i, the precautions of others, and policing, S. The probability of victimization is denoted $\phi_i \equiv \phi_i(X_1, X_2, \ldots, X_n, S)$ and ϕ_i is assumed continuous in all arguments and convex in X_i and in S. To finance policing, each individual pays a lump-sum tax, τ. We assume agents are in a Nash equilibrium, so that the beliefs of any one individual regarding the precautions of others is consistent with the beliefs of the others regarding the precaution of the one. For person i, we take expected utility to be given by

$$U_i \quad = \quad (y_i - k_i)\,\phi_i + y_i(1-\phi_i) = y_i - k_i\phi_i \tag{1}$$

where k_i is the cost of crime, $y_i = A_i - \tau - p_i X_i$ is after-tax wealth net of expenditures on precautions, A_i is initial wealth, and p_i is the price of precaution. We assume any goods that must be purchased in order to obtain precaution are produced under conditions of perfect competition, implying that the only social value of precaution is in lowering crime.[6] Our definition of expected utility can either be thought of as implying that society is comprised exclusively of potential victims or as implying that the social planner refuses to dignify the perpetrator's increased utility, as in Stigler (1970).[7]

Our social planner faces two types of constraints. The *financing* constraint is that total tax receipts for policing, $n\tau$, must equal total expenditures, wS, where w is the cost of hiring an additional officer. The *liberty* constraint is that the social planner is either unwilling or unable to dictate an individual's investments in precaution. To motivate the liberty constraint, note that a person installing a burglar alarm would not be held liable in tort for the burglary of her neighbor, even if it could be shown that the cause of her neighbor's burglary was the installation of the alarm. The liberty constraint is thus one that actual governments respect. To clarify that our social planner calculations are different from an unrestricted social planner's calculations where precautions could conceivably be dictated, we refer to the constrained social planner as the *state*. We define the state's problem as the maximization of average expected utility, $\frac{1}{n}\sum_{i=1}^{n} U_i$, subject to the financing and liberty constraints. This problem can be thought of as (1) delegating to each individual the choice of precaution; and (2) maximizing the average indirect utility function over policing. To solve the state's problem, then, we begin by solving the individual's problem.

Individuals adjust precautions to maximize expected utility. The first order necessary condition for this problem, which is also sufficient under our assumptions, is $p_i = -k_i\phi_{ii}$, where the second subscript indicates a

[6] Precaution may or may not involve a market transaction. For example, it could entail circumnavigating a dangerous neighborhood at the expense of extra travel time, or it could also involve the purchase of a burglar alarm. In these examples, the price of precaution is either the cost of the additional travel time or the market price of the alarm.

[7] See Cameron (1989) for a valuable discussion of these conceptual issues and extensive references to the relevant literature.

partial derivative. We assume that precautions and policing are both protective against crime, or that $\phi_{ii} < 0$ and $\phi_{iS} < 0$. Solving the first order condition for X_i leads to a reaction function, $X_i(X_{-i}, S)$, specifying the privately optimal level of precaution as a function of the precaution of others and policing, where X_{-i} is the vector of precautions for all agents other than i.[8]

Under the assumptions above, each agent has a unique best strategy for any given set of beliefs regarding the actions of other agents, and we obtain a Nash equilibrium in pure strategies (Dasgupta and Maskin 1986, Theorems 1, 2). Figure 1 shows individual reaction functions for the $n = 2$ case under high and low policing.[9] The equilibrium requirement that beliefs be mutually consistent implies a set of restrictions. These restrictions lead to equilibrium demand functions, or the level of precaution demanded by person i as a function of policing, prices, taxes, and assets alone (i.e., not the precautions of others). Write equilibrium demand for precaution as $X_i(S)$. Substituting the equilibrium demand functions into the individual's utility function yields equilibrium maximized expected utility for the individual, or

$$V_i(S) \quad = \quad A_i - \tau - p_i X_i(S) - k_i \phi_i\Big(X_1(S), X_2(S), \ldots, X_n(S), S\Big) \tag{2}$$

The state maximizes the average $V_i(S)$ subject to the financing constraint. Define $\mathcal{V}(S) \equiv \frac{1}{n}\sum_i V_i(S)$ where $\tau = wS/n$. The first order necessary condition, which is also sufficient, is $0 = \mathcal{V}'(S) = \frac{1}{n}\sum_i(-w/n + V_i'(S))$. In this framework, police affect expected utility for individuals through five distinct mechanisms:

1. additional police lower utility by increasing the tax burden $(-w/n < 0)$;
2. additional police increase utility by lowering expenditures on precaution $(-p_i X_i'(S) > 0)$;
3. additional police lower utility by crowding out precaution, thereby increasing the probability of crime indirectly $(-k_i \phi_{ii} X_i'(S) < 0)$;
4. additional police increase utility by reducing the probability of crime directly $(-k_i \phi_{iS} > 0)$; and
5. additional police either lower or increase utility by crowding out precautions by persons $\ell \neq i$, either increasing or decreasing, respectively, the probability of crime externally (the sign of $-k_i \phi_{i\ell} X_\ell'(S)$ is ambiguous because the sign of $\phi_{i\ell}$ is ambiguous)

The first order condition for the state's problem reflects these different mechanisms. Multiplying the first order condition by S/C, where $C = \frac{1}{n}\sum_{i=1}^n k_i \phi_i$ is the *crime index*, or the average expected cost of crime, does not change the sign of the derivative and yields a convenient elasticity representation. We have

$$\mathcal{V}'(S)\frac{S}{C} \quad = \quad -\frac{wS}{nC} - \sum_{i=1}^n \omega_i \rho_i \eta_i - \sum_{i=1}^n \omega_i \varepsilon_{ii} \eta_i - \sum_{i=1}^n \omega_i \varepsilon_{iS} - \sum_{i=1}^n \omega_i \sum_{\ell \neq i} \varepsilon_{i\ell} \eta_\ell \tag{3}$$

where $wS/(nC) = \tau/C$ is the tax burden relative to the expected cost of crime, $\omega_i = k_i\phi_i / \sum_{i=1}^n k_i \phi_i$ is the fraction of the expected cost of crime borne by person i, $\rho_i = p_i X_i(S)/(k_i \phi_i) < 1$ is the ratio of precaution expenses to the expected cost of crime, $\varepsilon_{iS} = \phi_{iS} S/\phi_i < 0$ is the partial elasticity of the probability of crime

[8] We suppress the dependence of the reaction function on prices, taxes, and initial assets to maintain a simple presentation.

[9] The example assumes $-\ln \phi_i(X_1, X_2, S) = \alpha X_i + \beta X_{-i} + \gamma S$, with $\beta < \alpha$, which leads to linear reaction functions $X_i(X_{-i}, S) = (1/\alpha)\left(\ln(\alpha k_i/p_i) - \beta X_{-i} - \gamma S\right)$. This formulation thus echoes the traditional textbook treatment of Cournot duopoly with linear demand (e.g., Tirole 1988, Chapter 5).

for person i with respect to policing, $\varepsilon_{i\ell} = \phi_{i\ell} X_\ell(S)/\phi_i$ is the partial elasticity of the probability of crime for person i with respect to precaution for person ℓ, and $\eta_i = X_i'(S)S/X_i(S)$ is the elasticity of precaution for person i with respect to policing. The five terms in equation (3) correspond to the five different mechanisms described above. Note that if individuals are taking optimal precautions, then the second and third mechanisms exactly offset, i.e., $-\sum_i \omega_i \rho_i \eta_i - \sum_i \omega_i \varepsilon_{ii} \eta_i = 0$, or the envelope theorem.

We now turn to the task of connecting the state's optimality condition to observable quantities, in particular the police elasticity of crime. Estimates of the police elasticity of crime are of two types. The first type is a *total* police elasticity, so called because it reflects both the direct reduction in crime due to increasing police as well as the indirect increase in crime due to reductions to precautions that result from hiring police. The second type is a *partial* police elasticity, so called because it holds precautions fixed and thus reflects only the direct reduction in crime due to increased police. Since our study focuses on changes in crime associated with year-to-year fluctuations in policing, we believe that our study most likely identifies a partial police elasticity, at least if most precautions are fixed investments, such as deadbolts and burglar alarms, or if precautions take the form of habits of potential crime victims that are slow to evolve. Because this is plausible but not demonstrable, however, we provide empirical calibrations both under the assumption that our study identifies the partial elasticity and under the assumption that it identifies the total elasticity.

To make these ideas explicit, note that the total and partial elasticities are given by

$$\widetilde{\theta} = \left(\frac{1}{n} \sum_{i=1}^{n} k_i \left\{ \phi_{ii} X_i'(S) + \phi_{iS} + \sum_{\ell \neq i} \phi_{i\ell} X_\ell'(S) \right\} \right) \frac{S}{C} = \sum_{i=1}^{n} \omega_i \left(\varepsilon_{ii} \eta_i + \varepsilon_{iS} + \sum_{\ell \neq i} \varepsilon_{i\ell} \eta_\ell \right) \quad (4)$$

$$\text{and} \quad \theta = \left(\frac{1}{n} \sum_{i=1}^{n} k_i \phi_{iS} \right) \frac{S}{C} = \sum_{i=1}^{n} \omega_i \varepsilon_{iS} \ , \quad (5)$$

respectively. Next, combining equations (3), (4), and (5), we have

$$\mathcal{V}'(S) \frac{S}{C} = -\frac{wS}{nC} - \sum_{i=1}^{n} \omega_i \rho_i \eta_i - \widetilde{\theta} \equiv -\frac{wS}{nC} - r - \widetilde{\theta} \quad (6)$$

$$= -\frac{wS}{nC} - \sum_{i=1}^{n} \omega_i \sum_{\ell \neq i} \varepsilon_{i\ell} \eta_\ell - \theta \equiv -\frac{wS}{nC} - e - \theta \quad (7)$$

where $r = \sum_{i=1}^{n} \omega_i \rho_i \eta_i$ is the *crowdout effect*, or the weighted average product of the ratio of precaution expenses to the expected cost of crime (ρ_i) and the elasticity of precaution with respect to policing (η_i), and $e = \sum_i \omega_i \sum_{\ell \neq i} \varepsilon_{i\ell} \eta_\ell$ is the *externality effect*, or the weighted average change in the crime index that results from policing crowding out precautions and externally impacting crime (i.e., the fifth mechanism affecting expected utility described above). The weights in the weighted average (ω_i) correspond to the fraction of the total expected cost of crime borne by person i.

The signs of the crowdout and externality effects will be important for some of our reasoning. Consider first the crowdout effect. While we can imagine that a given individual might perversely increase precaution with increased policing,[10] we believe that this is rare. We assume that, at least on average in the population, policing crowds out precautions. Since ρ_i cannot be negative, this means we assume $r \leq 0$.

The sign of the externality effect is somewhat more ambiguous. On the one hand, if forced to guess we would say that most precautions have beggar-they-neighbor effects (i.e., for most i and ℓ, $\varepsilon_{i\ell} \geq 0$), implying a negative overall externality effect, or $e \leq 0$. On the other hand, there are of course precautions that have positive externalities, such as LoJack®. Finally, many precautions have aspects of *both* positive and negative externalities.[11] Consequently, although we have a prior view, we will calibrate our empirical analysis allowing for both positive and negative externality effects.

As noted, equations (6) and (7) are both proportional to the first order condition for the state's problem of maximizing $\mathcal{V}(S)$. Consequently, the state's solution can be recast in terms of the total and partial police elasticities, taxes relative to the expected cost of crime, the externality effect, and the crowdout effect.

Consider first the possibility that our empirical analysis identifies the total elasticity, $\widetilde{\theta}$, i.e., that precautions adjust quickly. Rearranging equation (6) shows that

$$\mathcal{V}'(S) > 0 \quad\iff\quad \widetilde{\theta}\left(1 + r/\widetilde{\theta}\right) < -\frac{wS}{nC} \tag{8}$$

Suppose that increasing police is worthwhile in the provisional cost-benefit sense that

$$|\widetilde{\theta}|/\frac{wS}{nC} \equiv \widetilde{\kappa} \quad > \quad 1 \tag{9}$$

Since r and $\widetilde{\theta}$ share sign, the adjustment term $1 + r/\widetilde{\theta}$ is bigger than one, and if $\widetilde{\kappa} > 1$ then it is conservative to conclude that increasing police is welfare improving. Intuitively, this follows since increasing police under this scenario has two benefits for individuals—reduced crime and reduced expenditures on precaution—and only the first benefit is measured by the police elasticity.

Consider next the possibility that our empirical analysis identifies the partial elasticity, θ, i.e., that precautions are slow to adjust. Rearranging equation (7) shows that

$$\mathcal{V}'(S) > 0 \quad\iff\quad \theta\left(1 + e/\theta\right) < -\frac{wS}{nC} \tag{10}$$

Suppose now that increasing police is worthwhile in the provisional cost-benefit sense that

[10] For example, we can imagine an individual who does not think installing a camera is worth it, because she does not believe there are enough police to follow up on any leads she might give them.

[11] For example, the Club® has a negative externality in that it may displace car theft to another car (Ayres and Levitt 1998). On the other hand, each additional car using the Club® raises search costs for the car thief and provides a marginal disincentive to car theft. As a second example, consider a business installing a security camera. The camera could have a negative externality in displacing a burglary to another business and a positive externality in deterring a sidewalk robbery.

$$|\theta| / \frac{wS}{nC} \equiv \kappa \quad > \quad 1 \tag{11}$$

An analysis like that above shows that if $e \leq 0$, i.e., if precautions have beggar-thy-neighbor effects on average, then it is conservative to conclude that increasing police is welfare improving. This makes sense because under this scenario a typical person's precaution imposes a negative externality on others which government can mitigate through police hiring. Suppose instead that $e > 0$, or that precautions have positive externalities on average. In this scenario, government has an incentive to restrict public policing somewhat, in order to encourage precaution. We assume that externalities play a smaller role than the direct effect of policing, or that $e < |\theta|$.[12] We then have the bounds $0 < 1 + e/\theta < 1$, and the conclusion that

$$\mathcal{V}'(S) > 0 \quad \iff \quad \theta < -\frac{wS}{nC}\frac{1}{1+e/\theta} \quad \iff \quad |\theta| / \frac{wS}{nC} = \kappa > \frac{1}{1 - e/|\theta|} \tag{12}$$

Consequently, the provisional conclusion that increasing police is welfare improving remains correct if

$$\kappa > \frac{1}{1 - e/|\theta|} \quad \iff \quad \frac{e}{|\theta|} < \frac{\kappa - 1}{\kappa} \tag{13}$$

In words, if $\kappa > 1$, hiring police improves welfare as long as the externality effect is not too big relative to the partial elasticity. For example, if $\kappa = 2$, then additional police are socially valuable unless the externality effect is half as large as the partial elasticity, and if $\kappa = 1.5$, then additional police are socially valuable unless the externality effect is one-third as large as the partial elasticity.

This basic framework is readily extended in a variety of directions. One such direction pertains to multiple crime categories, which will be relevant for our empirical calibrations. For multiple crime categories, the crime index continues to be defined as the average expected cost of crime but no longer has the simple definition from above because there is more than one crime category. However, if we redefine the crime index as

$$C \quad = \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{J} k_i^j \phi_i^j \tag{14}$$

we retain the core conclusions of the above analysis with analogous redefinition of terms and greater notational complexity. In connecting our empirical results with this normative framework, we draw on the the literature seeking to estimate the cost of various crimes (e.g., Cohen 2000, Cohen and Piquero 2008). This literature can be understood as seeking to estimate k_i^j for a "typical person". With these estimates, we can take $\widehat{C} = \sum_j k^j N^j \big/ P$ as an approximation to the true crime index, where P is a measure of population, k^j is the cost of crime j, and N^j is the number of such crimes reported to police in a given jurisdiction in a given year, or an approximation for $\sum_i \phi_i^j$. These measurement considerations suggest that in empirical analysis one

[12] Since θ is negative and $wS/(nC)$ is positive, the second inequality in (10) cannot be satisfied if $e > |\theta|$. If $e > |\theta|$ regardless of the level of S, then $\mathcal{V}'(S) > 0$ is never satisfied, and the state is at a corner solution where it is optimal to have no police.

could either use the cost-weighted sum of crimes per capita as a dependent variable, or use the cost-weighted sum of crimes as a dependent variable provided there were population controls included as covariates. We follow the latter approach, as we detail below. Consequently, throughout our analysis, we will consider not just the effect of police on aggregate crime, as is typical of most crime papers, but also the effect of police on the *cost-weighted crime index*, or the weighted sum of crimes, where the weights are an estimate of the cost of the crime. We provide detail on these weights in Section VI, below.

III. Institutional Background and Identification Strategy

As noted above, the primary focus of much of the recent literature on police and crime has been the potential endogeneity of changes in police force strength. These concerns are rooted in the notion that a city ideally intertemporally adjusts its policing levels to smooth the marginal disutility of crime for the median voter, just as a consumer in a lifecycle model ideally intertemporally adjusts purchases to smooth the marginal utility of consumption. Such intertemporal adjustments to police would lead changes in police levels to be endogenous, i.e., to be correlated with unobserved determinants of changes in crime.

Our reading of the economics, political science, and public administration literatures is that the realities of city constraints and politics make intertemporal smoothing difficult, dampening the scope for endogeneity of this type. Cities labor under state- and city-level statutory and constitutional requirements that they balance their budgets annually,[13] they face tax and expenditure limitations,[14] they confront risks associated with hiring police due to legal and contractual obligations which encourages hiring as a means of solving long-term rather than short-term problems,[15] they may be operating under a consent decree or court order regarding racial, ethnic, or sex discrimination which may affect hiring decisions directly or indirectly and may affect retention,[16] they may suffer from inattention regarding staffing or may utilize staffing reductions as bargaining chips (e.g., bailout-seeking),[17] and cities may be hamstrung by unilateral changes to state and federal revenue sharing funds that are

[13]See Cope (1992), Lewis (1994), Rubin (1997), and City of Boston (2007).

[14]See Advisory Commission on Intergovernmental Relations (1977b, 1995), Joyce and Mullins (1991), Poterba and Rueben (1995), Shadbegian (1998, 1999).

[15]Regarding legal obligations, consider two examples: during 1972-1982, the federal government began pressuring departments to hire protected class group members with threat of withholding city and department revenues (Chicago Tribune 1972), and during 1972-1973, Massachusetts municipalities were unsure how to proceed with hiring in light of a constitutional challenge to a state statute allowing departments to favor city residents (Larkin 1973). Regarding contractual obligations, note that union contracts and state and local civil service ordinances may make it difficult to fire a police officer, even one who is substantially underperforming.

[16]For general background, see McCrary (2007).

[17]See, for example, LA Times (1966), Ireton (1976), or Recktenwald (1986a, 1986b). A common pattern is for police departments to have hired a large cohort of officers at some point. For some cities, this was after World War II, for other cities it was the late 1950s, and for other cities it was the 1960s crime wave. Combined with typical pension plans pegged to 20 years or 25 years of service, many departments face retirement waves roughly two decades after a hiring wave, setting the stage for a 20 to 25 year cycle unless the city exercises foresight. For example, in response to the famous Boston Police Strike of 1919, in which nearly three-quarters of the police department went on strike on September 9, then-governor Calvin Coolidge, having assumed control of the department on an emergency basis, refused to allow the strikers to return to work and replaced them all with veterans from World War I (Boston Police Department 1919, Russell 1975). This hiring burst, combined with the State-Boston

difficult to anticipate.[18] In addition, state and local civil service ordinances necessitate a lengthy and transparent hiring process making it difficult to adjust policing levels quickly or in great numbers.[19] Finally, cities may suffer from important principal-agent problems with elected officials having potentially quite different objectives from those of the median voter.[20] In short, if the city is analogous to a lifecycle consumer, it is most akin to one confronting liquidity constraints, limited information, inattention, and perhaps even self-commitment problems.

To amplify these points, consider the case of Chicago over the last five decades. Figure 2 presents an annotated time series of the number of sworn officers in the Chicago Police Department. In 1961, there were just over 10,000 sworn officers in Chicago. Crime and, in particular, the inadequacy of law enforcement was a major theme of the 1964 presidential election (Dodd 1964, Pearson 1964). As riots broke out in many U.S. cities between 1965 and 1968 (National Advisory Commission on Civil Disorders 1968), federal revenue sharing dollars made their way into Chicago budgets and the number of police increased rapidly (Varon 1975). By 1971, the number of sworn officers had risen to just over 13,000. A 1970 suit filed by the Afro-American Patrolmen's League against Chicago alleging *inter alia* discrimination in violation of 42 U.S.C. §1981, the modern legacy of §1 of the Civil Rights Act of 1866, was later joined by the Department of Justice in 1973 after the 1972 amendments to the 1964 Civil Rights Act expanded coverage of the Act to government employers.[21] Eventually, Judge Prentice H. Marshall, a self-described activist judge, reached a now-famous standoff with Mayor Richard Daley (Dardick 2004). Marshall ordered the department to use a quota system for future hiring in order to remedy discrimination in past hiring practices. Daley insisted that under such conditions, he did not intend to hire many officers. After impressive brinkmanship, Daley yielded when it became clear that failing to follow the court order would mean the loss of $100 million dollars in federal funds (Enstad 1976). Thereafter, the city faced a serious budget crisis (O'Shea 1981). The early 1980s saw the initiation of a long-term hiring freeze (Davis 1985), and with attrition the number of sworn officers fell from 12,916 in February 1983 to 11,945 in May 1986.

By summer 1986, the city faced a tidal wave of upcoming retirements. The department had added a large number of officers in the late 1950s, and those officers were nearing retirement. As of early 1987, fully 4,000

Retirement System which provides for a defined benefit pension after 10 years if over 55 and after 20 years if of any age, led to a highly persistent "lumpiness" in the tenure distribution of the department (Boston Police Department 1940, Table VI).

[18] Relevant federal programs over this time period include the Law Enforcement Assistance Administration (1968-1982), the Edward Byrne Memorial State and Local Law Enforcement Assistance programs (1988-2006), the Local Law Enforcement Block Grant program (1996-2006), the Justice Assistance Grant (2006-present), and the Community Oriented Policing Services (1994-present). For background on federal programs, see Varon (1975), Hevesi (2005), Richman (2006), and James (2008). At its peak in the late 1970s, LEAA funding accounted for roughly 5 percent of state and local criminal justice expenditures (Advisory Commission on Intergovernmental Relations 1977a). Background on state programs, which are ubiquitous, is much more scarce, but see Richardson (1980).

[19] See, for example, Greisinger, Slovak and Molkup (1979) and Koper, Maguire and Moore (2001).

[20] This perspective is particularly emphasized in the political science literature; see Banfield and Wilson (1963), Salanick and Pfeffer (1977), Schwochau, Feuille and Delaney (1988), and Clingermayer and Feiock (2001).

[21] For background on this litigation, see McCrary (2007) generally and more specifically *Robinson v. Conlisk*, 385 F. Supp. 529 (N.D. Ill. 1974), *United States v. City of Chicago*, 385 F. Supp. 543 (N.D. Ill. 1974), and *United States v. City of Chicago*, 411 F. Supp. 218 (N.D. Ill. 1976).

officers were eligible for retirement. The city tried to get ahead of the predictable decline in manpower, but it could not hire quickly enough to replace departing officers (Recktenwald 1986a, 1986b). Consequently, the department began shrinking again from 12,809 in April 1987 to 12,055 in November 1989 as the crack epidemic was roughly three years old.[22] The department managed to return to 12,919 sworn officers by January 1992, however, and policing levels were roughly stable until the beginning of the Community Oriented Policing Services (COPS) program. Between COPS funds and improving city revenues from the strong economy, the number of sworn police officers approached 14,000, reaching a peak of 13,927 in December 1996. The numbers were then stable during the crime decline of the 2000s, but in the wake of the 2008 financial crisis, the number of officers declined to 12,244, eroding nearly all the gains in police strength since 1990. Recently released data from the UCR program suggest that the number of sworn officers fell to 12,092 in 2011, and recent city payroll data indicate that the number of officers in October 2012 stood at 11,937.[23]

Overall, the key takeaways from Figure 2 are that: (1) Chicago's police strength has fluctuated a great deal over the past five decades, with swings of 10 percent being rather common, and (2) these fluctuations seem to respond to perceptions of lawlessness, but are also the product of political haggling, budgetary mismanagement, gamesmanship, and a seeming lack of attention on the part of city planners. By our reading, these cycles are not limited to Chicago, but are a pervasive feature of police hiring in cities across the United States (cf., Wilson and Grammich 2009).

Sometimes, these cycles are driven by fiscal crisis and bad luck. For example, in 1981, Boston confronted a sluggish to recessionary economy, Proposition $2\frac{1}{2}$, and a major Massachusetts Supreme Court decision that led to large reductions in Boston's property tax revenue.[24] Massachusetts, like other states, requires municipalities to balance their budgets annually.[25] Forced to balance its budget, the city reduced the police department budget by over 27 percent, eliminated all police capital expenditures, closed many police stations, and reduced the number of sworn officers by 24 percent (Boston Police Department 1982).

Other times, these cycles are driven by mayoral objectives that are unrelated to crime. For example, in the mid 1970s, Mayor Coleman Young sought to aggressively hire officers under an affirmative action plan (Deslippe 2004). The department hired 1,245 officers under the plan in 1977, increasing the size of the police force by some 20 percent, and the next year, a further 227 officers were hired under the plan. After Detroit

[22] Based on our own readings, Chicago newspapers begin mentioning the crack epidemic in 1986, and this is also the date identified more quantitatively by Evans, Garthwaite and Moore (2012).

[23] See `http://www.fbi.gov/about-us/cjis/ucr/` and `http://data.cityofchicago.org`, both accessed on October 24, 2012. The financial crisis led to force reductions in many cities, most famously Camden, which laid off 45 percent of its sworn officers in early 2011 (Katz and Simon 2011).

[24] *Tregor v. Assessors of Boston*, 377 Mass. 602, *cert. denied* 44 U.S. 841 (1979). For background on Proposition $2\frac{1}{2}$, see Massachusetts Department of Revenue (2007).

[25] *General Laws of Massachusetts*, Chapter 59, Section 23. Note that these cuts were partially offset by intervention from state government. See in this regard footnote 27 and Figure 3D, below.

hired those officers, the city confronted a serious budget crisis, forcing the city to lay off 400 and 690 officers in 1979 and 1980, respectively. In 1981 and 1982, the city was able to recall 100 and 171 of the laid off officers, respectively, but a new round of cuts in 1983 undid this effort, and 224 officers were again laid off. In 1984, 135 of those officers were recalled.[26] These sharp changes indicate liquidity problems or perhaps bargaining.

These anecdotal considerations suggest that short-run changes in police are, to a great extent, idiosyncratic. That case is strengthened by establishing that changes in police are only weakly related to changes in observable variables.[27] We now present statistical evidence on the exogeneity of changes in police to several key social, economic and demographic factors, conditional on some basic controls. Each column of Table 1 presents coefficients from 13 separate regressions of the growth rate in the UCR or ASG measure of police on the growth rate in a potential confounder, conditional on the growth rate in city population and either year effects or state-by-year effects, and weighted by 2010 city population.[28] We motivate and describe in greater detail these controls below. For now, it is sufficient to understand that these are the key covariates we will condition on later in the paper, where we model crime growth rates as a function of police growth rates and other covariates. Standard errors, in parentheses, are robust to heteroskedasticity.[29]

The table is divided into three panels, each of which addresses a different class of potential confounders. Panel A explores the relationship between police and the local economy, as measured by personal income, adjusted gross income, wage and salary income, county-level total employment, and the city's municipal expenditures exclusive of police. There are four sets of models, corresponding to the UCR or ASG measure of police and to year effects or state-by-year effects. The estimates in Panel A give little indication that police hiring is strongly related to local economic conditions. While the estimates based on year effects are all positive, they are generally small in magnitude. For example, the largest estimated elasticity is that of police with respect to total county

[26] *NAACP v. Detroit Police Officers Association*, 591 F. Supp. 1194 (1984).

[27] Note that we are *not* arguing that police levels fail to respond to crime in the medium- to long-run. Over a longer time horizon, cities may be able to overcome transaction costs and reoptimize, particularly when confronting severe crises. For example, cities facing a difficult crime problem may be able to obtain "emergency" funding from the state or federal government. Describing the situation in Washington, D.C., around 1994, Harriston and Flaherty (1994) note that the "hiring spree [in police] was a result of congressional alarm over the rising crime rate and the fact that 2,300 officers—about 60 percent of the department—were about to become eligible to retire. Congress voted to withhold the $430 million federal payment to the District for 1989 and again for 1990 until about 1,800 more officers were hired." As another example, in response to the 1980-1981 Boston police staffing crisis, "the Massachusetts Legislature enacted the Tregor Act [in 1982]... [providing] the city of Boston with new revenues... This legislative action terminated all layoffs and greatly diminished the risk that future layoffs might take place." *Boston Firefighters Union Local 718 v. Boston Chapter NAACP, Inc.*, 468 U.S. 1206, 1207 (1984). To the extent that even short-run fluctuations in police are partly responding to crime, it is likely that our estimates understate the effect of police on crime.

[28] As discussed in greater detail below, we include two separate measures of city population growth in these regressions to mitigate measurement error bias associated with errors in measuring city population. The UCR and ASG measures of police are described in Section VI, below. For details on the other variables used in this table, see the Data Appendix. We prefer not to control for these variables directly in our main analyses because they are missing for many years. However, after presenting our main results, we conduct a robustness analysis for the 1970-2002 subsample. During that time period, we can control for most of the potential confounders. These results, given in Table 7, show that our main effects are essentially unaffected by the inclusion of further covariates.

[29] For this and all other tables in the paper, we have additionally computed standard errors that are clustered at the level of the city. These are scarcely different from, and often smaller than, those based on Huber-Eicker-White techniques. The similarity in the standard errors suggests small intra-city residual correlations.

employment for the UCR measure given city population growth and year effects, where the elasticity estimate is 0.10. This would imply that a large 10 percent increase in total employment would result in a 1 percent increase in police. The estimates based on state-by-year effects are slightly smaller in magnitude and of varying signs. In addition to being economically small, these estimates are generally statistically indistinguishable from zero at conventional significance levels, or just barely on the cusp of statistical significance.

In Panel B of Table 1 we relate police hiring to several measures of social disorganization and a measure of the population at risk of arrest. The measures of social disorganization are the fraction of births where the mother is a teenager, for all babies as well as for African American babies, the fraction of births where the baby is low birthweight (less than 2500 grams, or about 5.5 pounds), and a proxy for the 12th grade dropout rate.[30] The first three measures are only available at the county level, while the fourth is measured at the city level. While not directly linked to crime, all four of these variables capture changes in local conditions which may correlate with a need for increased police. Notably, a casual examination of the city-specific time series for each of these variables shows that they are often strongly related to the onset of the "crack epidemic" that swept through U.S. cities during the late 1980s and early 1990s.[31] The results for these four variables are similar to those in Panel A, with little indication of a strong relationship between police and social disorganization, at least conditional on controls. The majority of the elasticities are negative, indicating that more social disorganization is associated with less policing rather than more, and are extremely small in magnitude. The elasticity of largest magnitude, that pertaining to the fraction of births where the mother is a teenager, among black births, for the ASG measure with year effects, is just -0.04.

The last variable in Panel B shows the relationship between police and a proxy for the number of persons at risk for arrest. This variable is constructed using Census data on county population for 16 detailed age-race-gender groups, weighted by the 2009 share of each of these groups among arrestees nationally.[32] For the UCR and ASG series, we estimate elasticities pertaining to the at-risk population of 0.11-0.12 and 0.05-0.10, respectively. As with the other estimates we have seen so far, controlling for year effects or even state-by-year effects matters little. These estimates suggest that a 10 percent increase in the at-risk population leads to a 1 to 1.3 percent increase in police, controlling for the covariates described. We note that a 10 percent increase in the population at risk of arrest is an extreme hypothetical representing 2.6 standard deviations for this variable. Overall, while there is some evidence that police hiring is responsive to demographic changes, we would characterize the relationship as fairly weak.

[30]The dropout rate proxy in year t is one minus the number of 12th graders in the city in year t relative to the number of 11th graders in the city in year $t-1$. See Data Appendix for discussion.

[31]The degree to which the "crack epidemic" represents an exogenous shock to local crime markets has been subject to considerable debate in the literature on the effect of state-level abortion policy on crime.

[32]Arrest data from the FBI were available through 2009 at the time of this writing.

Finally, Panel C presents elasticities of police with respect to three lagged crime aggregates: violent crimes, property crimes and a cost-weighted crime index which weights the prevalence of each crime by an estimate of the social damages associated with that crime (we describe these weights in detail in Section VI, below). These elasticities range from 0.004-0.012 for violent crimes to 0.008-0.019 for property crimes, suggesting that a 10 percent increase in crime would lead to no more than a 0.2 percent increase in police.[33] While the weak association reported here is somewhat at odds with the sense one gets from the existing academic literature, it is consistent with the limited available reports from interviews with police chiefs.[34]

Taken as a whole, the estimates in Table 1 suggest a more limited link between police hiring and potential confounders than the literature has presupposed, at least conditional on our preferred controls. While the results do not indicate a large discrepancy between models controlling for year effects and those controlling for state-by-year effects, in the remainder of the paper, we focus on models for crime that include state-by-year effects. These models are robust to any possible confounder that varies over time at the state level. This includes, for example, state-level policies that may affect crime, such as poor support, education policy, or penal policy. Most papers in the literature focus on models for year-over-year growth rates in crimes at the city level include year effects. Evans and Owens (2007) are unusual in focusing on a more flexible model involving group-specific year effects, where the groups are defined according to population and pre-COPS program crime trends. However, as far as we know, no paper in the literature has used state-by-year effects and thereby completely isolated the effect of police from the effect of state-level policies.

Before closing this section, we would like to address one final issue pertaining to interpretation. After reviewing the literature on police staffing fluctuations, including a non-random sampling of newspaper coverage for specific cities in specific years, we have the impression that policing increases are sometimes associated with the city council or mayor being pleased with the direction of the department. This may mean that the number of police partially proxies for changes in what police are doing, as well. For example, it is possible that increases in police are associated with the hiring of a popular new police chief or with the department being willing to transition to a community policing model. This might mean that our estimates, and those in the previous literature utilizing natural variation, capture an effect of police that is somewhat broader than just the effect of police manpower, *per se*.

[33] We note that a limitation of our analysis of possible confounders is that there are few variables which are collected systematically for a large number of cities for a long period of time. One suspected confounder in particular–calls for service–is sometimes reported in police department annual reports, but is not collected on a systematic basis by any organization. Consequently, we are unable to completely address the issue of possible confounders. On the other hand, calls for service likely does correlate strongly with other measures we do observe, such as crime.

[34] For example, Police Executive Research Forum (2005b) discusses the results of a focus group with four police chiefs (Largo, Scottsdale, Omaha, and Baltimore County) and a deputy chief (Charlotte-Mecklenburg). "Participants pointed out that the crime rate is usually not a major factor in budget success." (p. 18).

IV. Evidence on the Extent of Measurement Error in the Number of Police

A. Direct Evidence

We begin our discussion of the nature and extent of measurement errors in police personnel data using the case of New York City in 2003. The UCR data reports that the New York Police Department employed 28,614 sworn police officers in 2003.[35] Relative to the 37,240 sworn officers employed in 2002 and the 35,513 officers employed in 2004, this is a remarkably low number. If these numbers are to be believed, then the ranks of sworn officers in New York City fell by one-quarter in 2003, only to return to near full strength in 2004.

An alternative interpretion is that the 2003 number is a mistake. Panel A of Figure 3 compares the time series of sworn officers of the New York Police Department based on the UCR reports with that based on administrative data from 1990-2009.[36] These data confirm that the 2003 measure is in error and additionally suggest that the 1999 measure may be in error.[37]

Administrative data on the number of officers are difficult to obtain. More readily available are numbers from departmental annual reports. However, even these are not easy to obtain; annual reports are largely internal municipal documents and historically did not circulate widely. In recent years, many departments have begun a practice of posting annual reports online, but only a few cities post historical annual reports. Moreover, the annual report may or may not report the number of officers employed by the police department.[38]

Nonetheless, we have been able to obtain scattered observations on the number of sworn officers from annual reports for selected years for selected other cities: Los Angeles, Chicago, Boston, and Lincoln, Nebraska. The numbers for Chicago have been further augmented by the strength report data reported in Siskin and Griffin (1997).[39] The time series of sworn officers for these cities is given in Figure 3 in panels B through E. Treating the administrative and annual report data as the true measure, it seems that there is a broad range of fidelity in reporting to the UCR program, with Los Angeles being the most faithful, New York the least, and the others somewhere between those two bookends. While the series are highly correlated in levels or logs, the correlation is notably lower after taking first differences (results unreported). This is important because most of the recent literature analyzes the data in first-differences or with city effects.

Many people are suprised that there are errors in measuring the number of police officers. Errors could

[35] As discussed below, the UCR measurement protocol is a snapshot of the stock of officers as of October 31 of the survey year.

[36] See the Data Appendix for details on these data. Special thanks to Frank Zimring for pointing us towards public domain information on New York police staffing based on his work on the New York City crime drop (Zimring 2011).

[37] We have discussed the 2003 measure of police with other scholars of crime and police, both in economics and in criminology. To date, we have not heard a plausible account for this number, other than that it is a data entry error.

[38] For example, the annual reports for the Boston Police Department are available online beginning in 1885, but the reports stop detailing the number of officers between 1972 and 1981, when the number of officers fell by 40 percent. See `http://www.bpl.org/online/govdocs/bpd_reports.htm`.

[39] See the Data Appendix for details on the annual report and strength report data.

arise due to (1) transitory movements within the year in the number of police, (2) conceptual confusion, or (3) data entry errors. Regarding the first source of error, Figure 4 gives information on transitory movements in police for Chicago for the period 1979-1997. The figure displays the monthly count of the number of sworn officers, with the count for October superimposed as horizontal lines.[40,41] October is chosen because this is the reference month for the UCR data on police used in the literature. The figure demonstrates that there is a great deal of within-year volatility in the number of sworn officers. Overall, the series is characterized by hiring bursts followed by the gradual decline associated with losses due to retention or retirement. Transitory movements in police officers are relevant because surveys typically ask for a point-in-time measure, and the snapshot date differs across surveys. Among those we have been able to examine, internal police department documents use different reporting conventions, typically corresponding to the end of the municipal fiscal year, which varies across municipalities and over time. Perhaps responding to the ambiguities of point-in-time measures, the New York City Police Department uses average daily strength in internal documents.

In addition to transitory movements, there may also be conceptual ambiguity over who counts as a sworn police officer. First, there may be confusion between the number of total employees, which includes civilians, and the number of sworn officers. Second, newly hired sworn officers typically attend Police Academy at reduced pay for roughly 6 months prior to swearing in, and there may be ambiguity regarding whether those students count as sworn officers prior to graduation. Third, due to frictions associated with the hiring process, there is often a discrepancy between the number of officers the department has authority from city government to employ ("authorized strength") and the number of officers currently employed ("deployed strength").[42] For our main sample of cities, we have measures of the number of authorized and deployed sworn officers for selected recent years from the Law Enforcement Management and Administrative Statistics (LEMAS). These data show that the number of deployed sworn officers ranges from 62 to 128 percent of authorized strength.[43]

Finally, the UCR measure of sworn police has errors (e.g., New York in 2003) that are inconsistent with transitory movements within the year in the number of sworn police officers and inconsistent with conceptual

[40]We are not aware of any public-use data sets containing information on within-year fluctuations in police. During the period 1979-1997, a unique non-public dataset on sworn officers in Chicago is available to the authors, however, that allows the construction of monthly counts. These data are discussed in Siskin and Griffin (1997) and were previously used in McCrary (2007). See the Data Appendix for details.

[41]A natural question is whether there is seasonality to police hiring, particularly since summer months are typically high crime months. A regression of log police on an exhaustive set of year and month dummies over the period 1979 to 1997, where monthly data are available, yield an R-square of 0.95. This regression gives little indication of seasonality. While the set of 18 year dummies have an F-statistic of over 193, the set of 11 month dummies have an F-statistic of 1.08, with a p-value of 0.38.

[42]Typical steps include a written examination, a drug test, a background check, an interview, and a series of physical and psychological tests, among others (Police Executive Research Forum 2005a, Wilson and Grammich 2009).

[43]Numbers refer to a pooled analysis of data from 1987, 1990, 1993, 1997, 1999, 2000, and 2003. Population weighted mean and standard deviation are 97 percent and 5 percent, respectively. The LEMAS data also allow us to discount the possibility that there is error due to ambiguities among sworn officers, full-time sworn officers, or full-time-equivalent sworn officers, as only 1 to 2 percent of officers appear to work part-time.

confusion. For such errors, we have no other explanation than typographical or data entry error.[44]

B. Comparison of Two Noisy Measures

Police department internal documents are presumably more accurate than the information police departments report to the UCR program. However, as discussed, these are only available in selected cities and selected years. Trading off accuracy for coverage, we now present a comparison of the UCR series on the number of sworn officers with a series based on the ASG. We use the ASG data to construct an annual series on full-time sworn officers for all the cities in our main analysis sample. We define this sample and give background on the ASG data in Section VI, below.

Figure 5 provides visual evidence of the statistical association between the UCR and ASG series for sworn officers, measured in logs (panel A) and first differences of logs ("growth rates", panel B). In panel A, we observe a nearly perfect linear relationship between the two measures, with the majority of the data points massed around the 45° line. The regression line relating the log UCR measure to the log ASG measure is nearly on top of the 45° line, with a slope of 0.99. Panel B makes it clear that differencing the data substantially reduces the association between the two series; the slope coefficient for the data in growth rates is just 0.22. This much smaller relationship is the expected pattern when the true number of officers changes slowly (**?**, Section 26.2.5).

To appreciate the implications of these findings for quantification of the police elasticity of crime, we turn to a simple statistical model. Suppose the two observed series on police are related to true police as

$$S_i = S_i^* + u_i \tag{15}$$
$$Z_i = S_i^* + v_i \tag{16}$$

and suppose the outcome of interest, Y_i, is related to the true number of police and covariates X_i as

$$Y_i = \theta S_i^* + \gamma' X_i + \epsilon_i \tag{17}$$

Here, S_i is the log UCR measure in a given city and year, Z_i is the log ASG measure, S_i^* is the "true" log police or *signal*, X_i are other covariates measured without error, u_i and v_i are mean zero measurement errors that are mutually uncorrelated and uncorrelated with ϵ_i, S_i^*, and X_i, and ϵ_i is mean zero and uncorrelated with S_i^*, X_i, u_i, and v_i. Equations (15) through (17) and the stochastic restrictions just named constitute what is known as the classical measurement error model (Fuller 1987).

A famous result from the econometric literature on measurement errors (see, for example, Wooldridge

[44]It is worth noting that the crime data are the focus of the UCR system, with notably less attention paid to the police numbers. It is common to see a discussion of UCR crime figures in the local news and for local politicians to be under fire for any spikes in those numbers. However, neither of us have ever seen a local news discussion of the UCR measure of the number of officers. Perhaps because the release cycle used by the FBI for the UCR system involves releasing the numbers for police well after they release the numbers for crimes, reporters seem to ask cities directly for figures on police. This suggests that any lack of care in preparing the UCR police numbers would usually go unnoticed.

(2002, Section 4.4.2)) relates the probability limit of the least squares regression estimate of θ, based on using covariates S_i and X_i, to the scope of measurement errors and the relationship between the signal and the included covariates, under the assumptions of the classical measurement error model:

$$\begin{aligned} \text{plim}_{n\to\infty}\ \widehat{\theta}_{\text{OLS}} &= \frac{\mathbb{C}[\mathcal{M}S_i, \mathcal{M}Y_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{C}[\mathcal{M}S_i^* + \mathcal{M}u_i, \mathcal{M}Y_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{C}[\mathcal{M}S_i^*, \mathcal{M}Y_i]}{\mathbb{V}[\mathcal{M}S_i]} && (18)\\ &= \theta\ \frac{\mathbb{V}[\mathcal{M}S_i^*]}{\mathbb{V}[\mathcal{M}S_i]} = \theta\ \frac{\mathbb{V}[\eta_i]}{\mathbb{V}[\eta_i + \mathcal{M}u_i]} = \theta\ \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_u^2} && (19)\\ &= \theta\ \frac{\sigma_*^2(1-R^2)}{\sigma_*^2(1-R^2) + \sigma_u^2} && (20) \end{aligned}$$

where for a random variable A_i, we define $\mathcal{M}A_i = A_i - \mathbb{V}[X_i]^{-1}\mathbb{C}[X_i, A_i]$, i.e., applying $\mathcal{M}$ purges a random variable of its linear association with X_i, η_i is the associated linear projection residual, σ_η^2 is the variance of η_i, σ_u^2 is the variance of u_i from equation (15), and R^2 is the population R-squared from a regression of the signal S_i^* on the covariates X_i.[45]

This formula stands for three ideas. First, since $\sigma_u^2 > 0$, OLS will be too small in magnitude, or attenuated. Second, while it is a staple of empirical work to see whether a regression estimate is robust to the inclusion of various control variables, equation (20) 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. Third, since the estimates of θ and γ will generally covary, the bias in the estimate of θ will spill over to result in bias in the estimate of γ.

Now return to equation (17) and suppose that X_i is measured without any errors. Under the models in equations (15) and (16) and the associated assumptions on u_i and v_i, it is straightforward to estimate the reliability ratio. The probability limit of the coefficient on Z_i in a regression of S_i on Z_i and X_i is

$$\frac{\mathbb{C}[\mathcal{M}S_i, \mathcal{M}Z_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{C}[\mathcal{M}S_i^* + \mathcal{M}u_i, \mathcal{M}S_i^* + \mathcal{M}v_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{V}[\mathcal{M}S_i^*]}{\mathbb{V}[\mathcal{M}S_i]} \equiv \pi \qquad (21)$$

This implies that the ratio of the least squares estimate of the police elasticity of crime, relative to the estimate of π, is consistent for θ, suggesting a role for IV. This also implies that, in the context of the discussion of Figure 5, a regression of log crime on log police will not be importantly compromised by measurement errors in police, because in logs the reliability ratio is 0.99. However, a regression of growth rates in crime on growth rates in police and other covariates will be compromised, because in growth rates the reliability ratio is 0.22. Indeed, as we show below, once population growth rates and state-year effects are included in the model, the reliability ratio falls to 0.16. Consequently, even setting aside problems with simultaneity bias of the type discussed in the literature, measurement errors in police suggest that least squares estimates of the police

[45]Recall that $R^2 = 1 - \sigma_\eta^2/\sigma_*^2$, so that $\sigma_\eta^2 = \sigma_*^2(1-R^2)$. The formula for $\mathcal{M}$ assumes X_i is a vector with no linear dependencies. More generally, $\mathcal{M}A_i$ is A_i less the linear projection onto the column space of X_i.

elasticity in the literature are too small by a factor of 5 or more.

V. Econometric Approach

The three equation model introduced in Section IV.B leads naturally to a simultaneous equations model. Substituting equation (15) into equation (17) and linearly projecting S_i onto Z_i and X_i yields

$$Y_i = \theta S_i + \gamma' X_i + \varepsilon_i \tag{22}$$
$$S_i = \pi Z_i + \phi' X_i + \nu_i \tag{23}$$

where we now interpret Y_i as the year-over-year change in log crime in a given city and year, S_i as the year-over-year difference in observed log police, and X_i as a vector of control variables such as log revenues per capita, log population, the demographic structure of the population, all measured in first differences, as well as year effects or state-by-year effects. In this model, $\varepsilon_i = \epsilon_i - \theta u_i$, and ν_i is a linear projection error. This is then a standard simultaneous equations model where Z_i is potentially an instrument for S_i.

Estimation of the parameters in equations (22) and (23) proceeds straightforwardly by IV since the model is just-identified, and 2010 city population is used as a weight to obtain a police elasticity estimate representative of the typical person. Sufficient conditions for excluding Z_i from equation (22) are

$$\begin{array}{lr} (A1) & \mathbb{C}[u_i, \epsilon_i] = \mathbb{C}[v_i, \epsilon_i] = 0 \\ (A2) & \mathbb{C}[u_i, (S_i^*, X_i')'] = \mathbb{C}[v_i, (S_i^*, X_i')'] = 0 \\ (A3) & \mathbb{C}[u_i, v_i] = 0 \\ (A4) & \mathbb{C}[\epsilon_i, (S_i^*, X_i')'] = 0 \end{array}$$

where u_i and v_i are the measurement errors from equations (15) and (16) and ϵ_i is the structural error term from equation (17).[46]

Assumptions $(A1)$ through $(A3)$ assert that the measurement error in the UCR and ASG measures of police are not associated with the structural error term in equation (17), and are not associated with the true growth rate in police and the covariate vector X_i, and that the UCR and ASG measurement errors are mutually uncorrelated, respectively. We discuss empirical implications of assumptions $(A1)$ through $(A3)$ below. Assumption $(A4)$ is innocent if we maintain that we would be interested in running a regression of crime growth rates on police growth rates and controls X_i, were police growth rates observed without error. On the other hand, $(A4)$ may reasonably be called into question. In particular, city population growth rates are measured with error. City population growth is a sufficiently important confounder that we feel the (infeasible) regression model implied by equation (17) and assumption $(A4)$ would not be of interest unless X_i

[46] Assumptions $(A1)$ through $(A4)$ together imply that $\mathbb{E}[Z_i \epsilon_i] = \mathbb{E}[Z_i u_i] = 0$, which implies that $\mathbb{E}[Z_i \varepsilon_i] = 0$. Assumptions $(A2)$ and $(A4)$ imply that $\mathbb{E}[X_i \varepsilon_i] = 0$. Of course, $\mathbb{E}[(Z_i, X_i')' \varepsilon_i] = 0$ is one of the two familiar conditions for consistency of IV using Z_i as an excluded instrument and X_i as an included instrument. The other familiar condition, that the excluded instrument predict the endogenous regressor, i.e., that $\pi \neq 0$, is unremarkable in this context.

included it.[47] We discuss the challenges of mismeasurement of city population growth in greater detail below.

Under the classical measurement error, the exact same steps we used to motivate the simultaneous equations model in equations (22) and (23) can be used to motivate a second simultaneous model with the roles of S_i and Z_i reversed and identical parameters in equation (22).[48] We refer to IV models that use the ASG measure of police as an instrument for the UCR measure as *forward* IV estimates and to models that use the UCR measure of police as an instrument for the ASG measure as *reflected*. As noted, both IV estimates are consistent for the police elasticity of crime. This raises the possibility of pooling the estimates to increase efficiency. To do so, we stack the orthogonality conditions for the forward and reflected IV programs into the broader set of moments

$$g_i(\beta) = W_i \begin{pmatrix} Z_i(Y_i - \theta_1 S_i - \gamma_1' X_i) \\ X_i(Y_i - \theta_1 S_i - \gamma_1' X_i) \\ S_i(Y_i - \theta_2 Z_i - \gamma_2' X_i) \\ X_i(Y_i - \theta_2 Z_i - \gamma_2' X_i) \end{pmatrix} \tag{24}$$

where W_i is 2010 city population in levels and all other variables are as defined before, and we estimate the parameters using generalized method of moments (GMM). When the parameters θ_1 and θ_2 and γ_1 and γ_2 are allowed to differ, estimating those same parameters by GMM is equivalent to estimating them separately by IV and correcting the standard errors for the common dependent variable. We can also estimate the system imposing the restriction $\theta_1 = \theta_2 = \theta$.[49] This leads to an implicit averaging of the unrestricted IV estimates and potentially to efficiency gains.[50] An omnibus test of the classical measurement error model is also then available as the standard GMM test of overidentifying restrictions. Since these models are overidentified, there is *a*

[47]In times of population growth, police force size and crime both grow mechanically. For our sample, a population-weighted regression of the growth rate in a typical crime category on the growth rates of population as measured in the UCR and ASG yields a sum of population elasticities of roughly one or even higher. Replacing the dependent variable by the growth rate in police yields a sum of the population elasticities of roughly four-fifths. The resulting positive bias in the estimated police elasticity for specifications that omit population growth is quite large economically.

[48]Some well-known papers utilizing IV strategies to address measurement error have focused on the estimated return to education among samples of twins (see Card (1999) for a review of this literature). The set of econometric issues raised in those papers is slightly different than in our context, simply because twin number is randomly assigned in those studies. In our context, the labels "UCR" and "ASG" carry substantive meaning in a way that the twin labels do not.

[49]A somewhat technical issue arises if we additionally seek to impose the restriction that $\gamma_1 = \gamma_2 = \gamma$: redundancy of moments. When we do not impose any restrictions, we have a just-identified system with $2K$ parameters and $2K$ moments, all of which are linearly independent, where X_i has $K-1$ elements. However, once the restrictions $\theta_1 = \theta_2 = \theta$ and $\gamma_1 = \gamma_2 = \gamma$ are imposed, we have K parameters and only $K+1 < 2K$ linearly independent moments. This suggests two obvious approaches to estimation: (1) impose only the restriction $\theta_1 = \theta_2 = \theta$, in which case there is no moment redundancy; or (2) impose both sets of restrictions and drop $K-1$ moments, in which case GMM will embarassingly differ depending on which set of $K-1$ moments are dropped. An involved solution to the difficulty posed by the second approach is to estimate the models by empirical likelihood (EL; see Imbens 1993, Qin and Lawless 1994, and Imbens 2002 for an introduction and references to the literature), in which case estimates are invariant to the set of moments used to identify the model. EL may also be of interest for the first approach, as the model is (slightly) overidentified. We have used both approaches, using both GMM and EL for the sake of completeness, and there is hardly any difference across the four total possibilities. In our discussion, we focus on the first approach using GMM to maintain a simple presentation and additionally report EL estimates for the sake of completeness. We note that EL computation—a thorny issue—in our application was facilitated greatly by suggestions in Guggenberger and Hahn (2005).

[50]Indeed, a very good approximation to the GMM estimate is the weighted average of the forward and reflected IV estimates, with weights of the inverse squared standard errors. In most software packages, this average will be far easier to compute than GMM. However, the standard errors for GMM are notably larger than the square root of the sum of the weights, so for inference purposes the GMM computation may be necessary.

priori merit in considering empirical likelihood (EL) estimation as well. For overidentified models, EL has been shown to have smaller higher order bias than GMM (Newey and Smith 2004) and to enjoy other advantages as well (see, for example, Imbens, Spady and Johnson 1998). However, in our data, EL estimates and standard errors are nearly identical to two-step GMM estimates, as discussed below, and we focus on GMM.

A challenge we face in implementing the above ideas is that population growth is an important confounder, yet is also likely measured with error. As discussed above, measurement error bias may not have the attenuation bias form if more than one covariate is measured with error. Measurement errors in the population variable in the UCR data are, to the best of our knowledge, not discussed in the literature, but they are likely at least as bad as the measurement errors in police. As with police, any such problems will be particularly serious when the data are measured in growth rates. A potential solution to the measurement problems with city population growth is to again use the UCR measure as an instrument for the ASG measure since both surveys report city population. However, because the two measures are measured very similarly—both are essentially forecasts based on counts from the Census—there are good reasons to believe that the errors in the two measures are not independent of one another. Accordingly, we follow an approach suggested by Lubotsky and Wittenberg (2006) and include both the UCR and the ASG population measures in our main equation of interest. We argue below, based on an empirical comparison to models including data on alternative population controls, that this procedure is sufficient to control for the confounding influence of city population growth.

VI. Data

In this section, we introduce our sample of cities and describe the main sources of information for our data. Our sample of 242 cities is drawn from all cities with more than 50,000 population each year from 1960-2010.[51] In Figure 6 we present a map of the United States highlighting the location of our sampled cities. The shading of states provides information on the number of sampled cities in each state. Our sample contains at least one city in 45 of 51 U.S. states, inclusive of the District of Columbia.[52] In addition, there are 10 states for which our sample contains only a single city. This feature of our data will become relevant in understanding parameter estimates that condition on state-by-year effects.[53]

For each city in our sample, we collect information from public data sources on a variety of different measures. We obtain data on crimes and sworn police officers from the UCR. We collect information on sworn police officers from the ASG and from another survey, described below, that is available for selected years since 1987. These three types of data are the core of our analysis, but we also collect auxiliary data on city

[51] We exclude approximately 30 cities due to extensive missing data and various data quality issues. See Data Appendix for details.
[52] Alaska, Idaho, North Dakota, Vermont and Wyoming are unrepresented in our sample.
[53] States with only a single sampled city are dropped from the analysis when unrestricted state-by-year effects are included.

revenues, police payroll, and police operating budget from the finance files of the ASG; city demographic structure from the Census Bureau; county-level economic data from the Bureau of Economic Analysis; and proxies for social disorganization from the Centers for Disease Control and the National Center for Educational Statistics. Finally, we obtain data on city population from the UCR and ASG which we supplement with data from the National Cancer Institute's Surveillance Epidemiology and End Results (SEER) dataset and some limited information on city births from the National Center for Health Statistics. We now provide more detail regarding each of these data sources. We focus our discussion on our measures of crimes, police, and population, and provide more information regarding our auxiliary data in the Data Appendix.

The UCR crime data we collect are the standard measure used in the empirical literature. These data are collected annually by the FBI. Crime measures represent the total number of offenses known to police to have occurred during the calendar year and are part of the "Return A" collection. The offenses recorded in this system are limited to the so-called index offenses—murder, forcible rape ("rape"), robbery, aggravated assault ("assault"), burglary, larceny exclusive of motor vehicle theft ("larceny"), and motor vehicle theft. Time series for each of the crime rates utilized for each of our cities are shown in Web Appendix Figure 1.

Sworn police are included in both the Law Enforcement Officers Killed or Assaulted (LEOKA) collection and the Police Employees (PE) collection and represent a snapshot as of October 31st of the given year. Because of the late date of the measurement of the number of police, it is typical to measure police in year t using the measure from year $t-1$ (cf., Levitt 1997), and we follow that convention here. Consequently, although we have data on levels from 1960-2010, our regression analyses of growth rates pertain to 1962-2010.

As noted above, we augment data from the UCR with data from the employment files of the ASG. The ASG is an annual survey of municipal 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 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. 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.[54] As with the UCR system, the ASG reports a point-in-time measure of police. For 1960-1995 the reference date is November 1 and for 1997-2010 the reference date is June 30.[55]

The UCR data provide the number of full-time sworn police officers and the total number of police officers

[54]Full-time equivalent employees represent the number of full-time employees who could have been employed if the hours worked by part-time employees had instead been dedicated exclusively to full-time employees. The statistic is calculated by dividing the number of part-time hours by the standard number of full-time hours and then adding this number to the number of full-time employees.

[55]No annual ASG survey was conducted in 1996. We impute data for 1996 using the average of the 1995 and 1997 levels. Other than this one missing year and occasional missing data, information on police is available in both the UCR data and ASG data for each of these cities for the entire study period.

in each year. The ASG data provide the same information beginning in 1977. Prior to 1977, the ASG series reports only the number of full-time equivalent (FTE) police personnel, without differentiating between sworn officers and civilian employees. In order to extend the series, we use the UCR data to generate a city- and year-specific estimate of the proportion of police personnel who are sworn officers. This was accomplished by regressing the proportion of police personnel who are sworn on city and year indicators using the 1960-1977 sample and generating a predicted value for the sworn percentage in each city-year.[56] The ASG FTE numbers before 1977 were then multiplied by the estimated proportion.

For selected analyses we also draw upon a third measure of police. This measure is drawn from two additional sources: the Law Enforcement Management and Administrative Statistics (LEMAS) series and the Census of State and Local Law Enforcement Agencies. These data, which we refer to as the LEMAS series, have been collected at regular intervals from 1987-2008. For additional details, see the Data Appendix.

The measure of city population used in the majority of crime research is from the FBI's Return A file. While this series contains observations for nearly all city-years, it is potentially contaminated by measurement error, particularly in the years immediately prior to each decennial Census. The population entries are contemporaneous; while the FBI could retroactively correct any of the population figures used in the files, it does not. We augment the city population measure from the UCR with the city population measure from the ASG, as noted. As with the UCR, the ASG population measure is noisy and often not smooth across Census year thresholds. Because of the clear errors around Census years, we smooth both series using local linear regression with a bandwidth of 5 years and the triangle kernel (Fan and Gijbels 1996).[57,58,59] Intuitively, this is akin to taking a moving average of the underlying series.

In Section IX, we use data on the cost of police and the cost of crime to derive approximate benefit-cost

[56] Time series plots of the number of full-time sworn officers according to the UCR and ASG measures for each city are provided in Web Appendix Figure 2.

[57] Appendix Figures 1A and 1B provide evidence of the importance of smoothing the raw population measures. These figures present scatterplots of the growth rate in violent and property crimes against the growth rate in the the raw and smoothed population measures from both the UCR and the ASG file. In panel A of Appendix Figure 1A, we see that a 10 percent increase in the population growth rate is associated with a 2.5 percent increase in the number of violent and property crimes. While the crime-population elasticity need not equal 1, this population elasticity is surprisingly small. Panel B plots the crime growth rate against the smoothed UCR population measure. Here, the regression slopes for violent and property crime are 0.94 and 0.84, respectively, neither of which is statistically significantly different from 1. Appendix Figure 1B reports similar results for the ASG population measure. We interpret these findings as evidence that the smoothed population measures more accurately reflect changes in city population.

[58] Below, we test empirically our notion that using both the UCR and the ASG population series adequately controls for population growth using the number of births in a city-year. This can be thought of as a proxy for the size of the population. These data are available from the National Center for Health Statistics (NCHS) for the years 1960-1993 for all 242 cities in our sample and for the years 1960-2003 for 147 of the larger cities. These data are not available electronically, but are available as a series of scanned PDF files at an NCHS website. See `http://www.cdc.gov/nchs/products/vsus.htm`. We had the data on the number of births entered by workers from Amazon's Mechanical Turk service and reviewed them for accuracy. For an introduction to this service, see `http://www.mturk.com/mturk/welcome`. We note that the manner in which we had the data entered and the screening process we undertook together persuade us that there are no data entry errors on the part of the Mechanical Turk workers.

[59] Our population imputations, as well as the raw data underneath them, are shown for each city in the sample in Web Appendix Figures 4A and 4B.

ratios, both nationally and for specific cities. We pause here to describe these data briefly, with further detail provided in the Data Appendix. Data on the cost of police are taken from the ASG Finance and ASG Employment files from 2003-2010 and are used in conjunction with other public data to estimate a fully-loaded cost of hiring an additional police officer. As noted we also make use of the crime index, or the cost-weighted sum of crimes. The correct weight to use for connecting our empirical estimates to the framework of Section II is a measure of the *ex ante* cost of crime—i.e., the dollar amount a potential crime victim would pay to reduce their probability of victimization, relative to the change in the probability. While this is in principle a person-specific concept, we follow the literature in using an estimate for a representative person. Unfortunately, estimates of the *ex ante* cost of crime are not available except for the crime of murder, where we can take advantage of the rich literature on the value of a statistical life (VSL). For other crimes, we use estimates of the *ex post* costs of crime, which are typically derived from civil jury awards. The value of these civil jury awards captures both direct costs to crime victims arising from injuries sustained during the commission of the crime, as well as losses arising from reductions in a victim's quality of life.

We turn now to Table 2, which provides summary statistics for each of our two primary police measures as well as each of the seven index offenses. We additionally report summary statistics for the aggregated crime categories of violent and property crime, which simply add together the relevant corresponding individual crime categories, respectively, and for the cost-weighted crime index.

Descriptive statistics are reported for a sample of 10,589 observations, the universe of data for which measures of crime, police and population are nonmissing. The left-hand panel of Table 2 gives statistics for the levels of crime and police in per capita terms, specifically as a measure of the value per 100,000 population. The right-hand panel gives statistics for log differences of crime and police.

Several features of the data are worth noting. First, a typical city employs approximately 250 police officers per 100,000 population, one officer for every 4 violent crimes, and one officer for every 24 property crimes. There is considerable heterogeneity in this measure over time, with the vast majority of cities hiring additional police personnel over the study period. However, there is even greater heterogeneity across cities, with between city variation accounting for nearly 90 percent of the overall variation in the measure. The pattern is somewhat different for the crime data, with a roughly equal proportion of the variation arising between and within cities.

Second, the vast majority (91 percent) of crimes are property crimes with the most serious crimes (murder and rape) comprising less than 1 percent of all crimes reported to police. It is likewise important to note that each of the crime aggregates is dominated by a particular crime type, with assault comprising nearly half of all violent crimes and larceny comprising nearly sixty percent of all property crimes. This is particularly problematic since these are the two crime categories that are generally believed to be the least comparable across jurisdictions and time periods. Third, and turning to the growth rates, perhaps the most relevant feature of the data is that taking first differences of the series comes close to eliminating time invariant cross-sectional heterogeneity in log crime and log police. For each measure of crime and police, the within standard deviation in growth rates is essentially equal to the overall standard deviation. Moreover, in results not shown, the first difference of a log per capita measure exhibits essentially no cross-sectional heterogeneity.

Because of the prominence of the growth rate in police for our analysis, it is of interest to examine the marginal distribution of the growth rate in police for the UCR data and the ASG data separately. Both series exhibit a mass point at zero. In the UCR data, roughly 3.9 percent of the population-weighted observations have a growth rate of zero. The corresponding figure in the ASG data is 6.1 percent. Figure 7 presents estimates of the conditional density function for the growth rate in police, conditional on not being zero.[60] The figure indicates that the growth rate in police is roughly symmetric with a range of approximately -8 to 12 percent for both series. Compared to the UCR series, the ASG data has a greater prevalence of zero growth rates and a greater prevalence of extreme growth rates. For reference, the figure also shows normal density curves. These are generally close to the local linear density estimates.

Figure 8 highlights long-run trends in crime and police for our sample of 242 cities as well as for all cities in the United States, 1960-2010. The dotted lines in Panels A present the time series for total violent crimes per 100 thousand persons while the solid lines present the time series for cost of violent crimes per person.[61] Panel B presents the same time series evidence for property crimes while Panel C presents the time series for total sworn officers. Focusing on the trends among our sample of cities, we see that regardless of whether crimes are cost-weighted, the series show a remarkable 30 year rise in criminality from 1960 to 1990, followed by an equally remarkable 20 year decline in criminality from 1990 to 2010. These swings are spectacular in magnitude. For violent crime, costs in 2010 dollars per person rose from \$500 in 1960 to \$2,000 in 1990 before falling to less than \$1,000 in 2010. For property crime, costs in 2010 dollars per person rose from less than \$50 in 1960 to nearly \$150 in 1990 before falling to just above \$50 in 2010. Notably, our sample of cities, which covers approximately one third of the U.S. population over the 1960-2010 time period, closely parallels national trends.

Trends in policing in our sample of cities also closely track trends in policing nationally. The 1960s is a decade of strong gains in police strength, from roughly 160 officers per capita to just over 250 officers per capita, with some acceleration evident after the wave of riots in the period 1965-1968, followed by a slower rate of increase during the first half of the 1970s. During the second half of the 1970s, we see an era of retrenchment, perhaps related

[60]The conditional density function estimates are based on local linear density estimation (Fan and Gijbels 1996) and use a binsize of $b = 0.005$, a bandwidth of $h = 0.025$, and the Epanechnikov kernel. See McCrary (2008) for discussion of this density estimation technique and an application.

[61]This is simply the cost-weighted sum of crimes, computed for the subset of violent crimes, relative to the number of persons and is presented in units of dollars per person.

to urban fiscal problems. From 1980 to 2000, sworn police generally increase, with particularly strong increases in the 1990s. Since 2000 the numbers are roughly flat, with the exception of 2003, which is driven entirely by the erroneous estimate provided by the New York City Police Department to the UCR program (cf., Figure 5).

VII. Results

A. Main Results

To estimate the police elasticity of crime correcting for measurement error, we utilize IV estimates where one noisy measure of police is an instrument for another noisy measure. The logical starting point for this analysis is then an examination of the extent to which the UCR and ASG measures of the growth rate in police are related. The first two columns of Table 3 present coefficients and standard errors from models in which the growth rate in the UCR measure is regressed on the growth rate in the ASG measure. These models correspond to what we term our forward IV regressions, in which the UCR measure is the endogenous regressor and the ASG measure is the instrument. The final two columns correspond to what we term our reflected regressions, in which the roles are reversed, with the UCR measure as the endogenous regressor and the ASG measure as the instrument.

Column (1) presents a regression of the growth rate in the UCR measure on the growth rate in the ASG measure, conditional on two measures of the growth rate in the city's population (one from the UCR file and one from the ASG file) as well as a vector of year effects. In the interest of simplicity, we refer to including both population measures as "controlling for population" throughout the paper. In column (2), we condition on state-by-year effects. These capture the effect of any potential covariate that varies over time at the state-level, such as state welfare policy, penal policy, or education policy.[62]

Consistent with the scatterplots presented in Figure 5, the coefficients reported in Table 3 are relatively small in magnitude, indicating that both the UCR measure and the ASG measure contain a great deal of noise once measured in growth rates. Referring for example, to column (1) of Table 3, we observe that, conditional on the growth rate in population, a 10 percent increase in the ASG measure is associated with only a 1.8 percent increase in the UCR measure. Column (2) shows that this result is robust to the inclusion of the full set of state-by-year effects with the coefficient value falling by roughly 10 percent from 0.18 to 0.16.

Turning to columns (3) and (4), which present the results from the reflected first stage regressions, we see that these coefficients are substantially larger in magnitude than the coefficients in columns (1) and (2). These differing magnitudes are expected since the UCR measure of police growth rates exhibits less variance than the ASG measure, and since the first stage coefficient is the covariance between the two measures, relative to

[62] In Table 3, and in subsequent tables, we report Huber-Eicker-White standard errors that are robust to heteroskedasticity. We note that the heteroskedasticity robust standard errors are extremely similar in magnitude to robust standard errors, clustered at the city level. We favor the robust standard errors as they are generally slightly larger in magnitude.

the variance of the predicting variable. As with the forward first stage regressions, results differ only slightly when the state-by-year effects are added.[63]

The F-statistic on the excluded police measure is reported below the coefficient estimates. Since the sample size only affects the scaled distribution of the IV estimator through its impact on the F-statistic, it is often said that the F-statistic is the "effective sample size" of the IV estimator (Rothenberg 1984, Section 6). Since the F-statistics we report are all above 140, standard asymptotic approximations will be highly accurate in the context of our application (cf., Bound, Jaeger and Baker 1995). That is, weak instruments are not a concern in this context.

In Table 4, we present estimates of the police elasticity. The first four columns correspond to least squares models in which we regress the growth rate in crime on the growth rate in police, conditioning on population growth and either year or state-by-year effects. The final four columns of Table 4 correspond to IV regressions that are robust to measurement errors in either of the two police series. Elasticities are estimated for each of the seven index crimes as well as three crime aggregates—violent crimes, property crimes and the cost-weighted crime index.[64]

Turning to column (1) of Table 4, we see that using the UCR measure of police officers, the police elasticity of crime is largest for murder (-0.27), motor vehicle theft (-0.19) and robbery (-0.18). All three elasticities are statistically significant at conventional significance levels. Overall, the elasticity is greater for violent crime (-0.12) than for property crime (-0.07).[65] Reflecting the large weight on murder, the cost-weighted crime elasticity is -0.21 indicating that a ten percent increase in police is associated with a two percent decline in the cost of crime to victims. Referring to column (2), the estimated elasticities are largely similar when the full set of state-by-year effects are included in the model. Here, the elasticities are generally smaller though of

[63]First stage results are extremely similar when we condition additionally on a large number of local-level control variables.

[64]An alternative to using one measure as an instrument for the other is to try to restrict attention to observations that do not contain obvious errors. For example, out of our primary sample of 10,589 observations, roughly 1,000 are either zero (potentially consistent with simply filling out the survey with a copy of the numbers for last year) or are consistent with a growth rate in excess of 20 percent in absolute value (potentially consistent with a gross error such as New York in 2003). This approach is only somewhat successful in our application. For example, the OLS regression of the growth rate of murder on the UCR measure of the growth rate in police is -0.204 in the primary sample and -0.359 in the restricted sample of 9,616 observations where the UCR measure is neither zero nor larger than 0.2 in magnitude. The IV estimate using the ASG as an instrument is -0.889, or more than twice as large as the estimate from the restricted sample (all three estimates control for two measures of population growth rates and state-by-year effects). If we perform the same analysis with the ASG measure as the endogenous regressor, the analogous three estimates are -0.143, -0.171, and -0.572.

[65]In a recent working paper, Solon, Haider and Wooldridge (2012) note that using weighted least squares will not necessarily estimate the average partial effect in the presence of unmodeled heterogeneous effects. They suggest an alternate procedure whereby population is interacted with the main effect of interest. As a robustness check, we re-estimate the population-weighted estimates in Table 4 using this formulation, centering population around the population of the city in which a typical individual lives in our sample, which we write as $\overline{w}$, and including the population weight as an additional regressor. Under a linear approximation to the heterogeneity, i.e., $\theta(W_i) = \theta(\overline{w}) + (W_i - \overline{w})\theta'(\overline{w})$, where the prime indicates differentiation, the coefficient on the growth rate in police represents the average partial effect. The estimates we obtain are similar to those reported in Table 4, but slightly less negative for the forward estimates and somewhat more negative for the reflected estimates. For example, for violent crime, we obtain forward and reflected estimates (standard errors) of -0.123 (0.042) and -0.092 (0.037) for violent crime and -0.049 (0.030) and -0.030 (0.026) for property crime, respectively. The degree of similarity between these results and those in columns (2) and (4) of Table 4 provide little evidence in favor of important unmodeled heterogeneity in our primary models, and since the effects are opposite for forward and reflected models, this does not change our pooled estimates importantly.

the same order of magnitude. Conditioning on the state-by-year effects, the largest elasticities are for murder (-0.20), robbery (-0.20), and motor vehicle theft (-0.13). Elasticities for the aggregates are -0.12 for violent crimes, -0.06 for property crimes, and -0.14 for the cost-weighted crime index.

Columns (3) and (4) report results for models in which the growth rate in crimes is regressed on the growth rate in the ASG measure of police. To our knowledge, this is the first time a city-level panel data regression of crime on the ASG measure of police has been run.[66] While the coefficients in columns (3) and (4) are smaller in magnitude, they are also more precisely estimated with significant coefficients for murder (-0.15), motor vehicle theft (-0.11), and robbery (-0.09). While the violent crime elasticity (-0.05) remains significant, the property crime elasticity (-0.03) is no longer significant. Note that the smaller magnitude of the reduced form coefficient in columns (6)-(10) is expected; returning to equation (20), we recall that the degree of attenuation is greater when the reliability ratio is smaller, and the reliability ratio of the ASG measure is worse than that of the UCR measure. These elasticities are largely similar when the full set of state-by-year effects are included in column (4) with the exception of motor vehicle theft which falls by roughly half.

Taken as a whole, least squares estimates of the elasticity of crime with respect to police point to a persistent but modest relationship between changes in police and criminal activity. Regardless of whether we rely on the UCR or ASG measure, a 10 percent increase in the size of a city's police force (which would correpond to a large and costly change in the policy regime) is predicted to lead to only a 1 percent reduction in the rate of violent and property crimes.

In the final four columns of Table 4 we report IV estimates of each crime elasticity that correct for measurement error. These estimates are typically five times larger in magnitude than those estimated via least squares.[67] Referring to column (5), the largest elasticites are those for murder (-0.80), motor vehicle theft (-0.59), robbery (-0.46) and burglary (-0.22). In addition, we report elasticities for each of the two crime aggregates of -0.29 for violent crimes and -0.15 for property crimes, though the latter is not precisely estimated. The elasticity with respect to cost-weighted crimes is -0.61. The elasticities arising from the reflected IV regressions reported in column (7) exhibit a similar pattern with elasticities for murder, motor vehicle theft and robbery of -0.74, -0.51 and -0.49, respectively. Elasticities for the crime aggregates are -0.32 for violent crimes and -0.20 for property crimes.

Finally, in columns (6) and (8), we present IV results that condition on state-by-year effects. Here we report a violent crime elasticity that is approximately -0.35 and a property crime elasticity that is approximately -0.17. Depending on whether the forward or reflected estimates are used, the cost-weighted crime elasticity is between

[66] Marvell and Moody (1996) use the ASG police measure in regressions of the growth rate in crime on the growth rate in police at the state level.

[67] A familiar result is that the IV estimate can be recovered by dividing the "reduced form" estimate of the police elasticites in Table 4 by the first stage estimate presented in Table 3. In this context, to recover the forward IV coefficients presented in columns (5) and (6) of Table 4, we would divide the reflected least squares coefficients in columns (3) and (4) by the relevant first stage coefficient.

-0.40 and -0.61. With regard to the individual crimes, elasticities are largest for murder (between -0.57 and -0.89), robbery (between -0.52 and -0.57), motor vehicle theft (between -0.30 and -0.37) and burglary (between -0.17 and -0.34). While the coefficient on robbery does not change appreciably when conditioning on state-by-year effects, coefficients on motor vehicle theft are approximately 30 to 50 percent smaller with the inclusion of the unrestricted state-by-year effects as compared to the standard first differencing specification. We interpret this as evidence in favor of the presence of substantial time-varying unobserved heterogeneity at the state-level.

In Table 5, we present GMM and EL estimates of the elasticity of crime with respect to police. These estimates combine the information from the forward and reflected IV estimates presented in Table 4. For each crime type, the table reports an elasticity conditional on population growth and state-by-year effects. As before, robust standard errors are presented in parentheses.

The table shows that two-step GMM is more precise than, but hardly differs from, one-step GMM, and that EL and GMM are nearly indistinguishable. The two-step GMM estimates are -0.67 for murder, -0.56 for robbery, -0.34 for motor vehicle theft and -0.23 for burglary. With regard to the crime aggregates, we report an elasticity of -0.34 for violent crimes, -0.17 for property crimes and -0.47 for the cost-weighted crime index. These estimates represent our best guess regarding the police elasticity and are our preferred estimates.[68]

In the bottom panel of Table 5, we report Hansen's J-test of overidentifying restrictions, which provides a measure of the discrepancy between the two parameter estimates.[69] Under the null hypothesis of classical measurement error, the test statistic has a χ^2 distribution with one degree of freedom, which has a 95 percent critical value of 3.84. Table 5 reveals that we fail to reject the null hypothesis of classical measurement errors in each of ten tests. In fact, the largest of these test statistics is just 1.86. We thus interpret the differences in

[68] An alternative approach is to specify a distributional assumption for S_i^* and the errors in the model. Under normality and mutual independence of S_i^*, u_i, v_i, and ϵ_i, imposing zero means for u_i, v_i, and ϵ_i, but allowing a non-zero mean of μ_* for S_i^*, we obtain a log likelihood function of the form

$$\begin{aligned} L_i(\beta) &= \frac{1}{2}\ln(\omega_Y) + \frac{1}{2}\ln(\omega_S) + \frac{1}{2}\ln(\omega_Z) + \frac{1}{2}\ln(\omega_*) - 2\ln(2\pi) - \frac{1}{2}\ln\left(\theta^2\omega_Y + \omega_S + \omega_Z + \omega_*\right) \\ &- \frac{1}{2}S_i^2\omega_S - \frac{1}{2}Z_i^2\omega_Z - \frac{1}{2}\mu_*^2\omega_* - \frac{1}{2}(Y_i - \gamma' X_i)^2\omega_Y + \frac{1}{2}\frac{\{\theta(Y_i - \gamma' X_i)\omega_Y + S_i\omega_S + Z_i\omega_Z + \mu_*\omega_*\}^2}{\theta^2\omega_Y + \omega_S + \omega_Z + \omega_*} \end{aligned}$$

where $\omega_j = 1/\sigma_j^2$ for $j \in \{Y, S, Z, *\}$, corresponding to ϵ_i, u_i, v_i, and S_i^*, respectively, and where $\beta = (\theta, \gamma, \omega_Y, \omega_S, \omega_Z, \omega_*, \mu_*)$. The normal likelihood approach implicitly forms an estimate of S_i^*, given by a linear combination (call it μ_i) of S_i, Z_i, and $Y_i - \gamma' X_i$, and imposes orthogonality conditions akin to those for a regression of Y_i on μ_i and X_i, but adjusted for the fact that μ_i is a generated regressor. To economize on computing time, we apply the MLE to data de-meaned by state-year, just as with EL. This approach yields point estimates (standard errors) for the 10 crime categories in Table 5 of -0.614 (0.225), -0.233 (0.212), -0.530 (0.111), -0.101 (0.122), -0.207 (0.085), -0.079 (0.064), -0.331 (0.097), -0.327 (0.085), -0.166 (0.059), and -0.433 (0.166). Between the MLE estimates and the EL estimates, we favor the EL estimates because they are consistent under a weaker set of assumptions. Between the EL and GMM estimates, we observe small enough differences that in this application the distinction seems academic.

[69] Here, the test statistic is computed via two-step GMM. The results are nearly identical when the test is computed using an EL approach. Because we are unwilling to assert that the variance matrix of the errors is spherical, the two-step GMM estimator is no longer the efficient estimator in its class, which implies that the test of over-identifying restrictions is not equal to the minimized value of the objective function. However, the proper test statistic can nonetheless be constructed; see Newey (1985) for a discussion and the proper formula for this case.

the IV coefficients reported in columns (6) and (8) of Table 4 as providing little evidence against the classical measurement error hypothesis.

Generally speaking, Hansen's J is an omnibus test. Some insight into what aspects of the classical measurement error model are being tested by Hansen's J can be obtained by examining the null hypothesis in more detail. Abstracting from covariates and using the indirect least squares interpretation of IV, the null hypothesis for Hansen's J in this context is that the two IV estimators share a probability limit, or

$$\frac{\mathbb{C}[Y_i, Z_i]}{\mathbb{C}[S_i, Z_i]} = \frac{\mathbb{C}[Y_i, S_i]}{\mathbb{C}[Z_i, S_i]} \tag{25}$$

Since the denominators for the two ratios in equation (25) are the same, the ratios can only be equal if the numerators are. Hansen's J-test is thus a convenient way to test equality of covariances, which is implied by the classical measurement error model since it implies that both covariances simplify to $\mathbb{C}[Y_i, S_i^*]$.

A second characterization of Hansen's J suggests other testing possibilities as well. Write the null hypothesis for Hansen's J as

$$0 = \frac{\mathbb{C}[Y_i, Z_i]}{\mathbb{C}[S_i, Z_i]} - \frac{\mathbb{C}[Y_i, S_i]}{\mathbb{C}[Z_i, S_i]} = \frac{\mathbb{C}[Y_i, Z_i - S_i]}{\mathbb{C}[S_i, Z_i]} \quad \Longleftrightarrow \quad 0 = \frac{\mathbb{C}[Z_i - S_i, Y_i]}{\mathbb{V}[Y_i]} \tag{26}$$

where the logical equivalence follows since the ratio can only be zero if the numerator is zero. This latter characterization emphasizes that the null hypothesis for Hansen's J-test can also be understood as the requirement that the outcome not predict the difference in measures. This is implied by the classical measurement error model because the difference in measures is supposed to reflect only the difference in measurement errors, and each measurement error is supposed to be uncorrelated with the signal and the controls, which is $(A2)$, and with the structural error term, which is $(A1)$.

This is a helpful characterization because it clarifies what aspects of the classical measurement error model can and cannot be tested using Hansen's J-test. Hansen's J evidently does not test the validity of $(A3)$. This makes sense, because if $(A1)$ and $(A2)$ hold, but $(A3)$ does not, both IV estimators measure the incorrect quantity of $\mathbb{C}[Y_i, S_i^*] / (\mathbb{V}[S_i^*] + \mathbb{C}[u_i, v_i])$.

However, the analysis above suggests a natural method for testing $(A3)$ that takes advantage of the fact that for some years we possess a third measure of police from the LEMAS survey. Specifically, with 3 measures of police, we can see whether the difference between any two measures is predictable using a third measure. Since a third measure, say $\widetilde{Z}_i$, can be written as $\widetilde{Z}_i = S_i^* + \widetilde{v}_i$, where the same properties are asserted to hold for $\widetilde{v}_i$ as for u_i and v_i, $S_i - Z_i$ should be unrelated to $\widetilde{Z}_i$, $S_i - \widetilde{Z}_i$ should be unrelated to Z_i, and $Z_i - \widetilde{Z}_i$ should be unrelated to S_i. This method of testing $(A3)$ is really a joint test of $(A2)$ and $(A3)$, since each measure reflects both the signal and the measurement error. Access to a fourth measure would of course make such an approach even more powerful, but that is infeasible in our application.

Tests along these lines are presented in Table 6. These tests partially take advantage of the fact that, for selected years, we have three measures of police taken from the UCR, ASG, and LEMAS measurement systems, as discussed in Section VI, above.[70] Each column of Table 6 pertains to regressions of the the difference between the growth rates for two police measures. Column (1) pertains to the difference between the growth rate in the UCR and ASG measures for the full 1960-2010 sample. Columns (2)-(4) use only the subsample of years for which the LEMAS measure is available, with column (2) pertaining to the UCR and ASG series, column (3) pertaining to the UCR and LEMAS series, and column (4) pertaining to the LEMAS and ASG series. Each column of Table 6 presents coefficients from a regression of the growth rate in the measurement error gap on three categories of covariates: the growth rate in each of the seven index crimes (Panel A), the growth rate in the remaining police measure (Panel B) and the growth rate in each of our two population measures (Panel C).

Referring to Panel A, using the full sample in column (1), we find little evidence of a relationship between the growth rate in the measurement errors and the growth rate in crime for any of the seven index crimes. Of the seven t-ratios, only one is above 1 in magnitude. Columns (2)-(4) provide twenty-one tests of this hypothesis using only the subsample for which the LEMAS measure was collected. Each of these three columns uses a particular difference in measures as the dependent variable: $S_i - Z_i$, $S_i - \widetilde{Z}_i$, and $Z_i - \widetilde{Z}_i$. None of the 21 t-ratios in these columns in Panel A give evidence against the restrictions of the classical measurement error model. As noted, these t-ratio tests amount to joint tests of Assumptions ($A1$) and ($A2$), because crime growth rates reflect both the structural error ϵ_i and the signal S_i^*.

Panel B of Table 6 presents coefficients and standard errors from a regression of a difference in police measures on the police measure not involved in the difference (e.g., $S_i - Z_i$ being regressed on $\widetilde{Z}_i$). These are tests of Assumptions ($A2$) and ($A3$), because under the classical measurement error model, $S_i - Z_i$ is simply a difference in measurement errors, and the third measure reflects both the signal and a third measurement error. The results in this panel may contain some slight evidence against the classical measurement error model. Specifically, one of the three tests (UCR-LEMAS) rejects at the 1 percent level and this may be consistent with mean-reverting measurement error. On the other hand, the other two tests in Panel B provide little evidence against the classical measurement error model at the 5 percent level. More broadly, the magnitude of the covariance seems to be quite small—a 10 percent increase in the growth rate of a given police measure is associated with only a 0.8 percent change in the measurement error.

In Panel C, we present results from a series of regressions of the growth rate in the measurement errors on the growth rate of each of our two population measures. In all cases, we find little evidence of a systematic

[70]For a more detailed discussion of the LEMAS data, see the Data Appendix.

relationship between measurement errors and population growth rates.

Finally, in the bottom panel of Table 6, we present p-values from a series of F-tests on the joint significance of all of the variables in predicting the growth rate in the measurement errors. For the full sample, we fail to reject (p-value = 0.83) that the measurement errors are unrelated to crime, police, and population. For the LEMAS subsample, we fail to reject the null hypothesis in all three cases (p-values = 0.25, 0.17, and 0.07).

Overall, we interpret the evidence in Table 6 as furnishing little evidence against the assumptions of the classical measurement error model. There are 39 total tests presented in Table 6; only one of these tests rejects at the 5 percent level, and no joint test is significant at the 5 percent level.[71]

However, since these tests are not commonly used in the literature, there is a question regarding how powerful these tests are at detecting violations of the classical measurement error model. To address this point, we conducted a small simulation study pegged to our sample. We generate simulated data $(Y_i, S_i, Z_i, \widetilde{Z}_i)$ as

$$\begin{aligned} Y_i &= \theta S_i^* + \epsilon_i && (27)\\ S_i &= \lambda_1 S_i^* + u_i && (28)\\ Z_i &= \lambda_2 S_i^* + v_i && (29)\\ \widetilde{Z}_i &= \lambda_3 S_i^* + \widetilde{v}_i && (30) \end{aligned}$$

where the vector $(S_i^*, \epsilon_i, u_i, v_i, \widetilde{v}_i)$ is distributed jointly normal with zero mean and standard deviations calibrated to match key features of our data.[72] In the simulations, we allow five parameters of the data generating process (DGP) to vary: ρ_1, λ_1, λ_2, λ_3, and ρ_3, where ρ_1 is the (constant) correlation between u_i and ϵ_i, between v_i and ϵ_i, and between $\widetilde{v}_i$ and ϵ_i, and where ρ_3 is the (constant) correlation between u_i and v_i, between u_i and $\widetilde{v}_i$, and between v_i and $\widetilde{v}_i$. These parameters control the covariances among the elements of the vector $(S_i^*, \epsilon_i, u_i, v_i, \widetilde{v}_i)$.

Note that when $\rho_1 = \rho_3 = 0$ and $\lambda_1 = \lambda_2 = \lambda_3 = 1$, the DGP is consistent with the classical measurement error hypothesis. The parameter ρ_1 indexes the extent to which Assumption $(A1)$ is violated; λ_1, λ_2, and λ_3 index the extent to which Assumption $(A2)$ is violated; and ρ_3 indexes the extent to which Assumption $(A3)$ is violated.[73] We maintain Assumption $(A4)$ throughout. For each of 10,000 simulated data sets, we construct the tests performed in Table 6 and record whether the null hypothesis was rejected.[74] This allows

[71] While these tests are not independent, we note that a plot of the quantiles of the 39 t-ratios in Table 6 against the standard normal quantiles indicates similar distributions. Relatedly, the one-sample Kolmogorov-Smirnov test statistic (versus the standard normal distribution) has a p-value of 0.18.

[72] We set $\sigma_* = 0.044$, $\sigma_\epsilon = 0.260$, $\sigma_u = 0.047$, and $\sigma_v = 0.070$, and $\sigma_{\widetilde{v}} = 0.055$. This roughly matches the root mean squared error from IV models for the cost-weighted sum of crimes corresponding to Table 4, the first stage coefficients in Table 3 that condition on state-by-year effects, and the standard deviations of the various police measures after demeaning by state-year.

[73] Throughout, we maintain zero correlation between $(u_i, v_i, \widetilde{v}_i)$ and S_i^*. The parameters λ_j control the extent to which a composite error such as $u_i + (\lambda_1 - 1)S_i^*$ is correlated with S_i^*, where $S_i \equiv S_i^* + u_i + (\lambda_1 - 1)S_i^*$ and analogously for Z_i and $\widetilde{Z}_i$.

[74] To match our tests from Table 6, tests corresponding to column 1 are based on simulated data sets of size $n = 10,589$ and tests corresponding to columns 2 through 4 are based on $n = 1,752$.

us to examine the power of these tests against specific alternatives.

The results of this analysis are shown in Figure 9, which contains four panels. Each panel shows the impact of a departure from the classical measurement error model on the rejection rate for two tests ("Test A" and "Test B"). Test A is a t-ratio test in a bivariate regression of either $S_i - Z_i$, $S_i - \widetilde{Z}_i$, or $Z_i - \widetilde{Z}_i$ on an outcome Y_i (i.e., a test of the type discussed in Table 6, Panel A), and Test B is a t-ratio test where the covariate is not Y_i but a third measure of police (i.e., a test of the type discussed in Table 6, Panel B). The four panels in Figure 9 vary ρ_1, λ_1 and λ_2, and ρ_3, relative to the baseline of the classical measurement error model. The curves displayed are power curves corresponding to the tests which have power against the alternative being displayed. For reference, each panel also shows the average of the simulated GMM estimates. The true parameter in all scenarios is -0.5.

The figure shows that these tests have generally good power. For example, turning to Panel A, if the correlation between a measurement error and the structural error is 0.05, the rejection probability for Test A is roughly 30 percent. This is important, because even a small degree of correlation between a measurement error and the structural error leads to bias. The power of Tests A is very good for column 1, where we have our full sample size, but it is notably lower for columns 2 through 4. Our sense is that the measurement errors are unlikely to be correlated with the structural error, because we did not observe any rejections in any of the 28 tests in Panel A of Table 6, even those in column 1 where this test has quite good power.

Turning to the results in Panels B and C, we see that mean-reverting measurement error is quite likely to be detected as λ_1 or λ_2 depart from 1. Importantly, both Test A and Test B may detect mean-reverting measurement error. The curve labeled "A-any" is the power of a test which rejects at the 5 percent level if and only if one or more of the four Tests A reject at the $0.05/4 = 0.0125$ level. For a single crime outcome, this test has power approaching 20 percent for λ_1 or λ_2 equal to 0.7. We suspect that mean-reverting measurement errors in our data would thus be detected more decisively, either by rejections in columns 2, 3, and 4 of Table 6, or by at least threshold rejections for one or more crime categories.

Finally, in Panel D, we examine the power of Test B against alternatives rooted in correlated measurement errors. It is conceivable that the same core (mismeasured) information informs both the ASG and UCR measures of police. We suspect this happens rarely, as the UCR forms are filled out by employees of the police department and signed by the police chief, whereas the ASG forms are filled out by the mayor's office or city manager's office. However, it is of course true that the mayor could contact the police department for the information, in which case any measurement errors would be positively correlated. Nonetheless, Test B has power to detect correlated measurement errors. We note that to the extent the measurement errors in police are positively associated, we would understate the true effect of police on crime (cf., the expectation of the GMM estimates presented in Panel D).[75]

[75] Of course, as with any specification test, there will be a lack of power in specific directions. We have examined the power to

B. Robustness

Before turning to a discussion of the results presented above, we consider several robustness checks. The estimates in Tables 4 and 5 assume the exogeneity of police conditional on population growth and state-by-year effects. While state-by-year effects soak up important time-varying state-level variation, results will nevertheless be inconsistent if there are time-varying covariates measured at the city-level which are correlated with both growth in police and crime. In Table 1, above, we presented evidence that the growth rate in police is correlated with the growth rate in a number of city- and county-level covariates to only a limited degree. In Table 7, we explore the extent to which elasticities reported in Table 4 are robust to the inclusion of city-level covariates directly. The cost of this more direct analysis is that we are required for data availability reasons to restrict attention to the 1970-2002 subsample. The first six columns refer to estimates using the forward models while the last six refer to estimates from the reflected models.

We begin in column (1) by replicating the coefficients presented in column (6) of Table 4 for the 1970-2002 subsample of our data. These estimates condition on population growth and state-by-year effects. For the 1970-2002 subsample, the violent crime elasticity is -0.29 and the property crime elasticity is -0.26. The largest elasticites are for murder, robbery, and burglary (-1.1, -0.55, and -0.41, respectively). The elasticity for the cost-weighted crime index is -0.79. In column (2) we add a series of economic covariates that capture the growth rate in personal income and total employment as well as revenue and employment in four leading industrial sector (construction, manufacturing, wholesale trade and retail trade). We also include a variable that captures each city's public expenditures exclusive of police to capture the impact of all other municipal spending. In column (3), we include the lags of each of these variables to capture a potentially lagged response of crime to local macroeconomic conditions. In column (4), we capture changes in a city's demographic composition by adding control variables for the population share of sixteen age-gender-race groups within each city. In order to control flexibly for the effect of changes in a city's composition, in column (5) we add polynomials (to the second degree) and interactions for each of the demographic subgroups. Finally, in column (6) we add city-specific linear time trends that would capture long-standing crime trends that are independent of growth in police.

Referring to the forward models, it is apparent that the estimated elasticites change very little with the

detect local departures from the classical measurement error model, but one could instead imagine joint departures, and our tests will have little power against some of these joint alternatives. For example, if $(A2)$ is violated, but the λ_j parameters differ from 1 by the exact same amount, then the rejection rate for both Test A and Test B will be 5 percent. Similarly, if $(A1)$ is violated, but the measurement errors have the exact same covariance with the structural error, then the rejection rate for Test A will be 5 percent regardless of how large is the covariance with the structural error. This underscores, in our minds, the importance of validation studies based on administrative, rather than survey, data. We note that progress in labor supply and in the return to education, for example, occurred after several decades of hard work spent documenting fundamental properties of the measurement errors in the standard data sources on hours, earnings, and education. For a review of some of this literature, see Bound et al. (2001). Much more research along these lines is needed to obtain a clear picture of the proper inferences to be drawn from the crime literature.

inclusion of the controls. Referring, for example, to the cost-weighted crime index, the estimated elasticity moves from -0.79 when conditioning only on population and state-by-year effects to 0.76 when economic covariates are included. Conditioning also on the lags of the economic covariates brings the estimated elasticity up to -0.82 while controlling extensively for demographics brings the elasticity back to -0.79. When time trends are included, the elasticity increase to -0.82, just 2.5 percent higher than the original elasticity. A similar pattern holds for each of the other crime types with the largest change from column (1) to column (6) occuring for assault and larceny, both of which are imprecisely estimated. Referring to columns (7)-(12), the reflected estimates follow a similar pattern with the exception of murder which appears to be somewhat sensitive to the inclusion of lagegd economic covariates, a result which drives the difference between columns (7) and (12) for violent crimes and for the cost-weighted crime index. While the estimated murder coefficient changes with the inclusion of controls, it is nevertheless similar in magnitude to the estimate that conditions only on the state-by-year effects.[76]

As discussed, the elasticities reported in the paper condition on two measures of city population growth, taken from the UCR and ASG data systems, respectively. The motivation for including both measures is that we are persuaded there is measurement error in each series individually. As discussed in Section VI, above, it is necessary to smooth both series to circumvent clear measurement problems around Census years. However, even after smoothing, it may not be the case that true population growth rates can be represented as a linear combination of the growth rates of the UCR and ASG series. To assess the extent to which measurement error in population represents a source of bias for the estimated police elasticities, we take advantage of two additional proxies for a city's population: (1) population data from the Surveillance Epidemiology and End Results (SEER) dataset which has been compiled by the National Cancer Institute to track disease incidence and (2) the number of births in a city, drawn from the National Center for Health Statistics at the Centers for Disease Control. Births are correlated with population because, other things equal, the more individuals who are living in a city, the more children will be born. Moreover, unlike the UCR and ASG series on population, there is close to no measurement error in the number of births in a city, since births are estimated from no worse than 50 percent samples of birth certificates over the sample period and since birth certificates cover an estimated 99 percent of births in the U.S. over this time period.[77] Consequently, including the growth rate in births as a covariate should pick up on any remaining association between true population growth rates and crime growth rates.

Appendix Table 1 shows the sensitivity of our resulting estimates to the inclusion of SEER population

[76] We also consider whether the estimates are robust to the exclusion of the two largest cities in the sample–New York and Los Angeles–as well as whether the results are robust to the exclusion of cities with various data problems, namely those cities which have merged with their respective counties (e.g., Jacksonville, Nashville, Charlotte and Louisville) and cities which have been recently found to have misreported data to the FBI's Uniform Crime Reporting System (e.g., Milwaukee). When these cities are excluded from the sample, the estimates are nearly identical to those reported in Table 5.

[77] In the early 1970s, the NCHS transitioned to 100 percent samples of birth certificates.

data and the births data. For the years and cities for which data on births are available, the estimates change very little when the growth rate in births is added to the model. For example, referring to Panel A for which data on all cities in the sample are available for the 1960-1993 subsample, we see that pooled estimates of each of the crime elasticites are extremely similar with and without the inclusion of the births measure. For example, the violent crime elasticity moves from -0.16 to -0.18 while the property crime elasticity changes only in the third decimal place.

Panel B which provides estimates for a sample of 147 of our cities using the 1960-2003 window. For that sample, the largest impacts are on murder and motor vehicle theft where the elasticity changes from -0.548 to -0.565 and from -0.346 to -0.369, respectively, when births are included. Finally, in Panel C, we test the sensitivity of the pooled elasticities to inclusion of the SEER population data over the 1970-2008 time period. Again, the estimates are extremely similar when the SEER population measure is added, with motor vehicle theft showing the largest change (-0.320 versus -0.340). We interpret these findings as indicating that our estimates are not importantly compromised by measurement error in population. Indeed, to the extent that our estimates do change when additional population controls are added, they tend to get larger in magnitude with additional population controls, suggesting that our full sample estimates may be conservative.[78]

VIII. Discussion

The estimates reported in the previous section of this paper can be thought of as police elasticities that are robust to errors in the measurement of police. Pooled estimates in Table 5 represent our best guess regarding crime-specific police elasticities. Pooling via GMM or EL, we obtain precisely estimated elasticities of -0.34 for violent crimes and -0.17 for property crimes, with especially large elasticities for murder (-0.67), robbery (-0.56), motor vehicle theft (-0.34) and burglary (-0.23).

In this section, we contextualize these findings by comparing our reported elasticities to those in the prior literature. Table 8 presents selected police elasticities from eight recent papers that use U.S. data. Each of the papers explicitly seeks to correct for simultaneity bias, for which our estimates do not adjust. While these papers do not discuss the possibility of measurement error in police or in population, an IV estimator using exogenous instruments will correct for both simultaneity bias and measurement error bias under the classical measurement error hypothesis.

[78] A final issue that is worth mentioning is the possibility of displacement—an increase in policing in one jurisdiction might displace crime to a nearby jurisdiction. If this is the case, then our approach will tend to overestimate the social value of policing, since part of the apparent crime reduction associated with increasing policing would stem from a simple reshuffling of criminal activity. Appendix Table 2 addresses this concern. In this table, we contrast the GMM estimates presented in Table 5 with estimates based on aggregating up to the MSA level. The estimates in Table 2 indicate that there is not enough statistical power in these data to distinguish the estimates at the city-level from those at the MSA-level. If anything, the estimates at the MSA are larger, rather than smaller, than the estimates at the city level, which is the opposite of the expected pattern if displacement were a first-order phenomenon in these data.

Looking across the estimates from these papers in Table 8, four tendencies are evident. First, the estimates are generally negative. Some of the estimates are zero (e.g., Levitt (1997) for property crime), but virtually none are positive.[79] Second, the general tendency of these estimates is similar to, or perhaps slightly larger than, that of our own estimates. For example, the average of the murder elasticities is 1.18 in magnitude.[80] This is similar to the magnitude of our own estimated murder elasticity (roughly 1) when we replace our preferred state-by-year effects specification with a simpler year effects specification, which is more similar to most of the research designs employed in the previous literature. Similarly, the average of the elasticities for robbery, burglary, and auto theft is approximately 0.79, 0.35, and 0.77 in magnitude, respectively. Controlling for year effects, our estimates of the same quantities are roughly 0.50, 0.17, and 0.50, respectively. The differences between these estimates would likely not rise to the level of statistical significance, but the general tendency is for our estimates to be slightly smaller in magnitude. Some of this discrepancy stems from utilization of different time periods. For example, when we restrict our analysis to the years analyzed by Evans and Owens (2007), namely 1990 to 2001, our estimated elasticities are -0.83 and -0.31, for violent and property crimes, respectively. These are extremely close in magnitude to those in Evans and Owens (2007) (-0.99 for violent crimes and -0.26 for property crimes). Given the magnitude of the standard errors, the differences in estimates are likely consistent with the hypothesis of sampling volatility.

Third, there is a general tendency to find that police have a larger protective effect on violent crimes than on property crimes. This is a surprising finding if we conceive of the estimated effect of police on crime as being about deterrence. However, as noted in the introduction, the effect of police on crime operates through both a deterrence and an incapacitation channel. Moreover, police departments actively focus their resources on the incapacitation of individuals posing the greatest risk to society, which may make the incapacitation channel particularly important.

Fourth, the estimated elasticities tend to be quite imprecise, with estimated standard errors ranging from 0.2 to 0.7 for violent crimes and 0.2 to 0.9 for property crimes. As a result, it is often the case that even large elasticites (on the order of 1) cannot be rejected as being different from zero. Similarly, the cross-crime pattern of the elasticities is difficult to discern. For example, one of the more precise studies is that of Evans and Owens (2007). In that study, the magnitude of the estimated elasticities and standard errors suggest that it would be difficult to reject tests of the equality of various crime-specific elasticities. As a result, though

[79] The pooled estimates in Levitt (1997) are in error due to a mistake in the use of weights (McCrary 2002). The numbers in Table 8 listed as Levitt (1997) are actually the corrected numbers reported in McCrary (2002) that use Levitt's mayoral election year series. The numbers in Table 8 listed as McCrary (2002) are the numbers reported in McCrary (2002) that use McCrary's mayoral election year series.

[80] To avoid double-counting research designs, we count the average of the estimates from Levitt (1997) and McCrary (2002) as a single entry.

the general pattern of the elasticities is suggestive, it is difficult to draw inferences about even the most basic policy questions such as the relative effectiveness of police in reducing violent versus property crimes.

The elasticities we report in this research are estimated with considerably greater precision, with standard errors that are between one-quarter and one-half the size of those reported by Evans and Owens (2007) and up to an order of magnitude smaller than those reported in other papers. The result is that we are able to generate considerably stronger inferences regarding the cross-crime pattern of the elasticities.

In Table 9, we formalize this idea and test the equality of all pairs of individual crime elasticities. The table reports p-values from each of these tests, operationalized by stacking up crime categories into a broader GMM system. For a given row, a given column reports the p-value associated with a test of the equality of the coefficient for the crime category on the row and the coefficient for the crime category on the column. The pattern of the resulting p-values suggests that we can be confident that police reduce murder to a greater extent than assault and larceny and perhaps burglary. Likewise, the effect of police on robbery is greater than it is for assault, burglary and larceny and the effect of police on motor vehicle theft is greater than the effect of police on larceny. Referring to the aggregates, the elasticities for murder and robbery are greater than the property crime elasticity. We can also reject, at the 10 percent level, the equality of the violent and property crime elasticities. Despite a dominant pattern in the literature that suggests that the effect of police on crime is most concentrated among violent crimes, to our knowledge, this is the first paper that offers more than suggestive evidence on this point.

Whether our estimates are similar to or different from those in the preceding literature is important for getting the magnitude of police elasticities right, but is also interesting because it speaks to the broader issue of whether simple regression techniques are compromised by simultaneity bias. If our estimates are deemed to be similar to those reported in prior research, then our research implies a smaller role for simultaneity than has been suggested by prior studies.

Overall, our suspicion is that the estimates we have presented here are compromised somewhat by simultaneity bias, despite our best efforts to control for unobserved heterogeneity. The sign of the bias, as criminologists and economists have argued for several decades now, is likely positive, leading our approach to underestimate of the magnitude of the policing elasticity. Thus, the correct magnitude is likely at least as large as what our results indicate. As we turn in the next section to connecting our estimates to the state's optimal level of policing, these considerations should be kept in mind, as they suggest that our policy conclusions may well be conservative.

IX. Cost-Benefit Analysis

A. National Estimates

The results presented in Table 5 represent our best estimate of the elasticity of each type of crime with respect to police. These elasticities allow us to predict the change in reported crimes expected to arise from a given percent increase in the size of a city's police force. However, in allocating scarce resources among a large number of critical public services a potentially more relevant parameter is the ratio of the benefits to the costs of hiring additional police personnel. In Section II, we established that even in the presence of investment in precautions with externalities, the state's optimal choice of policing can be characterized by the the parameter θ, which represents the elasticity of the cost of crime with respect to police, holding precautions fixed. In particular, the rule-of-thumb outlined in Section II is that hiring police improves welfare when

$$|\theta| / \frac{wS}{nC} \equiv \kappa > 1 \tag{31}$$

In this section, we use the GMM approach described above to estimate the ratio of the benefits (as proxied by averted costs to potential victims) to the costs of police.

For a VSL of \$7 million, we estimate a police elasticity of the cost of crime of -0.47 (standard error = 0.17). This elasticity estimate is based on a model including state-by-year effects and two controls for population, analogous to our preferred specification in Table 5. Scaling this elasticity estimate by the ratio of mean victimization costs to mean police expenditures produces an estimate of the 2010 social dollars saved from increasing spending on police by one dollar, or the benefit-cost ratio (BCR).[81] Varying the VSL from \$1 to \$28 million, our GMM estimate of the police elasticity of the cost of crime ranges from -0.32 (standard error = 0.09) to -0.55 (standard error = 0.26).

An unfortunate feature of these types of estimates is that benefit-cost calculations are often extremely sensitive to the monetized value of an averted murder.[82] Figure 10 provides a visual presentation of the findings from this analysis. The figure plots the BCR that follows from this GMM procedure on the vertical axis against possible VSL estimates on the horizontal axis. The change in the BCR is linear with respect to the VSL employed since the VSL is simply the factor by which murders are scaled in the analysis. The

[81]To obtain the cost of increasing policing, we take the average of the UCR and ASG counts and scale it by \$130,000, an estimate of the fully-loaded cost of a police officer in 2010. As discussed in detail in Section VI, above, this estimate is based on data on the operating budget per officer, i.e., the ratio of the operating budget for the police department, in 2010 dollars, to the number of sworn officers. These figures are taken from the ASG Finance and ASG Employment files for 2003-2010. We use multiple years to get a clear picture of the finances for a department. These figures fluctuate a good deal from year to year. We use the city-specific median over time, and then compute a 2010 population weighted average of the city-specific medians; this weighted average is about \$130,000. Our estimate of the cost of hiring additional officers is notably higher than those used in some of the literature (e.g., Evans and Owens (2007) use \$55,000).

[82]In the literature, it is not uncommon for the results of a benefit-cost analysis of a given policy to depend on the researcher's choice between two reasonable alternative values of the cost of a murder.

BCR ranges from approximately 0.4 at a VSL of \$1 million to approximately 6.0 at a VSL of \$28 million. To further narrow down these estimates, we superimpose a kernel density estimate of the density of the 64 VSL estimates for the U.S. While the estimates vary considerably, approximately 80 percent of the data lies below \$10 million which is associated with an approximate BCR of 2. At \$7 million, the mean value of the VSL, the resulting BCR is 1.63, indicating that, in a typical U.S. city, an additional dollar allocated towards policing is predicted to save \$1.63 in costs to crime victims. This would be consistent with classical notions of the underprovision of public goods (Samuelson 1954). On the other hand, as noted there is substantial ambiguity regarding VSL estimates. The estimated VSL from Ashenfelter and Greenstone (2004) implies a BCR of roughly \$0.80, indicating substantial overpolicing.[83] If we revert to estimates from Cohen and Piquero (2008), the BCR is just below 1, suggesting that the political process may have arrived at the social optimum.

B. City-specific Estimates

The national benefit-cost ratios reported above answer the question: For a typical U.S. city in a typical year in our sample, what is the dollar value of crime reduction obtained by increasing spending on police by one dollar? A somewhat different question pertains to specific U.S. cities in 2010. For example, for many years Oakland, California, has had fewer police per capita than other cities, despite a relatively high crime rate. Journalists often note this fact and query whether Oakland should hire additional police (e.g., McKinley 2009). We now seek to answer the question: For specific U.S. cities in 2010, given that the value of a statistical life is \$7 million, what is the dollar value of crime reduction obtained by increasing spending on police by one dollar? Despite great interest in such issues, we emphasize that investigations along these lines are necessarily somewhat speculative, as data for individual cities are less reliable than data for a few hundred cities, taken as a whole. Indeed, such an analysis may be heroic, as it involves assuming that the police elasticity of crime is constant across cities, across time, and across possible adjustments to the size of the police force.

These limitations aside, we believe it is nonetheless of interest to characterize the heterogeneity across cities in a benefit-cost ratio, as a function of the prevalence of crime, the number of officers, and the cost to the city of hiring officers. There is extraordinary heterogeneity across cities in the prevalence of crime. For example, the cost-weighted sum of crimes per capita in the most dangerous city in our sample (Gary, Indiana), is nearly 40 times that of the safest city in our sample (Waltham, Massachusetts). Similarly, cities vary quite a lot in terms of the number of officers and the expense per officer. We suspect that our approach, while flawed, captures much of the variation from city to city in the true benefit-cost ratio.

[83]In fact, while estimates of the VSL arising from the study of individuals' labor market behavior tend to yield large values (on average, \$9.5 million), estimates of the VSL arising from the study of non-labor market behaviors tend to yield much smaller values (on average, \$4 million).

Table 10 presents the bottom and top 30 cities in our sample, based on our estimated benefit-cost ratio. The benefit-cost ratio, reported in the final column (column 9) of the table, is rooted in our overall estimated police elasticity of the crime index for all cities and all years, but is scaled by the mean cost of crime in the city from 2003-2010 relative to the product of the average number of officers from 2003-2010 and the estimated cost of hiring an additional officer (see Section VI and the Data Appendix for discussion of the construction of this variable). The mean cost of crime is reported in per capita terms in the table (column 6), as is the estimated cost per officer (column 8). Column 5 of the table reports per capita income in the city, and column 7 reports the cost of crime relative to income per capita ("fraction income at risk"). This last column encourages thinking of crime as a tax on the populace. For reference, we additionally report city population as of 2010 and the city's poverty rate.

The 30 cities listed in the top half of the table have the lowest benefit-cost ratios among our 242 cities, while the 30 cities listed in the bottom half have the highest benefit-cost ratios. For example, for Sunnyvale, California, we estimate that every dollar spent on policing yields only 20 cents in benefits in terms of crime reduction. In contrast, we estimate that every dollar spent on policing in Gary, Indiana, yields $14 in benefits in terms of crime reduction. The population weighted average of the city-specific benefit-cost ratio is about $1.78, or slightly higher than our estimate for the overall sample reported above.

Scanning down the table, we see several interesting patterns. Cities with low benefit-cost ratios are small, low-poverty cities, with low levels of crime and low to moderate levels of policing. Police officers in these cities often enjoy high salaries and benefits, leading to high employer costs per officer. Sunnyvale and Berkeley, for example, both have costs per officer of roughly $280,000.

Cities with high benefit-cost ratios are suprisingly representative of our broader sample in some regards. For example, cities with high benefit-cost ratios include both low and high population cities. Also, these cities sometimes have low policing levels (e.g., Oakland and Richmond, California, have 180 and 160 sworn officers per 100,000 population, respectively) and sometimes have high policing levels (e.g., Baltimore and Camden have 480 and 510 sworn officers per 100,000 population). On the other hand, cities with high benefit-cost ratios have high poverty rates and extraordinarily high crime rates—generally an order of magnitude higher than cities with low benefit-cost ratios. Crime costs residents in these cities anywhere from 5 percent of their annual income (Mobile, Alabama) to 34 percent of their annual income (Camden). In contrast, for cities with low benefit-cost ratios, crime costs resident at most 1 percent of their annual income.

Another interesting pattern is that California cities are prevalent among the lowest benefit-cost ratio cities, with 13 out of 30 spots, but also are represented among the highest benefit-cost ratio cities (Oakland, Richmond, and San Bernadino). The estimated cost per officer is very high among California cities generally. High costs per officer keeps several high-crime California cities from being among the highest benefit-cost ratio cities (Sacramento, Vallejo). Richmond's high estimated costs ($240,000) are particularly remarkable given its high poverty rate and low per capita income; wealthy Palo Alto's estimated cost per officer is slightly lower than that for Richmond. Factoring in base salary, overtime, and lump-sum payments, a police officer in Richmond makes an average of $148,000, or six times what a city resident makes.[84] It is worth noting that the estimated cost per officer does not include unfunded pension liabilities for the city, which is an ongoing issue for many cities that may lead the figures for officer expense to be understated (Gralla 2012).

As noted above, there is substantial ambiguity regarding some of the inputs to the city-specific benefit-cost ratios. However, we note that in many cases the benefit-cost ratios are sufficiently extreme that only gross errors in the inputs would alter the conclusion that the benefit-cost ratio was on the wrong side of 1. Our sense is that cities with benefit-cost ratios between 0.5 and 1.5 may well be near the optimal level of policing, but that the many cities outside this band are unlikely to be.

C. Police Incapacitation Effects and the Benefit-Cost Ratio

The estimates in the preceeding sub-sections are valid under the assumption that either (i) the decline in crime resulting from increased police is entirely due to deterrence or that (ii) the cost of incarcerating offenders is fixed in the short run so that the downstream cost of incapacitating offenders need not be counted as a cost of increased police personnel. Here, we re-frame the national benefit-cost analysis, treating the expected increase in incarceration resulting from more police as an additional cost of hiring a new officer. Because we are interested in the short-run costs and benefits of new police hiring, we count only the costs of incarceration that are borne in the first year. We begin with an estimate of the number of arrests per officer. Using our sample of 242 cities, an average officer made between 18.7 and 20.2 arrests in 2010, depending on whether the UCR or ASG officer count is employed.[85] Next, we employ an estimate of the conditional probability of a conviction given an arrest. In 2010, there were 13,120,947 arrests made by police officers in the United States while there were 1,132,290 convictions in state courts and another 81,934 convictions in federal courts. Dividing convictions by arrests yields an estimated conditional probability of a conviction of 9.3 percent. Of defendants sentenced in state courts, 40 percent were sentenced to state prison (with a mean sentence length of 4 years and 11 months), 28 percent were sentenced to a term in local jail (with a mean sentence of 6 months) and the remaining 32 percent were sentenced to a term of probation or an alternate penalty that did not involve incarceration. On average, offenders serve approximately 55 percent of their sentence. Thus, in steady state, a typical officer is associated

[84] Average take-home pay for officers based off of data pulled from the *San Jose Mercury News*. See footnote **??**.

[85] The working assumption here is that a new officer's productivity, and the lost productivity associated with laying off an officer, can be approximated using the productivity of an average officer.

with 20 new arrests, 1.85 new convictions and 0.87 incarceration-years.[86] At an incarceration cost of $25,000 per year, each new officer is thus associated with $21,738 in additional costs. Augmenting the salary figure with this estimate yields a benefit-cost estimate of $1.40 using the $7 million estimate of the value of a statistical life.

X. Conclusion

In this paper, we have presented estimates of the elasticity of crime with respect to police for index offenses: murder, rape, robbery, aggravated assault, burglary, larceny excluding motor vehicle theft, and motor vehicle theft. These estimates are based on annual data on year-over-year growth rates in crime and police in a panel data set of 242 cities observed from 1960-2010. Our primary specifications model growth rates in crime as a function of the growth rate in the number of sworn officers, population growth and a full set of state-by-year effects which render our estimates robust to arbitrary changes in state policy, such as penal policy, and other factors affecting cities in the same state similarly. In auxiliary regressions we show that our results are also robust to a wide array of local-level confounders.

We argue that a central problem in estimating the police elasticity of crime is measurement error in the number of police. These measurement errors are unimportant for specifications involving the level or log of police, but are first-order for specifications involving growth rates or city fixed effects. Problems with measurement errors in police have gone unaddressed in the crime literature, but add to a long list of literatures where measurement errors have been shown to be important (Ashenfelter and Krueger 1994, Bound, Brown, Duncan and Rodgers 1994, Kim and Solon 2005, Bollinger 2003, Black and Smith 2006, Edlin and Karaca-Mandic 2006, Nunn 2008).

A recent literature has focused on quasi-experimental estimates of the police elasticity of crime, and many of these estimates solve both for problems with measurement errors in police and for simultaneity bias (for a review, see Levitt and Miles 2006). We add to this literature by addressing the measurement error bias directly, utilizing independent measurements of the number of police departments collected annually by the Census Bureau. Correcting for measurement error increases least squares estimates of the police elasticity of crime by roughly a factor of 5 and results in estimates just slightly smaller than those estimated in the previous quasi-experimental literature. This may suggest a smaller role for simultaneity bias than has previously been emphasized. An advantage of bracketing the issue of simultaneity bias, and focusing instead on correcting for measurement errors, is that we obtain estimates that are demonstrably more precise than those from the previous literature and arguably conservative in magnitude. The received wisdom in this literature, going back to Nagin (1978) and before, is that police departments hire officers during and perhaps even in anticipation

[86] Regarding pre-trial detention, since the average length of time between arrest and sentence is approximately 5 months, we operate under the assumption that if an arrestee does not receive bail, their expected sentence amounts to time served.

of crime waves, leading the police elasticity of crime to be too small in magnitude.

Our best guess regarding the police elasticity of crime is -0.34 for violent crime and -0.17 for property crime. Crime categories where police seem to be most effective are murder, robbery, burglary, and motor vehicle theft, with estimated elasticities of -0.67, -0.56, -0.23, and -0.34, with standard errors of roughly 0.2 for murder and 0.1 for other crimes. The elasticity of the cost-weighted crime index is -0.47 with a standard error of 0.17.

To assess whether these magnitudes are small or large, we introduced a framework for assessing whether policing levels are socially desirable. This framework delivers a rule-of-thumb for optimal policing that pertains to a social planner unwilling to monitor precautions undertaken by individuals, where the social planner pays for policing using lump-sum taxes. This analysis, together with our empirical estimates of the police elasticity of the cost of crime, is suggestive of substantial underpolicing. However, as with many analyses pertaining to public investments and safety, our normative conclusions turn to a great extent on the price society is willing to pay for reductions in the probability of fatalities, or the value of a statistical life (VSL).

Despite the ambiguity regarding the appropriate quantity for the VSL, we note that federal and state regulatory authorities frequently undertake investments where the same tradeoff is confronted. Pegging policing investments to the typical federal standard suggests that society would receive approximately \$1.60 in benefits from an additional 2010 dollar spent on policing. This estimate is likely conservative if simultaneity bias in the police elasticity of crime is important.

This policy conclusion is most strongly justified if externalities in precautions are unimportant or if there is little scope for policing crowding out precautions. Precautions may however be important. Nonetheless, the framework we have introduced allows us to be clear about the assumptions supporting our policy conclusions in such a scenario. If precautions have negative externalities on average, with one individual's precaution displacing criminal activity to her neighbors, then the policy conclusion is conservative. If precautions have positive externalities on average, with one individual's precaution protecting her neighbors against criminal activity, then this channel has to be fully one-third as large as the direct effect of police on crime in order to overturn the policy conclusion.

References

Advisory Commission on Intergovernmental Relations, *Block Grants: A Comparative Analysis*, Washington, D.C.: US Advisory Commission on Intergovernmental Relations, October 1977.

——, *State Limitations on Local Taxes and Expenditures*, Washington, D.C.: US GPO, March 1977.

——, *Tax and Expenditure Limits on Local Governments: An Information Report*, Indiana University: Center for Urban Policy and the Environment, March 1995.

Ashenfelter, Orley and Alan Krueger, "Estimates of the Economic Returns to Schooling from a New Sample of Twins," *American Economic Review*, December 1994, *84* (5), 1157–1173.

—— **and Michael Greenstone**, "Using Mandated Speed Limits to Measure the Value of a Statistical Life," *Journal of Political Economy*, February 2004, *112* (S1), S226–S267.

Ayres, Ian and Steven D. Levitt, "Measuring Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack," *The Quarterly Journal of Economics*, 1998, *113* (1), 43–77.

Baltagi, Badi H., "Estimating an Economic Model of Crime Using Panel Data From North Carolina," *Journal of Applied Econometrics*, 2006, *21* (4), 543–547.

Banfield, Edward C. and James Q. Wilson, *City Politics*, Cambridge: Harvard University Press and MIT Press, 1963.

Beccaria, Cesare, *On Crimes and Punishments*, Oxford: Clarendon Press, 1764.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, March/April 1968, *76* (2), 169–217.

Bentham, Jeremy, *An Introduction to the Principles of Morals and Legislation*, Oxford: Clarendon Press, 1789.

Berk, Richard and John MacDonald, "Policing the Homeless: An Evaluation of Efforts to Reduce Homeless-Related Crime," *Criminology & Public Policy*, 2010, *9* (4), 813–840.

Black, Dan A. and Jeffrey A. Smith, "Returns to College Quality with Multiple Proxies for Quality," *Journal of Labor Economics*, 2006, *24* (3), 701–728.

Bollinger, Christopher R., "Measurement Error in Human Capital and the Black-White Wage Gap," *The Review of Economics and Statistics*, 2003, *85* (3), 578–585.

Boston Police Department, *Annual Report: 1919,* 1919.

——, *Annual Report: 1940,* 1940.

——, *Annual Report: 1981-1982,* 1982.

Bound, John and Alan B. Krueger, "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?," *Journal of Labor Economics*, January 1991, *9* (1), 1–24.

——**, Charles Brown, and Nancy Mathiowetz**, "Measurement Error in Survey Data," in James J. Heckman and Edward Leamer, eds., *Handbook of Econometrics*, Vol. 5, New York: Elsevier, 2001, pp. 3705–3843.

——**,** ——**, Greg J. Duncan, and Willard L. Rodgers**, "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics*, 1994, *12* (3), 345–368.

——**, David A. Jaeger, and Regina M. Baker**, "Problems with Instrumental Variables Estimation when the Correlation between the Instrument and the Endogenous Explanatory Variable Is Weak," *Journal of the American Statistical Association*, June 1995, *90* (430), 443–450.

Braga, Anthony, "Hot Spots Policing and Crime Prevention: A Systematic Review of Randomized Controlled Trials," *Journal of Experimental Criminology*, September 2005, *1* (3), 317–342.

Braga, Anthony A., "The Effects of Hot Spots Policing on Crime," *The Annals of the American Academy of Political and Social Science*, 2001, *578*, 104–125.

—— **and Brenda J. Bond**, "Policing Crime and Disorder Hot Spots: A Randomized Controlled Trial," *Criminology*, 2008, *46* (3), 577–607.

——**, David L. Weisburd, Elin J. Waring, Lorraine Green Mazerolle, William Spelman, and Francis Gajewski**, "Problem-Oriented Policing in Violent Crime Places: A Randomized Controlled Experiment," *Criminology*, 1999, *37* (3), 541–580.

——**, David M. Kennedy, Elin J. Waring, and Anne Morrison Piehl**, "Problem-Oriented Policing: Deterrence, and Youth Violence: An Evaluation of Boston's Operation Ceasefire," *Journal of Research in Crime and Delinquency*, 2001, *38* (3), 195–225.

Burdett, Kenneth, Ricardo Lagos, and Randall Wright, "An On-the-Job Search Model of Crime, Inequality, and Unemployment," *International Economic Review*, August 2004, *45* (3), 681–706.

Cameron, Samuel, "The Economics of Crime Deterrence: A Survey of Theory and Evidence," *Kyklos*, 1988, *41* (2), 301–323.

——, "On the Welfare Economics of Capital Punishment," *Australian Economic Papers*, 1989, *28* (53), 253–266.

Card, David E., "The Causal Effect of Education on Earnings," in Orley Ashenfelter and David E. Card, eds., *The Handbook of Labor Economics*, Vol. 3A, Amsterdam: Elsevier, 1999.

Chetty, Raj, "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics*, 2006, *90*, 1879–1901.

——, "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics, Annual Reviews*, 2009, *1* (1), 451–488.

Chicago Tribune, "Freeze U.S. Funds to Police: Robinson," *Chicago Tribune*, September 8, 1972.

City of Boston, "Financial Management of the City," *FY07 Adopted Budget by Cabinet (Volume I)*, 2007.

Clingermayer, James C. and Richard C. Feiock, *Institutional Constraints and Policy Choice: An Exploration of Local Governance*, Albany: State University of New York Press, 2001.

Cohen, Mark A., "Measuring the Costs and Benefits of Crime and Justice," *Criminal Justice*, 2000, *4*, 263–315.

—— **and Alex R. Piquero**, "New Evidence on the Monetary Value of Saving a High Risk Youth," *The Journal of Quantitative Criminology*, March 2008, *25* (1), 25–49.

Cope, Glenn H., "Walking the Fiscal Tightrope: Local Budgeting and Fiscal Stress," *International Journal of Public Administration*, 1992, *15* (5), 1097–1120.

Cornwell, Christopher and William N. Trumbull, "Estimating the Economic Model of Crime with Panel Data," *Review of Economics and Statistics*, May 1994, *76* (2), 360–366.

Dardick, Hal, "Prentice H. Marshall Sr., 77: "Activist Judge" and Proud of It," *Chicago Tribune*, May 26, 2004.

Dasgupta, Partha and Eric Maskin, "The Existence of Equilibrium in Discontinuous Economic Games, I: Theory," *Review of Economic Studies*, January 1986, *53* (1), 1–26.

Davis, Robert, "Police 'Boot Camp' Reopens: 1st Crop of City Recruits in 2 Years at Academy," *Chicago Tribune*, March 3, 1985.

Deslippe, Dennis A., ""Do Whites Have Rights?": White Detroit Policemen and "Reverse Discrimination" Protests in the 1970s," *The Journal of American History*, 2004, *91* (3), 932–960.

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

Dodd, Philip, "Goldwater Calls for Nation-wide War on Crime," *Chicago Tribune*, October 8, 1964, p. 7.

Eck, John E. and Edward R. Maguire, "Have Changes in Policing Reduced Violent Crime? An Assessment of the Evidence," in Afred Blumstein and Joel Wallman, eds., *The Crime Drop in America*, New York: Cambridge University Press, 2000, pp. 207–265.

Edlin, Aaron S. and Pinar Karaca-Mandic, "The Accident Externality from Driving," *Journal of Political Economy*, October 2006, *114* (5), 931–955.

Ehrlich, Isaac, "The Deterrent Effect of Criminal Law Enforcement," *Journal of Legal Studies*, June 1972, *1* (2), 259–276.

Enstad, Robert, "Daley Bows to Expediency in Agreeing to Police Quotas," *Chicago Tribune*, June 27, 1976, p. 28.

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

——**, Craig Garthwaite, and Timothy J. Moore**, "Stalled Progress: Black-White Education and Labor Market Differences and the Long-term Effects of the Crack Cocaine Epidemic," *Unpublished manuscript*, 2012.

Fajnzylber, Pablo, Daniel Lederman, and Norman Loayza, "What Causes Violent Crime?," *European Economic Review*, 2002, *46* (7), 1323–1357.

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

Fuller, Wayne A., *Measurement Error Models*, New York: Wiley, 1987.

Gralla, Joan, "U.S. Munis Face $2 Trillion in Unfunded Pension Costs," *Reuters*, July 2, 2012.

Greisinger, George, Jeffrey S. Slovak, and Joseph J. Molkup, *Civil Service Systems: Their Impact on Police Administration*, Washington, D.C.: U.S. GPO, 1979.

Guggenberger, Patrik and Jinyong Hahn, "Finite Sample Properties of the Two-Step Empirical Likelihood Estimator," *Econometric Reviews*, 2005, *24* (3), 247–263.

Harriston, Keith A. and Mary Pat Flaherty, "D.C. Police Paying for Hiring Binge," *Washington Post*, August 28 1994.

Hausman, Jerry A., "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left," *Journal of Economic Perspectives*, Autumn 2001, *15* (4), 57–67.

Hevesi, Alan G., *Revenue Sharing in New York State*, Albany: Office of the New York State Comptroller, 2005.

Imbens, Guido W., "A New Approach to Generalized Method of Moments Estimation," *Harvard Institute for Economic Research Working Paper*, 1993, *1633.*

——, "Generalized Method of Moments and Empirical Likelihood," *Journal of Business and Economic Statistics*, October 2002, *20* (4), 493–506.

———, **Richard H. Spady, and Phillip Johnson**, "Information Theoretic Approaches to Inference in Moment Condition Models," *Econometrica*, March 1998, *66* (2), 333–357.

Ireton, Gabriel, "Retirement Surge Grips City Police, Concerns Council," *Pittsburgh Post-Gazette*, May 10, 1976, p. 17.

James, Nathan, *Edward Byrne Memorial Justice Assistance Grant Program: Legislative and Funding History*, Washington, D.C.: Congressional Research Service, 2008.

Joyce, Philip G. and Daniel R. Mullins, "The Changing Fiscal Structure of the State and Local Public Sector: The Impact of Tax and Expenditure Limitations," *Public Administration Review*, May - June 1991, *51* (3), 240–253.

Katz, Matt and Darran Simon, "Daytime Camden Police Patrols Hit Hard by Layoffs," *Philadelphia Inquirer*, January 20, 2011.

Kim, Bonggeun and Gary Solon, "Implications of Mean-Reverting Measurement Error for Longitudinal Studies of Wages and Employment," *Review of Economics and Statistics*, February 2005, *87* (1), 193–196.

Klick, Jonathan and Alex Tabarrok, "Using Terror Alert Levels to Estimate the Effect of Police on Crime," *Journal of Law and Economics*, April 2005, *48* (1), 267–280.

Koper, Christopher S., Edward R. Maguire, and Gretchen E. Moore, *Hiring and Retention Issues in Police Agencies: Readings on the Determinants of Police Strength, Hiring and Retention of Officers, and the Federal COPS Program*, Washington, D.C.: Urban Institute, 2001.

LA Times, "Early Retirements Thin Police Force," *Los Angeles Times*, February 20, 1966, pp. SF–B1.

Larkin, Al, "Challenge to Law on Police Hiring Nearing Finish," *Boston Globe*, November 25, 1973, p. 59.

Lee, David S. and Justin McCrary, "The Deterrence Effect of Prison: Dynamic Theory and Evidence," July 2009. Unpublished manuscript, University of California, Berkeley.

Levitt, Steven D., "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *American Economic Review*, June 1997, *87* (3), 270–290.

———, "The Relationship Between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports," *Journal of Quantitative Criminology*, March 1998, *14*, 61–81.

——— **and Thomas J. Miles**, "Economic Contributions to the Understanding of Crime," *Annual Review of Law and Social Science*, 2006, *2*, 147–164.

Lewis, Carol W., "Budgetary Balance: The Norm, Concept, and Practice in Large U. S. Cities," *Public Administration Review*, November - December 1994, *54* (6), 515–524.

Lochner, Lance, "Education, Work, and Crime: A Human Capital Approach," *International Economic Review*, August 2004, *45* (3), 811–843.

Lubotsky, Darren H. and Martin Wittenberg, "Interpretation of Regressions with Multiple Proxies," *Review of Economics and Statistics*, August 2006, *88* (3).

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

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

Massachusetts Department of Revenue, *Levy Limits: A Primer on Proposition* $2\frac{1}{2}$ June 2007.

McCrary, Justin, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment," *American Economic Review*, September 2002, *92* (4), 1236–1243.

———, "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police," *American Economic Review*, March 2007, *97* (1), 318–353.

———, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, February 2008, *142* (2).

———, "Dynamic Perspectives on Crime," in Bruce Benson, ed., *Handbook of the Economics of Crime*, Northampton, MA: Edward Elgar, 2009.

McKinley, Jesse, "New Oakland Police Chief Inherits a Force, and a City, in Turmoil," *New York Times*, October 15, 2009.

Nagin, Daniel, "General Deterrence: A Review of the Empirical Evidence," in Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin, eds., *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, Washington, D.C.: National Academy of Sciences, 1978, pp. 95–139.

———, "Criminal Deterrence Research at the Outset of the Twenty-First Century," in Michael Tonry, ed., *Crime and Justice: A Review of Research*, Vol. 23, Chicago: University of Chicago Press, 1998, pp. 1–42.

National Advisory Commission on Civil Disorders, *Report of the National Advisory Commission on Civil Disorders*, Washington, D.C.: GPO, 1968.

Newey, Whitney, "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, September 1985, *29* (3), 229–256.

Newey, Whitney K. and Richard J. Smith, "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Econometrica*, January 2004, *72* (1), 219–255.

Nunn, Nathan, "The Long-term Effects of Africa's Slave Trades," *The Quarterly Journal of Economics*, 2008, *123* (1), 139–176.

O'Shea, Robert, "State, Federal Budget Cuts Mean Trouble for Chicago," *Chicago Tribune*, March 8, 1981.

Pearson, Drew, "GOP Plots Strategy on Crime," *Washington Post*, October 20, 1964.

Police Executive Research Forum, "The Cop Crunch: Identifying Strategies for Dealing with the Recruiting and Hiring Crisis in Law Enforcement," December 30, 2005. Grant Final Report.

—, "Police Department Budgeting: A Guide for Law Enforcement Chief Executives," November 2005. Final Draft.

Polinsky, A. Mitchell and Steven Shavell, "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature*, March 2000, *38* (1), 45–76.

Poterba, James M. and Kim S. Rueben, "The Effect of Property-Tax Limits on Wages and Employment in the Local Public Sector," *American Economic Review*, May 1995, *85* (2), 384–389.

Qin, Jing and Jerry Lawless, "Empirical Likelihood and General Estimating Equations," *The Annals of Statistics*, 1994, *22* (1), 300–325.

Recktenwald, William, "1 in 5 Hopefuls Gets Into Police Training," *Chicago Tribune*, November 12, 1986, p. A3.

—, "Surge Of Retirements Strips Streets Of Police," *Chicago Tribune*, November 9, 1986, p. B1.

Richardson, Charles, *The State of State-Local Revenue Sharing*, Washington, D.C.: US Advisory Commission on Intergovernmental Relations, 1980.

Richman, Daniel, "The Past, Present, and Future of Violent Crime Federalism," in Michael Tonry, ed., *Crime and Justice: A Review of Research*, Vol. 34, Chicago: University of Chicago Press, 2006, pp. 377–439.

Rothenberg, Thomas J., "Approximating the Distributions of Econometric Estimators and Test Statistics," in Zvi Griliches and Michael D. Intriligator, eds., *The Handbook of Econometrics*, Vol. 2, Amsterdam: North-Holland, 1984, pp. 882–935.

Rubin, Irene S., *The Politics of Public Budgeting: Getting and Spending, Borrowing and Balancing*, Chatham, NJ: Chatham House Publishers, 1997.

Russell, Rosalind, *A City in Terror: Calvin Coolidge and the 1919 Boston Police Strike*, Boston: Beacon Press, 1975.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68* (1), 205.

Salanick, Gerald R. and Jeffrey Pfeffer, "Constraints On Administrator Discretion: The Limited Influence of Mayors on City Budgets," *Urban Affairs Review*, June 1977, *12* (4), 475–498.

Samuelson, Paul, "The Pure Theory of Public Expenditure," *Review of Economics and Statistics*, November 1954, *36* (4), 387–389.

Schwochau, Susan, Peter Feuille, and John Thomas Delaney, "The Resource Allocation Effects of Mandated Relationships," *Administrative Science Quarterly*, September 1988, *33* (3), pp. 418–437.

Shadbegian, Ronald J., "Do Tax and Expenditure Limitations Affect Local Government Budgets? Evidence From Panel Data," *Public Finance Review*, March 1998, *26* (2), 1–18.

—, "The Effect of Tax and Expenditure Limitations on the Revenue Structure of Local Government, 1962-87," *National Tax Journal*, 1999, *52* (2), 221–238.

Sherman, Lawrence W. and David Weisburd, "General Deterrent Effcts of Police Patrol in Crime "Hot Spots": A Randomized Controlled Trial," *Justice Quarterly*, 1995, *12* (4), 625–647.

— **and Dennis P. Rogan**, "Effects of Gun Seizures on Gun Violence: "Hot Spots" Patrol in Kansas City," *Justice Quarterly*, 1995, *12* (4), 673–693.

Siskin, Bernard R. and David W. Griffin, *Analysis of Distributions by Rank, Race, and Gender: City of Chicago Police Department, 1987-1991*, Philadelphia: Center for Forensic Economic Studies, 1997.

Skogan, Wesley and Kathleen Frydl, *Fairness and Effectiveness in Policing: The Evidence*, Washington, D.C.: National Academies Press, 2004.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge, "What Are We Weighting For?," *Unpublished manuscript, Michigan State University*, March 8, 2012.

Stigler, George J., "The Optimum Enforcement of Laws," *Journal of Political Economy*, May/June 1970, *78* (3), 526–536.

Tirole, Jean, *The Theory of Industrial Organization*, Cambridge: MIT Press, 1988.

Varon, Jay N., "A Reexamination of the Law Enforcement Assistance Administration," *Stanford Law Review*, 1975, *27* (5), 1303–1324.

Weisburd, David, "Hot Spots Policing Experiments and Criminal Justice Research: Lessons from the Field," *The Annals of the American Academy of Political and Social Science*, 2005, *599*, 220–245.

———, **Cody W. Telep, Joshua C. Hinckle, and John E. Eck**, "Is Problem-Oriented Policing Effective in Reducing Crime and Disorder?," *Criminology & Public Policy*, 2010, *9* (1), 139–172.

Wilson, James Q. and Barbara Boland, "The Effect of the Police on Crime," *Law and Society Review*, 1978, *12* (3), pp. 367–390.

Wilson, Jeremy M. and Clifford A. Grammich, *Police Recruitment and Retention in the Contemporary Urban Environment: A National Discussion of Personnel Experiences and Promising Practices from the Front Lines*, Santa Monica: RAND, 2009.

Witt, Robert, Alan Clarke, and Nigel Fielding, "Crime and Economic Activity: A Panel Data Approach," *British Journal of Criminology*, 1999, *39* (3), 391–400.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, Cambridge: MIT Press, 2002.

Zimring, Franklin E., *The City That Became Safe: New York's Lessons for Urban Crime and Its Control*, New York: Oxford University Press, 2011.

Figure 1. Private Precaution Reaction Functions: Two Person Case, Low and High Public Policing

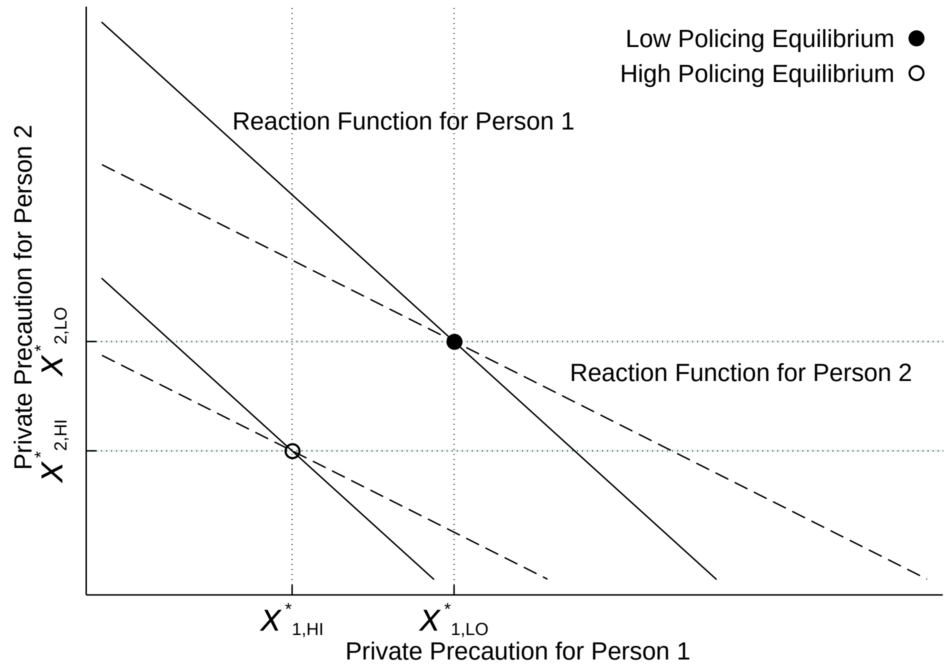
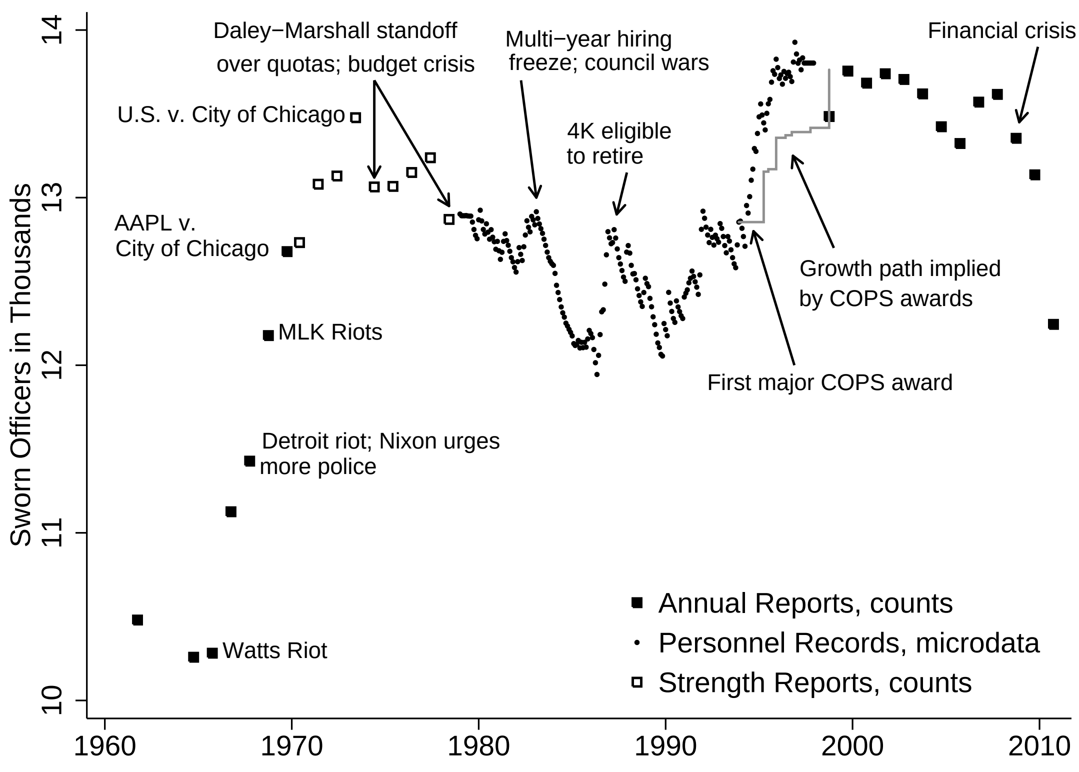


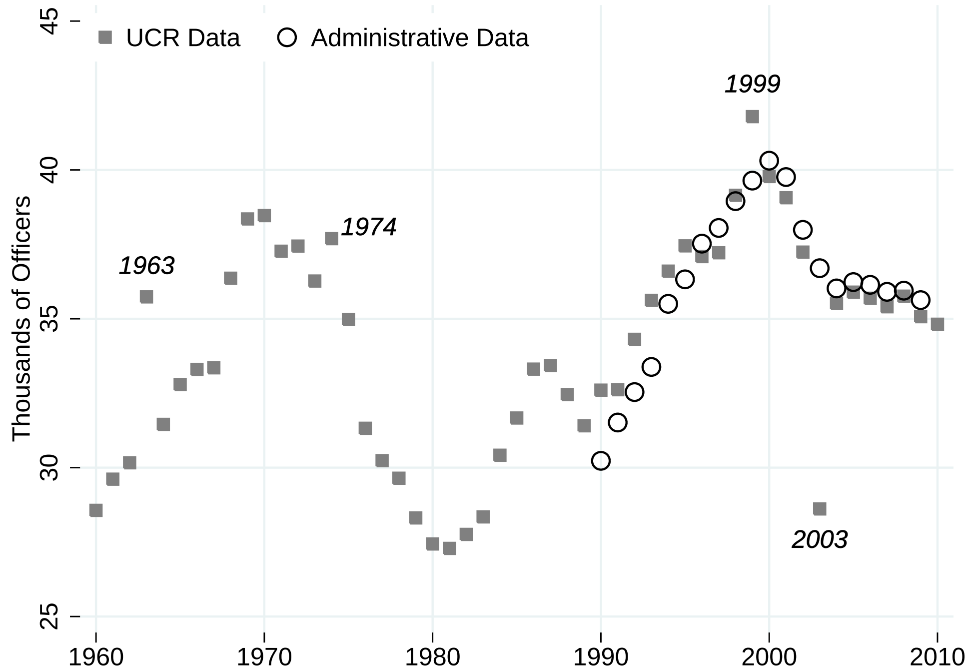
Figure 2. Police Hiring in Chicago, 1960-2010



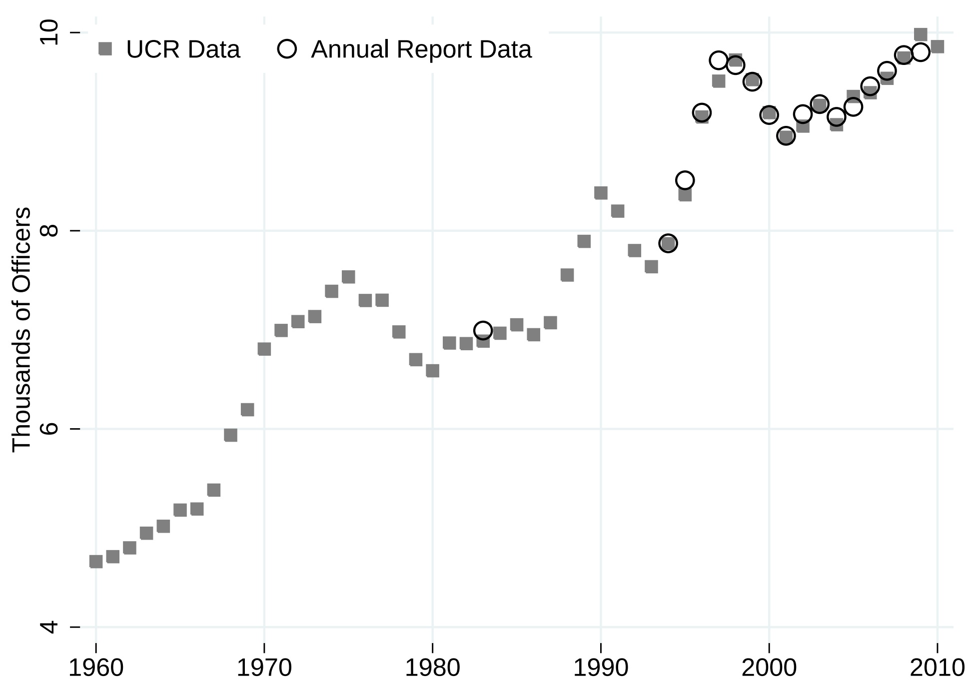
Note: See text for details.

Figure 3. Sworn Officers in Five Cities: the Uniform Crime Reports and Direct Measures from Departments

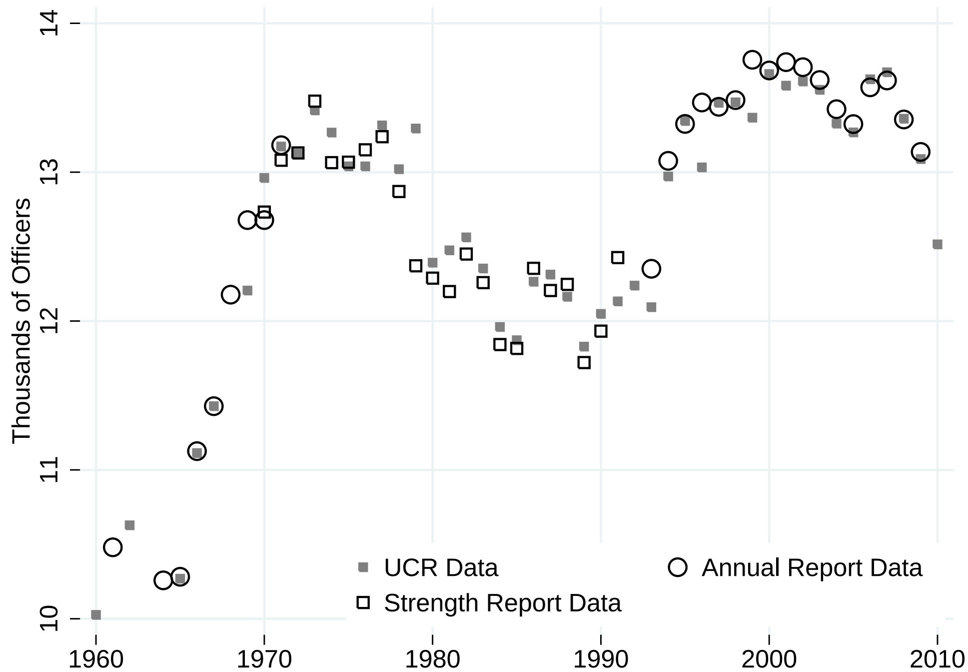
A. New York



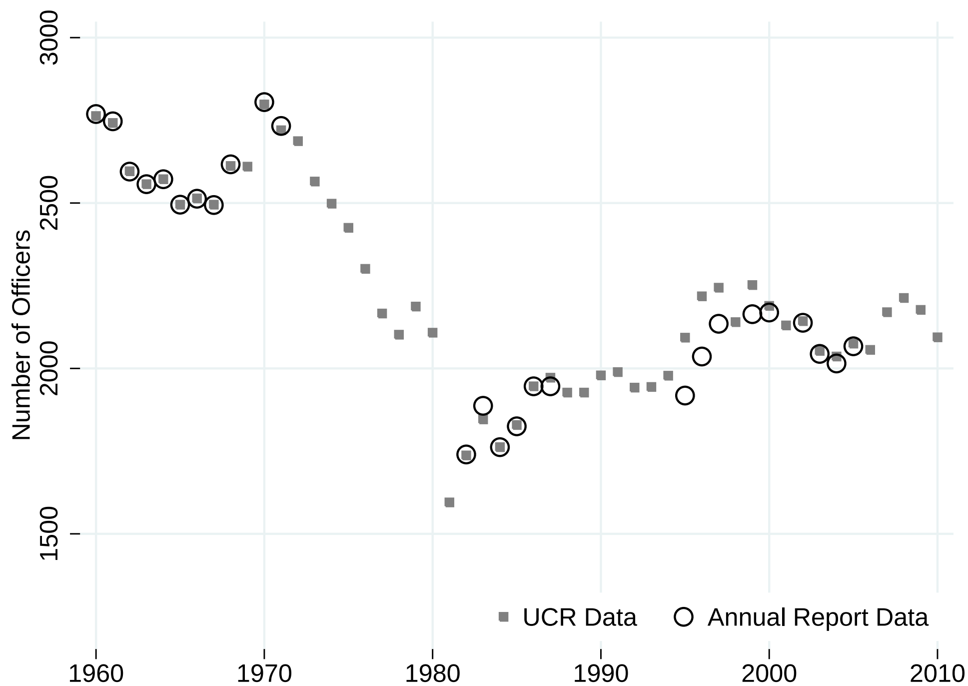
B. Los Angeles



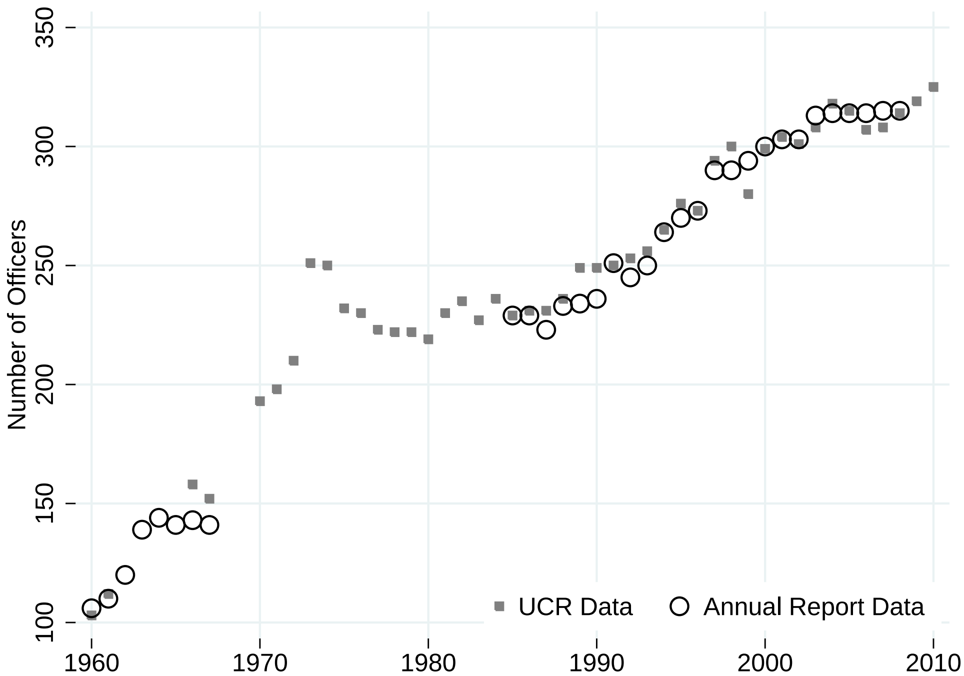
C. Chicago



D. Boston

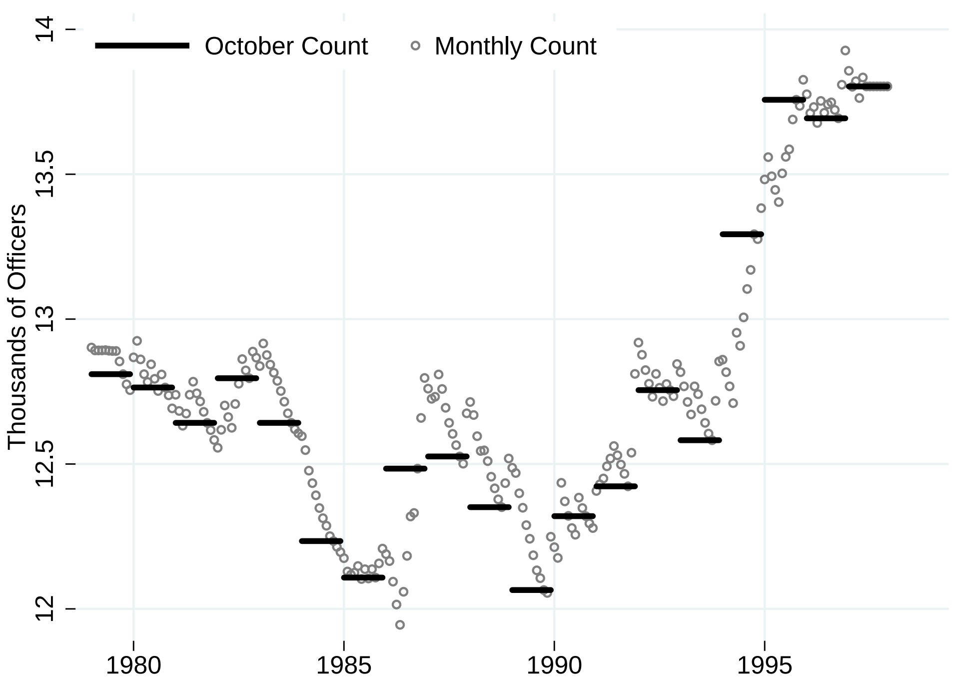


E. Lincoln, Nebraska



Note: In panel A, numbers for 1960-1994 are adjusted to account for the 1995 merger of NYPD with housing and transit police. See Data Appendix for details.

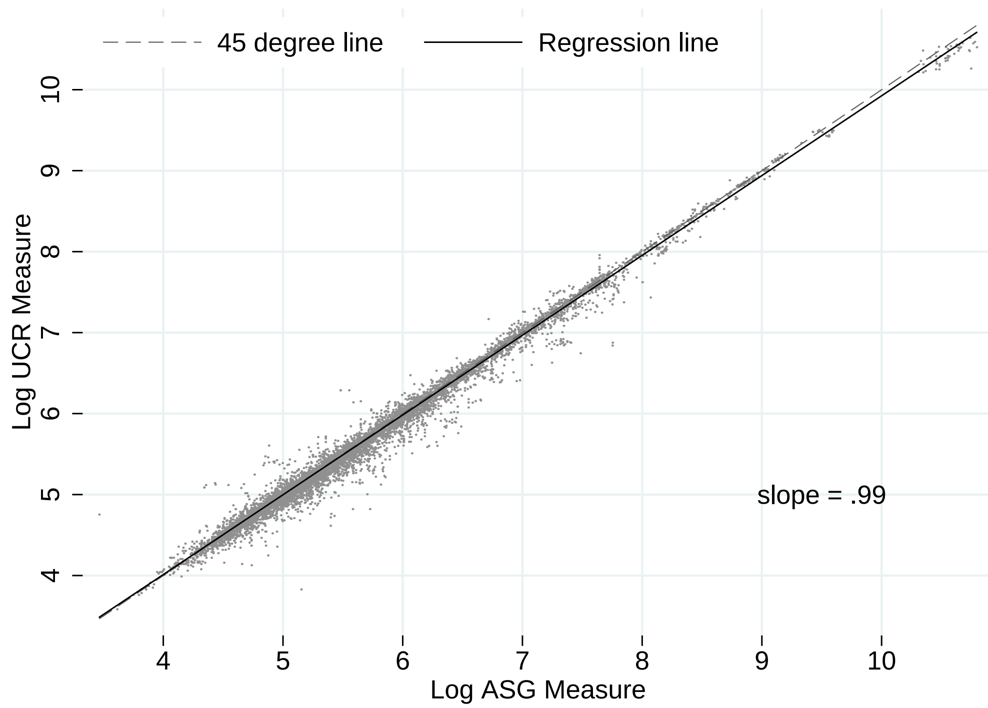
Figure 4. Sworn Officers in Chicago 1979-1997, by Month



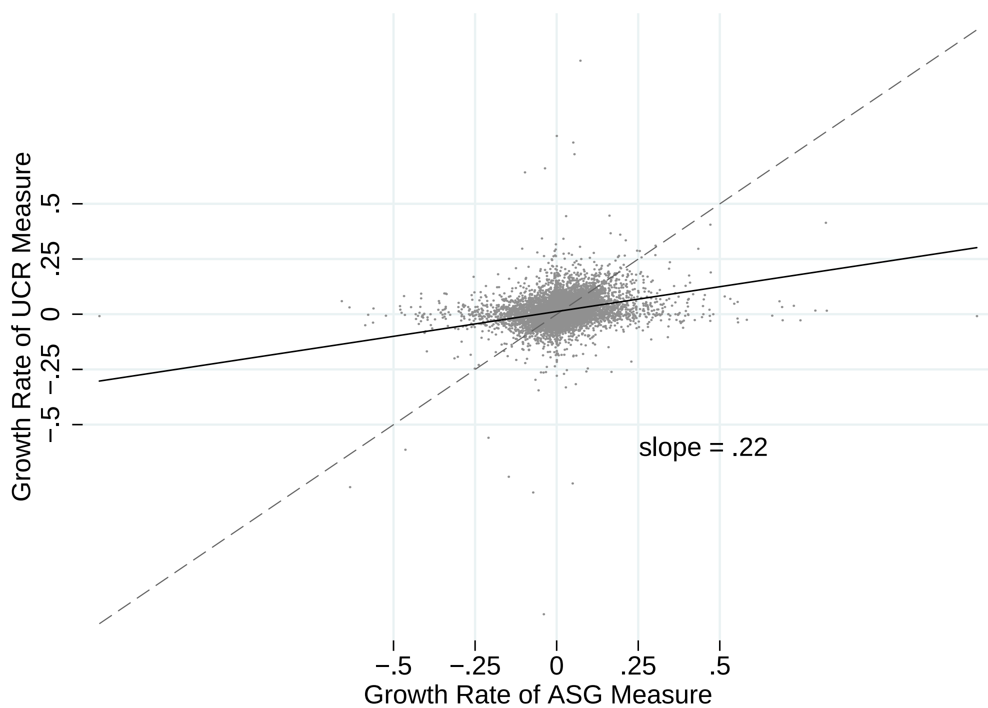
Note: See text for details.

Figure 5. Two Leading Measures of Sworn Officers: the Uniform Crime Reports and the Annual Survey of Government

A. Logs



B. Log Differences



Note: See text and Data Appendix for details.

Figure 6. Location of Cities in Sample

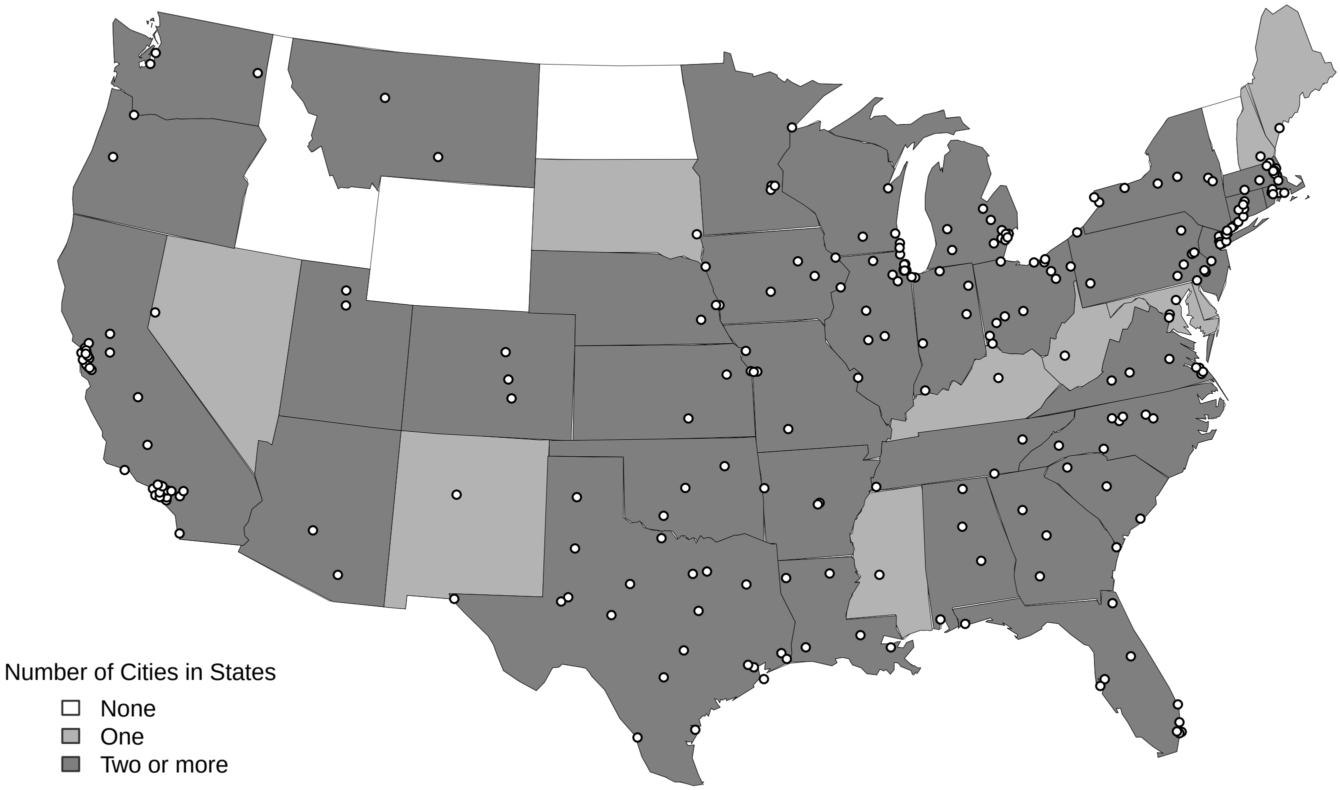
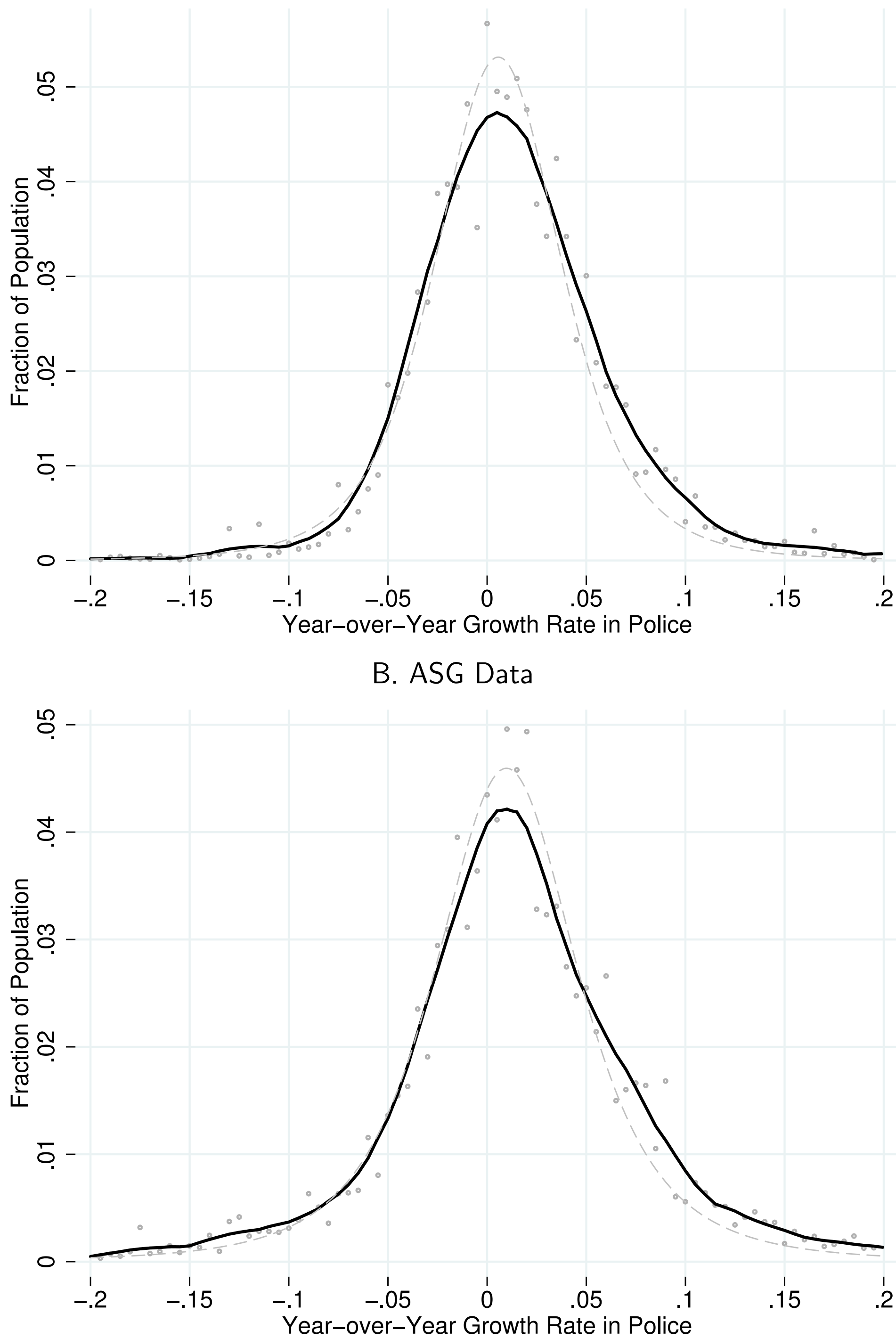


Figure 7. Distribution of Growth Rates in Police

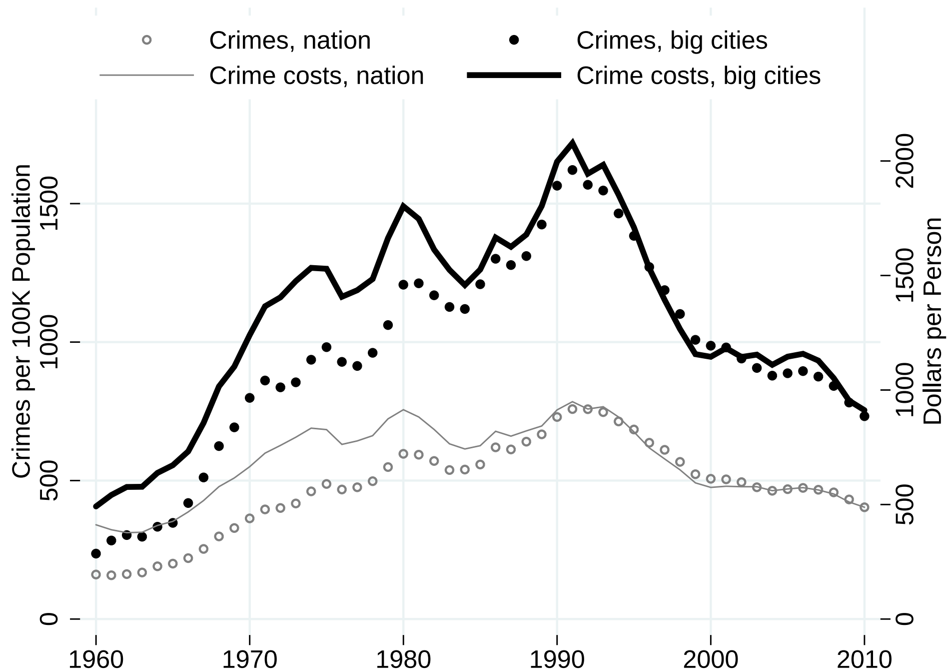
A. UCR Data



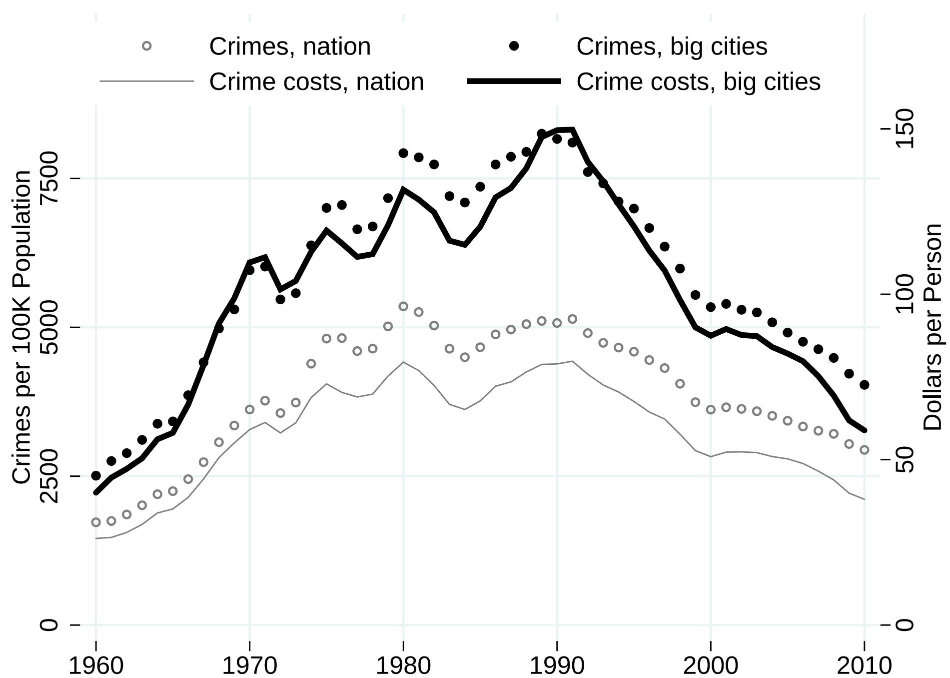
Note: Curves are proportional to the population weighted conditional density function of growth rates in police, conditional on not having been equal to zero. In the UCR (ASG) data, 3.8 (6.1) percent of person-weighted city-years have exactly zero growth rate. Gray circles are undersmoothed histogram heights. The gray dashed line is represents the normal density plot. The scale for the y-axis is percent of person-weighted city-years. See text for details of density estimation.

Figure 8. Aggregate Trends in Violent and Property Crime and Police: Evidence from the Uniform Crime Reports

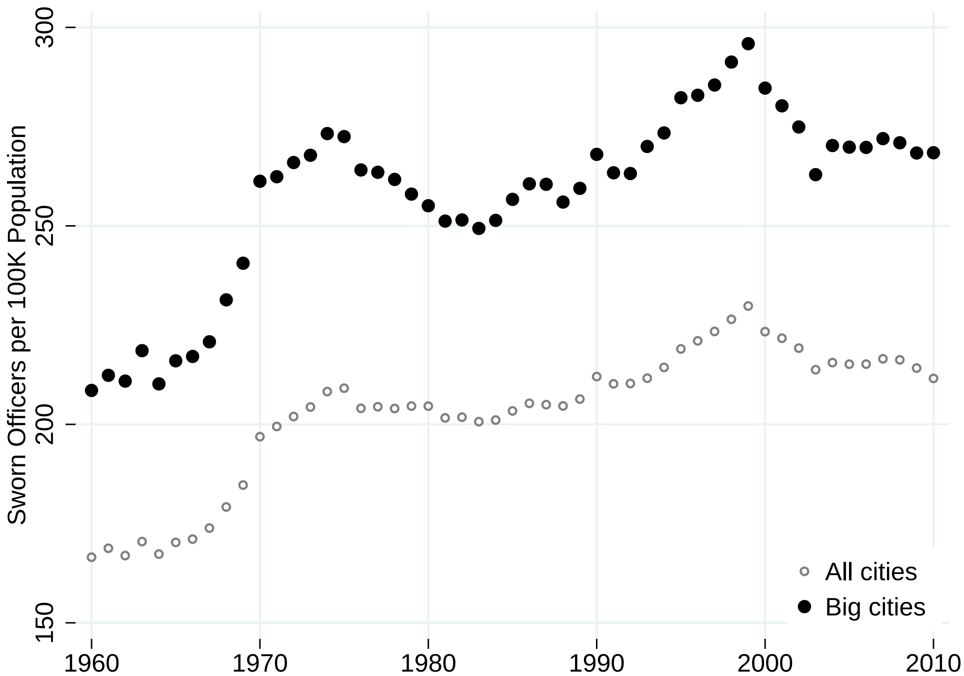
A. Violent Crime: Murder, Forcible Rape, Robbery, Aggravated Assault



B. Property Crime: Burglary, Larceny, Motor Vehicle Theft

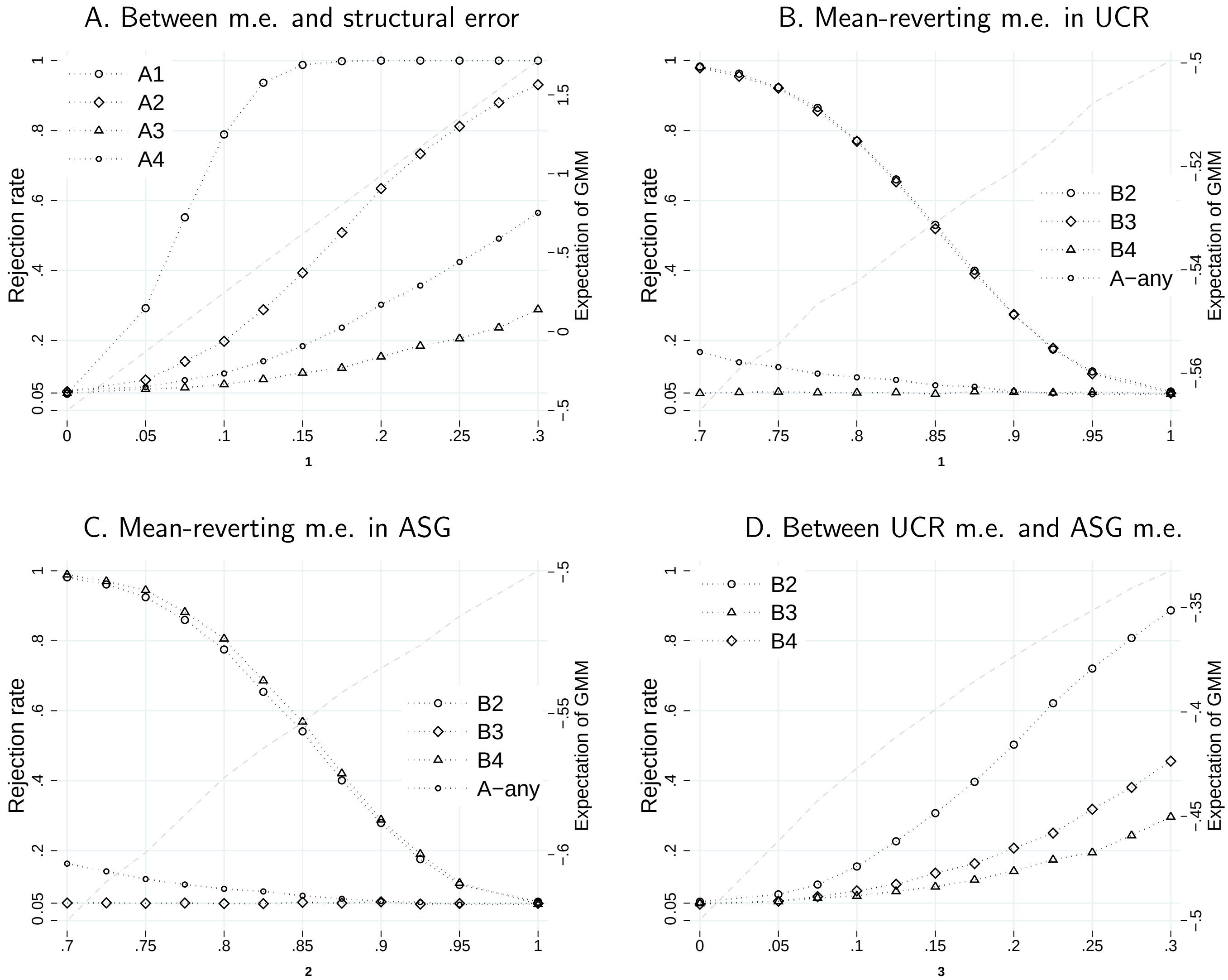


C. Sworn Police



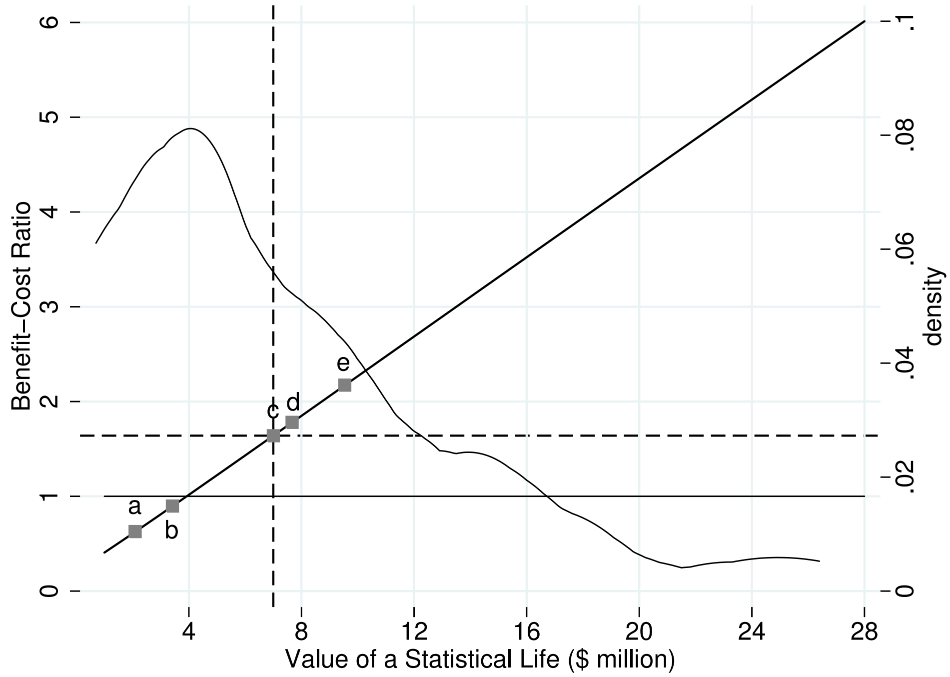
Note: In panels A and B, data on crimes nationally are taken from `http://www.ucrdatatool.gov`. In panel C, no such data are available, and we construct an index using all municipalities ever reporting to the UCR system 1960-2010 and imputation. See text and Data Appendix for details.

Figure 9. Power of Tests of Classical Measurement Error Model



Note: m.e. = measurement error. Each panel presents the fraction of 10,000 simulated data sets for which a t-ratio test rejects, as a function of a particular parameter. The parameter ρ_1 indexes the correlation between the measurement error and the structural error (panel A). The parameters λ_1 and λ_2 index the degree of mean reversion in the mesaurement errors (panels B, C). The parameter ρ_3 indexes the degree of correlation between the measurement errors themselves (panel D). Two types of t-ratio tests are presented, corresponding to the tests presented in the top two panels of Table 6. Test A is the t-ratio on the outcome in a regression of the difference in two measures of the variable of interest on the outcome and corresponds to Panel A of Table 6. Test B is the t-ratio on a third measure in a regression of the difference in measures on the third measure and corresponds to Panel B of Table 6. There are four such tests examined corresponding to the columns of Table 6. For example, "A1" corresponds to the t-ratio tests of Panel A of Table 6 for the first column, i.e., the full sample, whereas "B2" corresponds to the t-ratio tests of Panel B of Table 6 for the second column, i.e., the LEMAS subsample. The curves with open circles, diamonds, and triangles correspond to rejection rates for the given scenario. The curve labeled "A-any" is a rejection rate assuming "reject" occurs if any of the four Tests A reject at the 0.05/4 level. The dashed line with no symbols overlaid is the simulation estimate of the expectation of two-step GMM (right axis). The true parameter in all scenarios is -0.5, so departures from -0.5 capture estimator bias. See text for details.

Figure 10. Cost-Benefit Analysis
Benefit-Cost Ratio as a Function of the Value of a Statistical Life



Note: The table plots the value of the benefit-cost ratio calculated using the two-step GMM procedure that pools the "forward" and "reflected" IV regressions of the growth rate in each of nine crime rates on the first lag of the growth rate in the number of sworn police officers, conditional on both the UCR and the ASG measure of the growth rate in the population size and a vector of unrestricted state-by-year dummies. For each measure of police, expenditures on personnel are estimated by multiplying the number of personnel by $130,000, an estimate of the "fully-loaded" annual salary of a police officer. Victimization costs for rape, robbery, assault, burglary, larceny and motor vehicle theft are drawn from Cohen and Piquero (2008). As there is a great deal of variation in extant estimates of the value of a statistical life, the cost of murder is allowed to vary. Using the solid black line, we plot the benefit-cost ratio on the vertical axis as a function of the value of a statistical life, plotted on the horizontal axis in millions of dollars. The horizontal line correpsonds to a benefit-cost ratio of 1. In addition, we superimpose a kernel density function that plots the distribution of the extant estimates of the value of a statistical life. Key estimates include the $2.1 million VSL estimated by Ashenfelter and Greenstone (2004) ("a"), $3.4, the mean VSL among studies of non-labor market behavior ("b"), $7 million, the mean VSL in our sample ("c"), $7.7 million, the mean VSL used by various federal agencies for the 2004-2010 period ("d") and $9.5 million, the mean VSL among studies of U.S. labor market behavior ("e"). The dotted lines show the BCR ($1.63) at the mean value of a statistical life ($7 million). The majority of these estimates are drawn from Viscusi and Aldy (2003). We supplement these estimates with several that are drawn from the more recent literaure.

Table 1. Elasticity of Police with Respect to Selected Covariates

Variable	Source	Years	UCR measure	ASG measure	UCR measure	ASG measure
A. Economic Characteristics						
Personal income	Bureau of Economic Analysis (county-level)	1969-2010	0.053	0.055	-0.002	-0.052
			(0.040)	(0.046)	(0.034)	(0.050)
Total employment			0.097	0.079	0.030	-0.042
			(0.047)	(0.065)	(0.043)	(0.064)
Adjusted gross income	Internal Revenue Service (county-level)	1990-2009	0.012	-0.012	-0.018	-0.039
			(0.035)	(0.034)	(0.036)	(0.050)
Wage and salary income			0.044	0.037	-0.009	-0.047
			(0.042)	(0.042)	(0.043)	(0.058)
Municipal expenditures exclusive of police	Annual Survey of Gov't Finances (city-level)	1960-2010	0.019	0.016	0.011	0.018
			(0.007)	(0.008)	(0.006)	(0.008)
B. Social Disorganization and Demographics						
Share of births to teenage mothers, overall	Centers for Disease Control (county-level)	1968-2002	-0.019	-0.033	-0.022	-0.026
			(0.011)	(0.015)	(0.012)	(0.015)
Share of births to teenage mothers, black births only			-0.012	-0.038	-0.018	-0.029
			(0.011)	(0.015)	(0.012)	(0.015)
Share of births that are low birthweight			0.012	0.004	0.011	-0.007
			(0.012)	(0.016)	(0.011)	(0.017)
12th grade dropout rate	Nat'l Center for Educational Statistics (city-level)	1986-2008	0.006	0.005	-0.012	0.006
			(0.015)	(0.012)	(0.009)	(0.015)
Estimated population at risk of arrest	U.S. Census (county-level)	1970-2010	0.113	0.050	0.124	0.096
			(0.026)	(0.042)	(0.038)	(0.051)
C. Lagged Crimes						
Violent crimes	Federal Bureau of Investigation, Uniform Crime Reports (city-level)	1960-2010	0.012	0.006	0.009	0.004
			(0.005)	(0.006)	(0.005)	(0.006)
Property crimes			0.019	0.012	0.013	0.008
			(0.011)	(0.011)	(0.009)	(0.012)
Cost-weighted crimes			-0.001	0.002	0.001	-0.000
			(0.003)	(0.003)	(0.002)	(0.003)
year effects			yes	yes	—	—
state-by-year effects			no	no	yes	yes

Note: Each column reports coefficients from 13 separate regressions of the growth rate in a measure of the number of police officers on the the growth rate in a covariate, controlling for two measures of the growth rate in city population and either year effects or state-by-year effects. The coefficient reported corresponds to the variable given in the first column, and Huber-Eicker-White standard errors are given in parentheses. The first police measure is from the Uniform Crime Reports and the second is from the Annual Survey of Government Employment. All regressions are weighted by 2010 city population. The municipal budget cycle, the 12th grade dropout rate, and the lagged crime rates are all measured at the city level, but city level data was unavailable for other variables, which were measured instead at the county level. See Data Appendix for details. Population weighted means (standard deviations) for variables in Panel A are 0.068 (0.042), 0.015 (0.029), 0.040 (0.050), 0.041 (0.038), and 0.030 (0.119); for those in Panel B are -0.014 (0.083), -0.016 (0.086), -0.002 (0.065), 0.007 (0.111), and 0.011 (0.038); and for those in Panel C are 0.037 (0.160), 0.017 (0.110), and 0.023 (0.270).

Table 2. Summary Statistics on Police and Crime

Variable		Levels				Log Differences			
		Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Sworn police, UCR	O	245.5	111.6	54.4	786.6	0.016	0.058	-1.359	1.148
(per 100K pop)	B		105.7				0.012		
	W		36.0				0.056		
Sworn police, ASG	O	257.7	128.0	50.1	779.8	0.016	0.078	-1.401	1.288
(per 100K pop)	B		120.4				0.012		
	W		42.4				0.078		
Violent crimes	O	972.7	630.5	8.2	4189.0	0.035	0.162	-1.804	1.467
(per 100K pop)	B		440.3				0.019		
	W		451.4				0.161		
Murder	O	14.6	10.6	0.0	110.9	0.014	0.382	-2.792	2.446
(per 100K pop)	B		8.4				0.021		
	W		6.5				0.382		
Rape	O	49.0	29.6	0.0	310.5	0.035	0.291	-4.384	4.199
(per 100K pop)	B		17.4				0.028		
	W		23.9				0.289		
Robbery	O	438.0	344.5	1.1	2,358.0	0.035	0.202	-1.792	1.946
(per 100K pop)	B		257.5				0.019		
	W		228.9				0.201		
Assault	O	471.1	329.5	1.2	2,761.4	0.037	0.213	-2.833	3.129
(per 100K pop)	B		209.5				0.024		
	W		254.4				0.212		
Property crimes	O	6,223.4	2,355.0	667.3	18,345.2	0.015	0.113	-1.304	1.248
(per 100K pop)	B		1,366.2				0.014		
	W		1,918.2				0.112		
Burglary	O	1,671.9	810.9	143.0	6,713.5	0.010	0.149	-1.549	1.411
(per 100K pop)	B		433.8				0.018		
	W		685.1				0.148		
Larceny	O	3,655.4	1,500.2	84.2	11,590.7	0.017	0.122	-1.435	2.146
(per 100K pop)	B		982.6				0.015		
	W		1,133.7				0.121		
Motor vehicle	O	896.0	574.5	42.5	5,294.7	0.014	0.169	-1.516	1.447
theft	B		428.6				0.016		
(per 100K pop)	W		435.3				0.169		
Cost-Weighted	O	1,433.9	904.9	15.36	8,909.2	0.019	0.271	-2.363	3.033
Crimes	B		699.6				0.018		
($ per capita)	W		573.9				0.270		

Note: This table reports descriptive statistics for the two measures of sworn police officers used throughout the article as well as for each of the seven crime categories and three crime aggregates. For each variable, we report the overall mean, the standard deviation decomposed into overall ("O"), between ("B"), and within ("W") variation, as well as the minimum and maximum values. Summary statistics are reported both in levels per 100,000 population and in growth rates. All statistics are weighted by 2010 city population. The sample size for all variables is N=10,589.

Table 3. First Stage Estimates

	Forward Models		Reflected Models	
Endogenous Regressor	(1)	(2)	(3)	(4)
UCR measure	0.184	0.161		
	(0.014)	(0.013)		
ASG measure			0.364	0.356
			(0.029)	(0.029)
F-statistic	169.1	144.7	154.2	146.4
Instrument:	ASG		UCR	
year effects	yes	—	yes	—
state-year effects	no	yes	no	yes

Each column reports results of a least squares regression of the growth rate in a given measurement of the number of police officers on the the growth rate in the other measurement of police. Columns (1) and (2) report results for the forward regressions in which the UCR measure is employed as the endogenous covariate and the ASG measure is employed as the instrumental variable while columns (3) and (4) report results for the reflected regressions in which the ASG measure is employed as the endogenous covariate and the UCR measure is employed as the instrumental variable. For each set of models, the first column reports regression results, conditional on both the UCR and the ASG measures of the growth rate in the city's population and a vector of year dummies. The second column adds a vector of state-by-year effects. All models are estimated using 2010 city population weights. Huber-Eicker-White standard errors are reported in parentheses below the coefficient estimates.

Table 4. Estimates of the Effect of Police on Crime

	Least Squares Estimates				2SLS Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	UCR Measure		ASG Measure		*Forward Models* UCR Measure		*Reflected Models* ASG Measure	
Violent crimes	-0.117	-0.120	-0.053	-0.058	-0.289	-0.361	-0.321	-0.336
	(0.037)	(0.040)	(0.024)	(0.023)	(0.128)	(0.143)	(0.100)	(0.106)
Murder	-0.270	-0.204	-0.148	-0.143	-0.804	-0.889	-0.742	-0.572
	(0.071)	(0.097)	(0.047)	(0.059)	(0.260)	(0.364)	(0.198)	(0.262)
Rape	-0.066	-0.074	-0.038	-0.054	-0.208	-0.339	-0.181	-0.208
	(0.069)	(0.092)	(0.043)	(0.050)	(0.234)	(0.301)	(0.188)	(0.248)
Robbery	-0.180	-0.204	-0.085	-0.084	-0.459	-0.521	-0.493	-0.572
	(0.048)	(0.047)	(0.032)	(0.029)	(0.176)	(0.177)	(0.128)	(0.125)
Assault	-0.052	-0.037	-0.010	-0.013	-0.052	-0.079	-0.143	-0.104
	(0.044)	(0.050)	(0.030)	(0.035)	(0.164)	(0.209)	(0.120)	(0.136)
Property crimes	-0.071	-0.059	-0.028	-0.030	-0.152	-0.189	-0.195	-0.167
	(0.028)	(0.026)	(0.020)	(0.015)	(0.109)	(0.090)	(0.077)	(0.068)
Burglary	-0.061	-0.062	-0.041	-0.054	-0.222	-0.339	-0.166	-0.174
	(0.043)	(0.037)	(0.027)	(0.021)	(0.144)	(0.128)	(0.118)	(0.098)
Larceny	-0.038	-0.025	-0.002	-0.018	-0.012	-0.113	-0.103	-0.070
	(0.031)	(0.027)	(0.021)	(0.017)	(0.115)	(0.103)	(0.085)	(0.074)
Motor vehicle theft	-0.187	-0.131	-0.109	-0.047	-0.592	-0.292	-0.514	-0.367
	(0.049)	(0.043)	(0.031)	(0.025)	(0.169)	(0.151)	(0.130)	(0.115)
Cost-Weighted crime	-0.213	-0.144	-0.112	-0.099	-0.605	-0.614	-0.583	-0.403
	(0.054)	(0.071)	(0.034)	(0.041)	(0.184)	(0.250)	(0.147)	(0.192)
year effects	yes	—	yes	—	yes	—	yes	—
state-year effects	no	yes	no	yes	no	yes	no	yes
Instrument:	—	—	—	—	ASG		UCR	

Note: Columns (1)-(4) reports results of a least squares regression of the growth rate in each of ten crime rates on the first lag of the growth rate in the number of sworn police officers. For each set of models, the first column reports regression results, conditional on both the UCR and the ASG measures of the growth rate in the city's population and year effects. The second column adds state-by-year effects. Columns (5)-(8) reports results from a series of 2SLS regressions of the growth rate in each of nine crime rates on the first lag of the growth rate in the number of per capita sworn police officers. All models are estimated using 2010 city population weights. Huber-Eicker-White standard errors that are robust to heteroskedasticity are reported in the second row below the coefficient estimates.

Table 5. Pooled Estimates of the Effect of Police on Crime: Within-State Differences

	Violent Crimes				Property Crimes			Aggregates		
Estimator	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft	Violent Crime	Property Crime	Cost-Weighted Crimes
One-step GMM	-0.727 (0.242)	-0.272 (0.231)	-0.547 (0.123)	-0.092 (0.142)	-0.255 (0.093)	-0.091 (0.072)	-0.331 (0.105)	-0.348 (0.103)	-0.178 (0.066)	-0.506 (0.173)
Two-step GMM	-0.666 (0.238)	-0.255 (0.219)	-0.559 (0.117)	-0.099 (0.127)	-0.225 (0.089)	-0.083 (0.067)	-0.343 (0.101)	-0.344 (0.096)	-0.174 (0.062)	-0.473 (0.171)
Empirical Likelihood	-0.667 (0.236)	-0.256 (0.221)	-0.559 (0.117)	-0.099 (0.127)	-0.221 (0.087)	-0.082 (0.067)	-0.341 (0.103)	-0.344 (0.096)	-0.173 (0.061)	-0.473 (0.170)
Test statistic:	0.78	0.19	0.09	0.02	1.86	0.18	0.24	0.03	0.06	0.71

Note: Each column reports generalized method of moments (GMM) or empirical likelihood (EL) estimates of the growth rate in each of ten crime rates on the first lag of the growth rate in the number of sworn police officers, conditional on both the UCR and ASG measures of the growth rate in population and a vector of unrestricted state-by-year effects. Computationally, however, we de-mean the data by state-year rather than computing the fixed effects parameters, even when this is an approximation (e.g., EL). The one-step GMM estimator uses the identity weighting matrix while the two-step GMM estimator uses a weighting matrix based on the updated variance estimator from the one-step procedure. Following Guggenberger and Hahn (2005), we interpret EL as a particular just-identified GMM estimator and compute it using 5 Newton iterations using the two-step GMM estimates as starting values. All models use 2010 city population weights, and Huber-Eicker-White standard errors are reported in parentheses. Below the parameter estimates and the standard errors, we report the value of of the overidentification statistic from the two-step GMM procedure. The test statistic corresponds to the pooling restriction that we estimate a common parameter on the growth rate in police and refers to a test of the equality of the "forward" and "reflected" IV coefficients. The test statistic is distributed χ_1 under the null hypothesis of classical measurement error. The 95 percent critical value of the test is 3.84.

Table 6. Further Tests of Classical Measurement Errors

	Full Sample	LEMAS Subsample		
	UCR-ASG (1)	UCR-ASG (2)	UCR-LEMAS (3)	ASG-LEMAS (4)
A. Growth Rate in Crimes				
Murder	-0.002	-0.009	0.003	0.010
	(0.002)	(0.007)	(0.005)	(0.007)
Rape	0.001	-0.009	-0.002	0.008
	(0.004)	(0.016)	(0.010)	(0.014)
Robbery	-0.004	-0.021	0.009	0.026
	(0.006)	(0.020)	(0.016)	(0.022)
Assault	-0.002	0.004	-0.010	-0.014
	(0.005)	(0.018)	(0.013)	(0.018)
Burglary	0.014	-0.001	-0.013	-0.011
	(0.010)	(0.027)	(0.022)	(0.028)
Larceny	0.002	-0.012	0.006	0.020
	(0.012)	(0.035)	(0.023)	(0.033)
Motor vehicle theft	-0.007	-0.021	0.025	0.040
	(0.008)	(0.023)	(0.017)	(0.023)
B. Growth Rate in Police				
LEMAS police measure		-0.083		
		(0.049)		
ASG police measure			-0.077	
			(0.026)	
UCR police measure				-0.082
				(0.049)
C. Growth Rate in Population				
UCR population measure	-0.055	0.571	0.364	-0.124
	(0.124)	(0.308)	(0.213)	(0.201)
ASG population measure	-0.035	-0.406	-0.200	0.172
	(0.108)	(0.348)	(0.236)	(0.236)
D. Joint Tests of Significance				
F-test: all variables	0.83	0.25	0.17	0.07
F-test: crime variables	0.82	0.39	0.67	0.08
F-test: police variable		0.09	0.00	0.09
F-test: population variables	0.11	0.39	0.11	0.76

Note: Each column corresponds to a particular difference between measures of the growth rate in police. The heading for each column gives the sources of the two measures being differenced. The LEMAS data are only available for years 1987, 1990, 1992, 1993, 1996, 1997, 1999, 2000, 2003, 2004, 2007 and 2008, and so columns involving the LEMAS data correspond to a limited subsample. Each column reports coefficients and heteroskedasticity-robust standard errors (parentheses) from a single regression of the difference in measures reported in the column heading on the variables listed in the first column. The coefficients are grouped substantively into panels A, B, and C. Panel D gives p-values from a series of heteroskedasticity-robust F-tests on the joint significance of each set of variables. Each of the models controls for state-by-year effects and is weighted by 2010 city population.

Table 7. Robustness of Results to the Inclusion of Covariates
1970-2002 Sample

	"Forward" Models						"Reflected" Models					
	Endogenous Covariate: UCR Instrument: ASG						Endogenous Covariate: ASG Instrument: UCR					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Violent crimes	-0.286	-0.291	-0.319	-0.293	-0.316	-0.332	-0.238	-0.233	-0.174	-0.159	-0.159	-0.173
	(0.152)	(0.154)	(0.160)	(0.161)	(0.166)	(0.168)	(0.114)	(0.115)	(0.114)	(0.115)	(0.117)	(0.117)
Murder	-1.068	-1.029	-1.086	-1.058	-1.044	-1.079	-0.523	-0.467	-0.361	-0.336	-0.342	-0.366
	(0.392)	(0.394)	(0.411)	(0.411)	(0.419)	(0.424)	(0.298)	(0.301)	(0.296)	(0.297)	(0.304)	(0.305)
Rape	-0.165	-0.144	-0.264	-0.235	-0.180	-0.193	0.121	0.127	0.156	0.176	0.187	0.168
	(0.258)	(0.259)	(0.266)	(0.269)	(0.274)	(0.277)	(0.210)	(0.212)	(0.211)	(0.214)	(0.219)	(0.221)
Robbery	-0.550	-0.556	-0.547	-0.554	-0.567	-0.578	-0.686	-0.686	-0.602	-0.601	-0.594	-0.603
	(0.185)	(0.186)	(0.190)	(0.193)	(0.197)	(0.200)	(0.135)	(0.136)	(0.134)	(0.135)	(0.138)	(0.139)
Assault	0.095	0.080	0.032	0.082	0.040	0.022	0.138	0.139	0.189	0.215	0.205	0.197
	(0.205)	(0.207)	(0.217)	(0.219)	(0.224)	(0.228)	(0.144)	(0.146)	(0.148)	(0.150)	(0.150)	(0.151)
Property crimes	-0.262	-0.281	-0.295	-0.285	-0.298	-0.314	-0.152	-0.148	-0.122	-0.117	-0.119	-0.124
	(0.102)	(0.102)	(0.107)	(0.107)	(0.108)	(0.108)	(0.086)	(0.087)	(0.088)	(0.089)	(0.090)	(0.090)
Burglary	-0.408	-0.414	-0.450	-0.437	-0.440	-0.461	-0.118	-0.114	-0.099	-0.095	-0.078	-0.081
	(0.146)	(0.147)	(0.156)	(0.157)	(0.161)	(0.160)	(0.117)	(0.119)	(0.123)	(0.125)	(0.127)	(0.126)
Larceny	-0.174	-0.204	-0.224	-0.222	-0.249	-0.261	-0.074	-0.071	-0.040	-0.035	-0.047	-0.049
	(0.121)	(0.121)	(0.126)	(0.126)	(0.127)	(0.128)	(0.093)	(0.093)	(0.095)	(0.095)	(0.097)	(0.095)
Motor vehicle	-0.357	-0.337	-0.328	-0.303	-0.301	-0.320	-0.385	-0.369	-0.348	-0.341	-0.351	-0.356
	(0.160)	(0.160)	(0.164)	(0.163)	(0.166)	(0.168)	(0.147)	(0.147)	(0.147)	(0.147)	(0.151)	(0.155)
Cost-Weighted Crimes	-0.786	-0.763	-0.816	-0.783	-0.790	-0.815	-0.345	-0.312	-0.208	-0.183	-0.187	-0.206
	(0.257)	(0.259)	(0.267)	(0.267)	(0.271)	(0.274)	(0.207)	(0.209)	(0.200)	(0.201)	(0.206)	(0.208)
state-by-year effects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
economic covariates	no	yes	yes	yes	yes	yes	no	yes	yes	yes	yes	yes
lagged economic covariates	no	no	yes	yes	yes	yes	no	no	yes	yes	yes	yes
demographic variables	no	no	no	yes	yes	yes	no	no	no	yes	yes	yes
polynomials and interactions	no	no	no	no	yes	yes	no	no	no	no	yes	yes
linear time trends	no	no	no	no	no	yes	no	no	no	no	no	yes

Note: Each column reports results of a 2SLS regression of the growth rate in each of ten crime rates on the first lag of the growth rate in the number of sworn police officers. Columns (1)-(6) report results for the "forward" regressions in which the UCR measure is employed as the endogenous covariate and the ASG measure is employed as the instrumental variable while columns (7)-(12) report results for the "reflected" regressions in which the ASG measure is employed as the endogenous covariate and the UCR measure is employed as the instrumental variable. For each set of models, the first column reports regression results, conditional on both the UCR and the ASG measures of the growth rate in the city's population and a vector of unrestricted state-by-year dummies. The second column adds a vector of economic covariates while the third column adds the first lag of each of these covariates. In the fourth column, we add demographic controls which capture the proportion of a city's population that is comprised of each of sixteen age-gender-race groups. In the fifth column, we add polynomial terms and interactions of the demographic variables. Finally, in column (6), we add city-specific linear time trends. All models are estimated using 2010 city population weights. Huber-Eicker-White standard errors that are robust to heteroskedasticity are reported below the coefficient estimates.

Table 8. Comparison of Estimates of the Police Elasticity of Crime

Source	Years	Cross-Sectional Units	Research Design	Violent Crime	Murder	Robbery	Property Crime	Burglary	Motor Vehicle Theft
Marvell and Moody (1996)	1973-1992	56 cities	lags as control variables		-0.13	-0.22		-0.15	-0.30
					(0.05)	(0.06)		(0.04)	(0.07)
Levitt (1997)	1970-1992	59 cities	mayoral elections	-0.79	-3.03	-1.29	0.00	-0.55	-0.44
				(0.61)	(2.03)	(1.00)	(0.34)	(0.67)	(0.98)
McCrary (2002)	1970-1992	59 cities	mayoral elections	-0.66	-2.69	-0.98	0.11	-0.47	-0.77
				(0.65)	(2.07)	(1.09)	(0.43)	(0.77)	(1.08)
Levitt (2002)	1975-1995	122 cities	number of firefighters	-0.44	-0.91	-0.45	-0.50	-0.20	-1.70
				(0.23)	(0.33)	(0.26)	(0.24)	(0.26)	(0.57)
Corman and Mocan (2005)	1974-1999	NYC	monthly time series		-0.50	-0.39		-0.28	-0.58
					(1.04)	(0.45)		(0.23)	(0.25)
Klick and Tabarrok (2005)	2002-2003	DC	high terrorism alert days	0.00				-0.30	-0.84
				(na)				(0.18)	(0.25)
Evans and Owens (2007)	1990-2001	2,074 cities	COPS grants	-0.99	-0.84	-1.34	-0.26	-0.59	-0.85
				(0.33)	(0.47)	(0.52)	(0.16)	(0.18)	(0.35)
Lin (2009)	1970-2000	51 states	state sales tax	-1.13	-2.73	-1.86	-2.18	-1.59	-4.14
				(0.74)	(1.31)	(1.12)	(0.93)	(0.80)	(1.82)
Our preferred estimates	1960-2010	242 cities	measurement error correction	-0.34	-0.67	-0.56	-0.17	-0.23	-0.34
				(0.10)	(0.24)	(0.12)	(0.06)	(0.09)	(0.10)

Note: Table reports implied elasticities and standard errors from recent articles employing a novel identification strategy to estimate the effect of police on crime. In place of the original elasticities reported in Levitt (1997), we have included elasticity estimates from McCrary (2002) which correct for a coding error in the original paper; McCrary (2002) estimates refer to the estimates using a different measure of mayoral elections. Our preferred estimates which account for the presence of measurement errors are shown below.

Table 9. Tests of the Equality of Cross-Crime Elasticities

Type	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft	Violent Crimes	Property Crimes
Murder	-	.213	.649	.036	.058	.015	.181	-	.035
Rape	-	-	.181	.485	.917	.452	.689	-	.731
Robbery	-	-	-	.002	.008	.001	.120	-	.001
Assault	-	-	-	-	.382	.922	.114	-	.554
Burglary	-	-	-	-	-	.109	.287	.295	-
Larceny	-	-	-	-	-	-	.010	.010	-
Motor vehicle theft	-	-	-	-	-	-	-	.997	-
Violent crimes	-	-	-	-	-	-	-	-	.075

Note: Each element of the table reports a p-value for a test of the equality between the two-step GMM parameters reported in Table 5 for an exhaustive combination of any two crime categories. For example, the p-value arising from a test of the equality of the pooled murder and burglary elasticities is 0.058. The p-values are generated using a GMM procedure in which we stack data pertaining to each of the two crime categories. All models are estimated using 2010 city population weights and condition on two measures of population as well as an unrestricted vector of state-by-year effects.

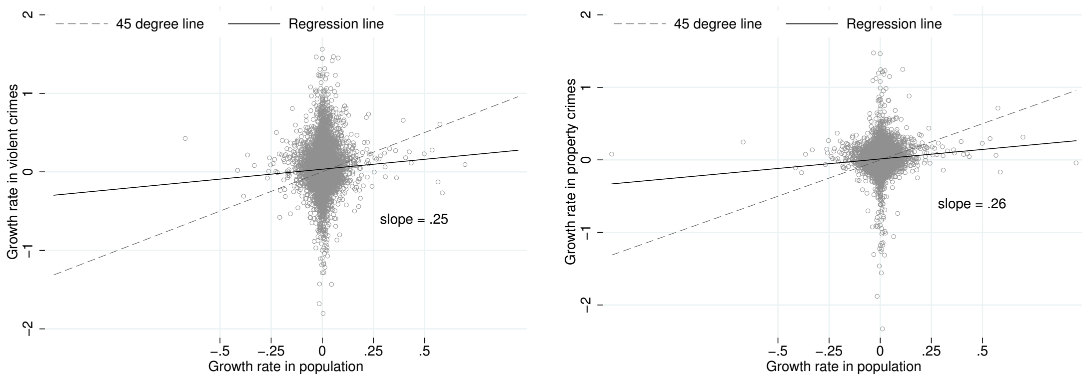
Table 10. Heterogeneity Across Cities in Relative Costs and Benefits as of 2010

Rank	City	Pop.	Poverty Rate	Income per capita	Cost of Crime per capita	Fraction Income at Risk	Officers per 100K Pop.	Est'd Cost per Officer	Benefit Cost Ratio
A. Most Overpoliced Cities									
1	Sunnyvale, CA	140,081	6.2	$43,828	$169	0.004	168.1	$280,434	0.2
2	Waltham, MA	60,632	11.5	$32,014	$123	0.004	246.8	$110,926	0.2
3	Torrance, CA	145,438	6.3	$36,007	$184	0.005	161.5	$204,302	0.3
4	Palo Alto, CA	64,403	5.7	$70,242	$211	0.003	148.6	$235,566	0.3
5	Bayonne, NJ	63,024	12.3	$28,698	$242	0.008	369.2	$107,971	0.3
6	Cambridge, MA	105,162	15.0	$44,717	$284	0.006	260.9	$157,353	0.3
7	Santa Monica, CA	89,736	11.1	$58,399	$434	0.007	230.4	$257,434	0.3
8	Stamford, CT	122,643	11.1	$44,667	$271	0.006	239.9	$139,771	0.4
9	Warwick, RI	82,672	7.6	$30,422	$194	0.006	203.3	$113,538	0.4
10	Burbank, CA	103,340	8.2	$33,320	$271	0.008	148.9	$212,202	0.4
11	Ann Arbor, MI	113,934	20.2	$30,498	$172	0.006	129.3	$151,331	0.4
12	Cranston, RI	80,387	8.4	$27,752	$183	0.007	180.0	$108,909	0.4
13	Alameda, CA	73,812	10.1	$38,434	$299	0.008	134.9	$233,762	0.4
14	Royal Oak, MI	57,236	6.8	$37,095	$185	0.005	143.6	$129,915	0.5
15	Alexandria, VA	139,966	7.8	$54,345	$335	0.006	225.6	$149,329	0.5
16	Lakewood, OH	52,131	15.1	$27,452	$244	0.009	170.8	$135,616	0.5
17	San Mateo, CA	97,207	5.9	$44,949	$312	0.007	119.3	$247,161	0.5
18	Bloomington, MN	82,893	7.3	$34,400	$231	0.007	137.5	$157,981	0.5
19	Livonia, MI	96,942	5.0	$31,632	$204	0.006	156.0	$121,898	0.5
20	Santa Barbara, CA	88,410	14.1	$36,601	$381	0.010	155.6	$216,493	0.5
21	Meriden, CT	60,868	13.8	$27,625	$311	0.011	203.6	$132,685	0.5
22	Santa Clara, CA	116,468	8.6	$38,422	$380	0.010	129.2	$251,667	0.5
23	Somerville, MA	75,754	14.7	$32,517	$305	0.009	163.4	$152,442	0.6
24	Honolulu, HI	953,207	8.8	$29,516	$310	0.010	222.8	$111,955	0.6
25	Berkeley, CA	112,580	18.9	$36,498	$610	0.017	176.5	$276,945	0.6
26	Eugene, OR	156,185	20.7	$24,917	$292	0.012	123.1	$189,450	0.6
27	Yonkers, NY	195,976	13.8	$29,191	$516	0.018	311.4	$132,415	0.6
28	Glendale, CA	191,719	13.0	$29,823	$256	0.009	128.0	$157,361	0.6
29	Fullerton, CA	135,161	11.3	$30,580	$300	0.010	115.3	$204,863	0.6
30	Alhambra, CA	83,089	12.7	$24,327	$276	0.011	95.0	$224,309	0.6
B. Most Underpoliced Cities									
242	Gary, IN	80,294	34.2	$15,383	$4,376	0.284	266.2	$54,893	14.0
241	New Orleans, LA	343,829	24.4	$24,929	$3,963	0.159	405.9	$55,337	8.2
240	Flint, MI	102,434	36.6	$14,910	$3,291	0.221	190.0	$107,507	7.5
239	Saginaw, MI	51,508	37.4	$14,157	$2,970	0.210	187.7	$125,772	5.9
238	Youngstown, OH	66,982	32.7	$14,451	$3,057	0.212	238.9	$106,274	5.6
237	Detroit, MI	713,777	34.5	$15,062	$3,691	0.245	360.9	$95,934	5.0
236	Birmingham, AL	212,237	26.4	$19,775	$3,106	0.157	350.4	$85,184	4.8
235	Jackson, MS	173,514	26.6	$19,095	$2,198	0.115	261.0	$81,255	4.8
234	Baton Rouge, LA	229,493	25.5	$23,195	$2,286	0.099	288.2	$86,510	4.3
233	St. Louis, MO	319,294	26.0	$21,406	$3,486	0.163	400.7	$104,741	3.9
232	Pontiac, MI	59,515	32.0	$15,957	$1,918	0.120	177.2	$134,235	3.8
231	San Bernardino, CA	209,924	27.4	$15,616	$1,969	0.126	157.7	$156,615	3.7
230	Macon, GA	91,351	30.6	$17,110	$1,774	0.104	298.8	$75,503	3.7
229	Richmond, VA	204,214	25.3	$26,034	$2,614	0.100	351.6	$96,612	3.6
228	Richmond, CA	103,701	16.4	$24,847	$3,010	0.121	163.1	$240,409	3.6
227	Little Rock, AR	193,524	17.4	$29,229	$2,181	0.075	275.7	$105,633	3.5
226	Baltimore, MD	620,961	21.3	$23,333	$3,428	0.147	478.5	$97,815	3.4
225	Camden, NJ	77,344	36.1	$12,807	$4,352	0.340	510.1	$122,706	3.2
224	Columbus, GA	189,885	18.2	$22,514	$1,041	0.046	210.3	$71,691	3.2
223	Memphis, TN	646,889	25.4	$21,007	$2,019	0.096	310.4	$99,214	3.1
222	Atlanta, GA	420,003	22.6	$35,453	$2,133	0.060	333.3	$97,859	3.0
221	Shreveport, LA	199,311	22.1	$22,047	$1,618	0.073	281.4	$89,395	3.0
220	Albany, GA	77,434	32.0	$17,416	$1,170	0.067	234.8	$77,827	3.0
219	Oakland, CA	390,724	18.7	$30,671	$2,516	0.082	182.2	$221,429	2.9
218	Tulsa, OK	391,906	19.3	$26,069	$1,565	0.060	205.4	$122,882	2.9
217	North Little Rock, AR	62,304	22.5	$23,578	$1,723	0.073	317.8	$87,544	2.9
216	Knoxville, TN	178,874	23.4	$21,964	$1,292	0.059	211.7	$99,491	2.9
215	Dayton, OH	141,527	31.0	$16,702	$1,955	0.117	272.0	$117,528	2.8
214	Indianapolis, IN	820,445	17.9	$24,334	$1,431	0.059	200.3	$117,840	2.8
213	Mobile, AL	195,111	21.5	$22,401	$1,122	0.050	209.1	$89,803	2.8

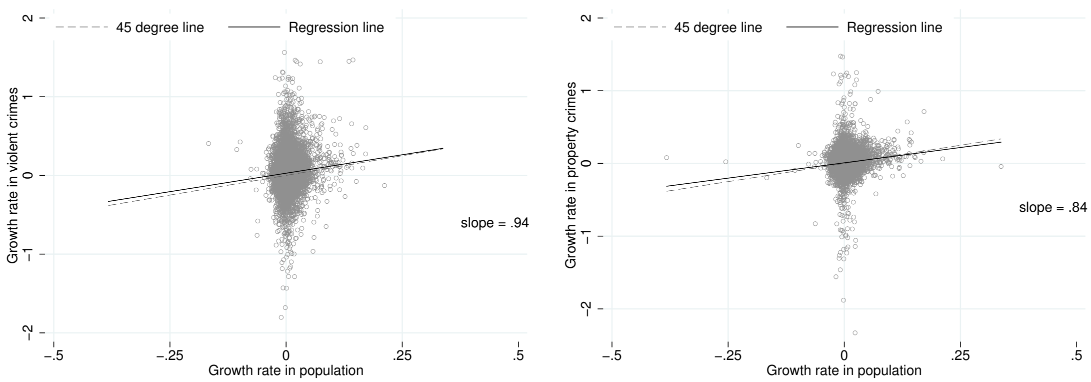
Note: Population and dollar figures pertain to 2010. Poverty rate and per capita income are from the Ammerican Community Survey 5-year estimates. Cost of crime is cost-weighted sum of crimes per capita. Fraction income at risk is cost of crime relative to income per capita. Estimated cost per officer is a cross-sectional regression-predicted operating budget per sworn officer, where the cross-sectional data are themselves city-specific medians from 2003-2010 from the Annual Survey of Government data. For further details see text.

Appendix Figure 1A. Correlations Between UCR Population Measure and Crime

Panel A. Raw Population Measure

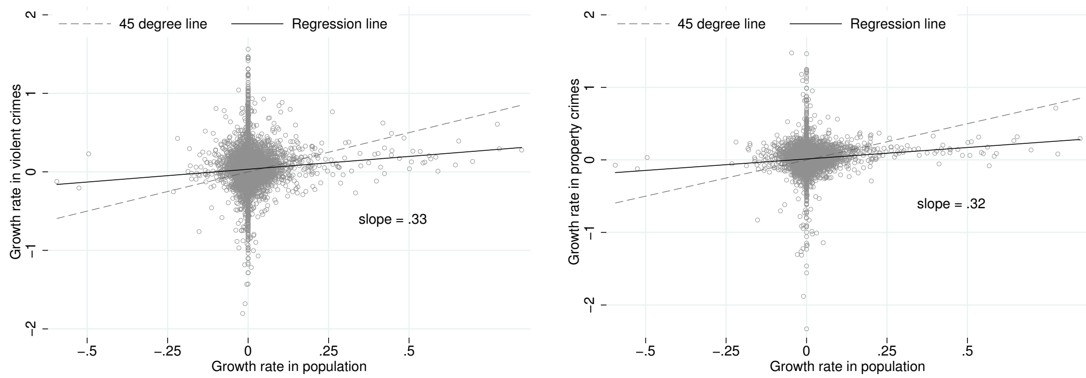


Panel B. Smoothed Population Measure

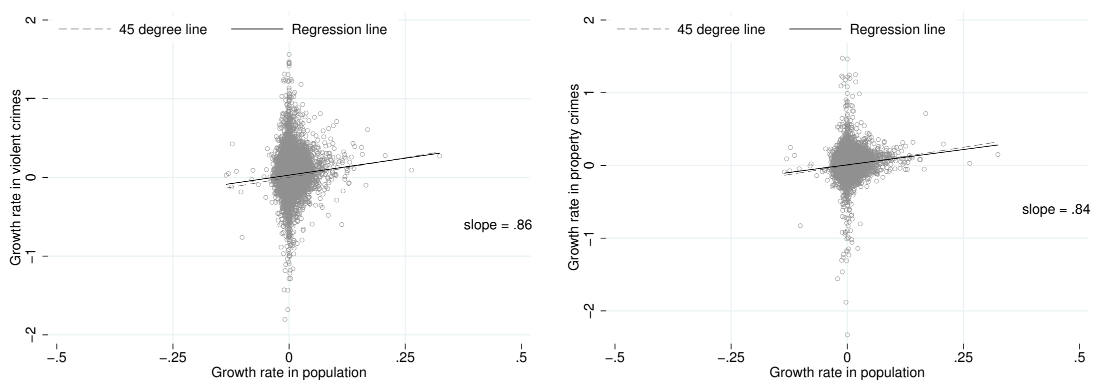


Appendix Figure 1B. Correlations Between ASG Population Measure and Crime

Panel A. Raw Population Measure



Panel B. Smoothed Population Measure



Appendix Table 1. Sensitivity of Pooled GMM Estimates to the Inclusion of Additional Population Measures

	Violent Crime	Murder	Rape	Robbery	Assault	Property Crime	Burglary	Larceny	Motor Vehicle Theft
Panel A. Data, 1960-1993 [Full Sample]									
without births measure	-0.162 (0.096)	-0.784 (0.240)	0.075 (0.173)	-0.585 (0.119)	0.216 (0.123)	-0.145 (0.069)	-0.200 (0.098)	-0.042 (0.078)	-0.313 (0.115)
with births measure	-0.177 (0.096)	-0.801 (0.238)	0.064 (0.173)	-0.602 (0.119)	0.203 (0.123)	-0.153 (0.068)	-0.208 (0.097)	-0.048 (0.077)	-0.338 (0.112)
Panel B. Data, 1960-2003 [N=147 cities]									
without births measure	-0.251 (0.095)	-0.548 (0.246)	0.141 (0.168)	-0.661 (0.119)	0.112 (0.121)	-0.195 (0.070)	-0.216 (0.096)	-0.120 (0.078)	-0.346 (0.120)
with births measure	-0.263 (0.095)	-0.565 (0.245)	0.127 (0.167)	-0.675 (0.119)	-0.102 (0.121)	-0.203 (0.069)	-0.224 (0.095)	-0.126 (0.078)	-0.369 (0.117)
Panel C. Data, 1970-2008 [Full Sample]									
without SEER measure	-0.259 (0.089)	-0.793 (0.234)	0.027 (0.160)	-0.628 (0.108)	0.072 (0.115)	-0.194 (0.064)	-0.242 (0.089)	-0.109 (0.072)	-0.320 (0.108)
with SEER measure	-0.270 (0.089)	-0.807 (0.233)	0.017 (0.159)	-0.640 (0.108)	0.062 (0.115)	-0.201 (0.064)	-0.248 (0.088)	-0.114 (0.071)	-0.341 (0.105)

Note: Each column reports results of an estimation procedure in which the “forward” and “reflected” IV regression coefficients are pooled via GMM. In each panel, the top row of estimates report parameter estimates that are conditional on the UCR and ASG population meausures. The bottom row reports models which also include a third proxy for the growth rate in population (either using NCHS natality data or the SEER population measure). In Panel A, we estimate models on the full sample of data using the years 1960-1993, the years for which births data are available for all cities. In Panel B, we estimate models using the years 1960-2003 using the sample of 147 cities for which we have births data for this period. In Panel C, we estimate models on the full sample of data using the years 1970-2008, the years for which the SEER population measure is available. All models are conditional on a vector of unrestricted state-by-year dummies and are estimated using 2010 city population weights. As clustered and robust standard errors are very similar across all models, we report Huber-Eicker-White standard errors that are robust to heteroskedasticity in parentheses below the coefficient estimates.

Appendix Table 2. Testing the "Displacement" Hypothesis

	Violent Crimes				Property Crimes			Aggregates		
Sample	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft	Violent Crime	Property Crime	Cost-Weighted Crimes
City-level (N=242 cities)	-0.666 (0.238)	-0.255 (0.219)	-0.559 (0.117)	-0.099 (0.127)	-0.225 (0.089)	-0.083 (0.067)	-0.343 (0.101)	-0.344 (0.096)	-0.174 (0.062)	-0.473 (0.171)
MSA-level (N=167 MSAs)	-0.914 (0.366)	-0.268 (0.235)	-0.611 (0.157)	-0.060 (0.176)	-0.161 (0.149)	-0.135 (0.106)	-0.358 (0.151)	-0.335 (0.129)	-0.165 (0.092)	-0.667 (0.267)

Note: Each column reports generalized method of moments (GMM) estimates of the growth rate in each of ten crime rates on the first lag of the growth rate in the number of sworn police officers, conditional on both the UCR and ASG measures of the growth rate in population and a vector of unrestricted state-by-year effects. The top panel of the table replicates the two-step GMM estimates reported in Table 5 using our original sample of 242 cities. The bottom panel of the table reports two-step GMM estimates arising from a dataset in which we have collapsed our original sample to the level of the MSA. All models use 2010 city population weights, and Huber-Eicker-White standard errors are reported in parentheses.