Are U.S. Cities Underpoliced?: Theory and Evidence*

Aaron Chalfin
School of Criminal Justice
660 Dyer Hall
University of Cincinnati
Cincinnati, OH 45221
Phone: (510) 219-5718
E-Mail: aaron.chalfin@uc.edu

Justin McCrary
Berkeley Law
586 Simon Hall
University of California
Berkeley, CA 94720
Phone: (510) 643-6252
E-Mail: jmccrary@law.berkeley.edu

May 30, 2013

Abstract

The socially optimal number of police hinges on the extent to which police reduce the most costly crimes, which are also the most difficult to model econometrically because they are rare. In the hope of minimizing simultaneity bias, papers in the recent literature have focused on quasi-experimental approaches that disregard most of the variation in police staffing levels, compounding the modeling difficulty. We argue that the central empirical challenge in this literature is not simultaneity bias, as has been supposed, but measurement error bias. 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, with much greater parameter certainty for the most costly crimes. Our analysis suggests that U.S. cities are in fact underpoliced.

JEL Classification: K42, H76, J18

Keywords: Police, crime, measurement error

*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, Jeff 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. An earlier draft of this manuscript circulated under the title "The Effect of Police on Crime: New Evidence from U.S. Cities, 1960-2010."

I. Introduction

One of the most intuitive predictions of deterrence theory is that an increase in an offender's chances 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, a large literature focuses on the effect of police on crime, where police are viewed as a primary factor influencing the chances of apprehension.[2] This literature is ably summarized by Cameron (1988), Nagin (1998), Eck and Maguire (2000), Skogan and Frydl (2004), and Levitt and Miles (2006, 2007), all of whom provide extensive references.

The early panel data literature tended to report small elasticity estimates that were rarely distinguishable from zero and sometimes even positive, suggesting perversely that police increase crime.[3] The ensuing discussion in the literature was whether police reduce crime at all. Starting with Levitt (1997), an emerging quasi-experimental literature has argued that simultaneity bias is the culprit for the small elasticities in the panel data literature.[4] The specific concern articulated is that if police are hired in anticipation of an upswing in crime, then there will be a positive bias associated with regression-based strategies, masking the true negative elasticity. This literature has therefore focused instead on instrumental variables (IV) or difference-in-difference strategies designed to overcome this bias. These strategies consistently demonstrate that police do reduce crime. However, the estimated elasticities display a wide range, roughly -0.1 to -1, depending on the study and the type of crime.

In this paper, we argue that the central empirical issue confronting the literature is not whether police affect crime, but the extent to which police reduce violent crime, particularly murder. We formalize this

[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.

[3] 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).

[4] Prominent quasi-experimental papers after Levitt (1997) include Di Tella and Schargrodsky (2004), Klick and Tabarrok (2005), Evans and Owens (2007), Draca, Machin and Witt (2011), Machin and Marie (2011), and Vollaard and Hamed (2012).

point in Section II with a social welfare analysis. The analysis shows that at current staffing levels U.S. cities are almost surely underpoliced if police appreciably reduce violent crimes, but are almost surely overpoliced if they do not. This conclusion is rooted in the extraordinary cost of most violent crimes and the comparatively minor cost of most property crimes.

Unfortunately, papers in the recent quasi-experimental literature present suggestive but not persuasive evidence regarding the effect of police on violent crime. Compounding the fact that quasi-experimental research designs purposefully disregard most of the variation in police staffing levels, a further empirical challenge is that the most costly crimes are rare. Rare crimes have highly variable crime rates and even more variable growth rates, leading to parameter uncertainty. Consequently, we still know little about the elasticities that are central to a social welfare evaluation.

The leading example is the police elasticity of murder. Two prominent papers using U.S. data are Levitt (1997, murder elasticity of -3.05 ± 4.06) and Evans and Owens (2007, elasticity of -0.84 ± 0.94).[5] Both confidence intervals are wide enough to incorporate very large elasticities (e.g., -1.5) as well as zero. Meanwhile, another prominent study estimates a police elasticity of violent crime of zero and argues that it is implausible that police affect murder (Klick and Tabarrok 2005, fn. 24).

As noted, many recent studies disregard most of the variation in police due to concerns over simultaneity bias. An obvious way to improve the precision of police elasticities is to return to regression-based methods with appropriate controls, as in Marvell and Moody (1996), for example. Importantly, however, this type of approach has the potential to run afoul of the "iron law of econometrics," or the tendency of regression coefficients to be too small because of errors in the measurement of the variable of interest (Hausman 2001). Most quasi-experimental approaches, such as IV, do not suffer from the same bias (Bound, Brown and Mathiowetz 2001), at least under the hypotheses of the classical measurement error model.

The implications of measurement errors in police for the estimated police elasticity of crime has, prior to this work, gone unaddressed in the crime literature.[6] We document that there is a high degree of measurement error in the basic dataset on police used in the U.S. literature, the Uniform Crime Reports (UCR). We estimate that prior regression-based estimates are too small by a factor of four or five, providing an alternative explanation for the small size of the elasticities from the prior panel data literature.

[5] For Levitt (1997), we cite the corrected elasticity estimate and confidence interval from McCrary (2002).

[6] In contrast, a great has been written on measurement errors in crime. Non-classical measurement errors in crime are the focus of an extensive literature within criminology (see Mosher, Miethe and Hart (2011) for a review). Within economics, non-classical measurement errors in crime are the subject of two papers using U.S. data (Levitt 1998a,b), and a paper using British data (Vollaard and Hamed 2012). None of these papers contemplates measurement errors in police.

The core of our empirical analysis 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. These estimates take advantage of measures of city police from the Annual Survey of Government (ASG) and maintain the hypotheses of the classical measurement error model. A methodological contribution of the paper is that we obtain a more efficient estimator of the policy parameter by exploiting the inherent symmetry of the classical measurement error model. A second methodological contribution is that we show how that symmetry implies new tests for the restrictions of the classical measurement error model. We find little evidence against the restrictions of the classical measurement error model in our data.

Our estimated police elasticities are substantively large, roughly four times as large as those from the traditional literature using natural variation and in line with some of the larger estimates from the quasi-experimental literature. For example, our best guess regarding the elasticity for murder is -0.67 ± 0.48. Combining our empirical analysis with the social welfare framework suggests reduced victim costs of \$1.63 for each additional dollar spent on police in 2010, implying that U.S. cities are in fact underpoliced. To the extent that lingering simultaneity bias affects our estimates, this conclusion is conservative. Importantly, our estimates are robust to controlling for the confounders mentioned in the quasi-experimental literature, including demographic factors, the local economy, city budgets, social disorganization, and any possible state-level policy changes that have the same effect across cities (e.g., sentencing reform, education policy 10-20 years ago, and so on).

The remainder of the paper is organized as follows. Section II presents our social welfare framework. Section III presents direct evidence on the degree of measurement error. We then outline our econometric methodology in Section IV, discuss our primary data in Section V, and report estimated police elasticities of crime in Section VI. In Section VII, we compare our results to those from the previous literature. Section VIII connects our estimated elasticities to the social welfare framework from Section II, and Section IX concludes.

II. Conceptual Framework

In this section, we outline a framework for deriving the optimal number of police. We begin by considering a simple case in which individuals are identical and there is only one type of crime. We then extend the model to accommodate multiple crime types which vary according to their social cost. Finally, we extend the model to accomodate externalities in private precautions, crowd-out, and heterogeneity. This framework emphasizes two main points. First, reasonable calibrations suggest that additional investments in police are unlikely to be socially beneficial unless police reduce violent crimes to at least a moderate degree. Reductions in property crime are simply not sufficiently costly to justify the expense of additional police officers. Violent crimes, on the other hand, are extremely costly and consequently even relatively small effects of police on violent crime would be sufficient to justify additional investment in police. Second, empirical estimates of the police elasticity of crime can be used in conjunction with prices to assess social optimality, even in a relatively complicated scenario in which public police crowd out private precautions and those precautions involve externalities.

A. Base Case

Suppose society is comprised of n identical individuals, each of whom confronts a probability of criminal victimization $\phi(S)$, where S is the number of police employed by the government.[7] Each individual faces a victimization cost of k and has assets A that could be spent on consumption. To keep the presentation as simple as possible, we restrict attention to the case of linear utility.[8] Individuals pay a lump-sum tax τ to pay for police, and the cost of an officer is w. Table 1 shows our estimate of the annual "fully-loaded" cost of a police officer in 2010 of \$130,000.[9] On a per capita basis, this works out to \$341, or about 1.3 percent of income.

The social planner maximizes the expected utility of the representative agent, subject to the *financing constraint* that tax receipts must equal the total wages paid police, or $n\tau = wS$. This implies a social welfare function of

$$V(S) \quad = \quad y(S) - C(S) \tag{1}$$

where $C \equiv C(S) = k\phi(S)$ is the expected cost of crime and $y(S) = A - \tau = A - wS/n$ is consumption in the absence of crime and subject to the financing constraint.[10] The first-order necessary condition for this problem, which is also sufficient, is of course $0 = V'(S)$, but it is convenient to analyze instead the proportional condition

$$0 \quad = \quad V'(S)\frac{S}{C} = y'(S)\frac{S}{C} - C'(S)\frac{S}{C} \equiv -\frac{wS}{nC} - \varepsilon \tag{2}$$

where $\varepsilon \equiv \partial \ln C/\partial \ln S$ is the police elasticity of the cost of crime, and $y'(S) = -w/n$. Next, note that

[7] We assume that $\phi(\cdot)$ is differentiable and strictly convex.

[8] More generally, a third-order Taylor approximation to utility in conjunction with typical estimates of the coefficients of relative risk aversion and prudence (Chetty 2006) suggest that linear utility is a good approximation.

[9] This estimate, which is specific to the 242 large U.S. cities we study empirically below, is based on total police operating budgets relative to the total number of officers. This is closer to the concept employed in Levitt (1997) (who obtains \$133,000 in 2010 dollars) than to the pure marginal cost concept employed in Evans and Owens (2007) (who obtain \$73,000 in 2010 dollars). The data on operating budgets are taken from the Annual Survey of Government (ASG) Finance files, and the data on the number of officers are taken from the ASG Employment files. To accomodate outliers in the budget data, which are prevalent, we compute a city-specific median of the per sworn officer budget from 2003 to 2010, after adjusting each year's budget to 2010 dollars.

[10] 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). See Cameron (1989) for a valuable discussion of these conceptual issues.

in this framework, an increase in policing improves the welfare of the representative agent when policing passes a cost-benefit test. Formally,

$$V'(S) > 0 \qquad \Longleftrightarrow \qquad |\varepsilon| > \frac{wS}{nC} \tag{3}$$

B. Multiple Crime Categories

Now suppose there are multiple crime categories.[11] The probability of victimization is $\phi_j(S)$ and the cost of crime is k_j, where j ranges from 1 to J. This leads to a redefinition of the expected cost of crime: $C(S) = \sum_{j=1}^{J} k_j \phi_j(S)$. With these redefinitions, equations (1), (2), and (3) remain the same as above.

However, it is useful to rewrite the aggregate police elasticity, ε, in terms of the elasticities for specific crime categories. Minor rearrangement shows that the aggregate elasticity is a weighted average of elasticities for individual crime categories, or

$$\varepsilon \;=\; \frac{\sum_{j=1}^{J} k_j \phi_j(S) \varepsilon_j}{\sum_{j=1}^{J} k_j \phi_j(S)} \tag{4}$$

where the weights, $k_j \phi_j(S)$ are the expected cost of the crime categories and $\varepsilon_j = \partial \ln \phi_j(S) / \partial \ln S$ is the police elasticity for crime type j.

The crime-specific elasticities ε_j are the focus of most of the empirical literature on the effect of police on crime. Estimates are available for the seven so-called "index offenses" captured by the Uniform Crime Reports (UCR) system of the Federal Bureau of Investigation (FBI). Table 1 displays the costs associated with these crimes (k_j) as well as their prevalence in the population (ϕ_j scaled by 100,000) and the expected cost ($k_j \phi_j$).[12] Totalling across crime categories yields $C = \$995$, which is about 3.8 percent of per capita income in our sample. The cost figures in Table 1 thus imply that $wS/(nC)$ is about 0.34.

Some simple arithmetic using the cost figures in Table 1 in connection with the framework sketched above shows that the key policy question for this literature is not *whether* police affect crime, but the extent to which police affect *violent* crime, particularly murder.[13] To appreciate this point, suppose that

[11] Without loss of generality, we define crime categories to be mutually exclusive so that the probability of being victimized by no crime is $1 - \sum_{j=1}^{J} \phi_j(S)$.

[12] Analogous to the figure for the cost of police, the prevalence of crime is specific to our sample of 242 cities and pertains to 2010. The figures on the cost of crime are drawn from the literature, the most recent of which is Cohen and Piquero (2009), augmented by estimates of the value of a statistical life (VSL). As we discuss below, the \$7 million figure we use for murder is approximately the VSL number used by most large federal and state agencies in calibrating public safety investments. The *ex ante* perspective adopted in constructing VSL figures is the appropriate one for this context. Unfortunately, for crimes other than murder, the only study to utilize an *ex ante* perspective is Cohen, Rust, Steen and Tidd (2004). Their methodology involved a contingent valuation survey in which individuals were asked to choose from among several different hypothetical dollar amounts in order to protect themselves from crime. The resulting cost estimates are much larger—often 1 to 2 orders of magnitude larger—than those given in Table 1. We use the more conventional victim cost approach to be conservative.

[13] Levitt (1997) makes a similar point in emphasizing the reliance of his cost-benefit calculation on the magnitude of the

the police elasticity of crime was -1 for each property crime category, but 0 for each violent crime category. Then using equation (4) and the cost and incidence figures from Table 1, we see that the cost-weighted elasticity would be a scant -0.07—a notable departure from -0.34, the value of the cost-weighted elasticity that would justify hiring additional police.[14]

Indeed, if police had no effect on violent crime, then for the cost-weighted elasticity to be -0.34 would require the elasticities for property crimes to be at least as negative as -6. Elasticities as large as that are clearly implausible. Nearly all credible studies report property crime elasticity estimates of between 0 and -1. Fixing all property crime elasticities somewhat optimistically at -1, for the cost-weighted elasticity to be -0.34 would require the elasticities for violent crime to be at least as negative as -0.31.

Violent crime elasticities are decisive in a social planner's calculations because violent crimes are so much more costly than property crimes. The extreme case is murder. Even though it is exceedingly rare—occurring at a rate one-third that of the second rarest crime, rape, and one-fiftieth that of motor vehicle theft—murder accounts for fully sixty percent of the per capita expected cost of all crime (Table 1). This means that the murder elasticity is particularly important to a social planner's calculations. If the police elasticities of crime were -1 for all crimes, but 0 for murder, then the cost-weighted elasticity would still be only -0.3. If more plausibly the police elasticity was -0.75 for all crimes except murder, the murder elasticity would have to be at least as negative as -0.2 to lead to a cost-weighted elasticity of -0.34.

C. Crowd-out, Externalities, and Heterogeneity

Investments in public police are not the only way individuals seek to protect themselves against crime. Individuals may purchase a car alarm, upgrade a deadbolt, or pay the higher rent of a safer neighborhood, all in an attempt to provide safety to themselves and those near them. Public investments in police could crowd out some of these private investments in precaution. Moreover, private precautions may involve externalities. These dual considerations suggest that a more detailed analysis is warranted before accepting the provisional conclusions of the baseline models sketched above.

Consider then a richer model in which heterogeneous individuals undertake private precautions that may involve externalities. We continue to assume linear utility and lump-sum taxes, and we further assume that

murder elasticity.

[14] We note that there are certainly benefits from policing that are not captured by the seven index offenses (e.g., arrests for other crime categories, or emergency medical response). In this section, we are pointing out the effect of hypothetical changes to elasticities on the cost-weighted elasticity. We do not mean to say that if the cost-weighted elasticity is not as negative as -0.34 then there should be fewer police. Rather, we mean simply to say that a cost-benefit analysis focused on the seven index offenses would not justify the existing number of police.

individuals 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.

Analogous to the base case discussed above, expected utility is given by $y_i(X_i, S) - C_i(X_1, X_2, \ldots, X_n, S)$ where $y_i(X_i, S) = A_i - wS/n - p_i X_i$ is after-tax wealth net of expenditures on precautions, A_i is assets, p_i is the price of precaution, and $C_i(X_1, X_2, \ldots, X_n, S) = k_i \phi_i(X_1, X_2, \ldots, X_n, S)$ is the expected cost of crime.[15] 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.[16]

In addition to the financing constraint mentioned above, we impose a *liberty* constraint that the social planner is unwilling or unable to dictate an individual's precautions. To motivate the liberty constraint, note that in the real world 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. That is, the liberty constraint is one that actual governments respect. To clarify that these 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*.

The state seeks to maximize average expected utility subject to the financing and liberty constraints. This can be thought of as (1) delegating to each individual the choice of precaution; (2) restricting beliefs regarding precautions to be mutually consistent; and (3) maximizing the average equilibrium indirect utility function over policing. To solve the state's problem, then, we begin by solving the individual's problem.

The individual's first order necessary condition for precaution, which is also sufficient, is $p_i = -k_i \phi_{ii}$, where at some small risk of confusion the second subscript indicates a partial derivative.[17] Solving the first order condition for X_i leads to a reaction function, $X_i(X_{-i}, S)$, where X_{-i} is the vector of precautions for individuals other than i.[18]

Under the assumptions above, each individual has a unique best strategy for any given set of beliefs regarding the actions of other individuals, and we obtain a Nash equilibrium in pure strategies (Fudenberg and Tirole 1991, Theorem 1.2). Figure 1 shows reaction functions with two individuals under functional

[15]We assume that ϕ_i is differentiable in all arguments and strictly convex in X_i and in S. As before, in the case of multiple crimes the expected cost of crime can be redefined as $\sum_{j=1}^{J} k_{ij} \phi_{ij}(X_1, X_2, \ldots, X_n, S)$.

[16]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 bicycle lock. In these examples, the price of precaution is either the cost of the additional travel time or the market price of the lock.

[17]We assume that precautions and policing are both protective against crime, or that $\phi_{ii} < 0$ and $\phi_{iS} < 0$.

[18]We suppress the dependence of the reaction function on prices, taxes, and initial assets.

form assumptions under high and low policing.[19] Write equilibrium demand for precaution as $X_i(S)$.[20] Substituting the equilibrium demand functions into the individual's utility function yields the equilibrium indirect utility function $V_i(S)$.[21] The state then maximizes $\mathcal{V}(S) \equiv \frac{1}{n}\sum_i V_i(S)$. In this framework, police affect expected utility 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 own precaution ($-p_i X_i'(S) > 0$);
3. additional police lower utility by crowding out own 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, which is also sufficient, reflects these different mechanisms. For convenience we again analyze the first order condition times S/C. We have

$$\mathcal{V}'(S)\frac{S}{C} = -\frac{wS}{nC} - \sum_{i=1}^{n} \omega_i \rho_i \eta_i - \sum_{i=1}^{n} \omega_i \varepsilon_{ii} \eta_i - \sum_{i=1}^{n} \omega_i \varepsilon_{iS} - \sum_{i=1}^{n} \omega_i \sum_{\ell \neq i} \varepsilon_{i\ell} \eta_\ell \tag{5}$$

where $\omega_i = k_i \phi_i / \sum_{i=1}^{n} k_i \phi_i$ is the fraction of the social expected cost of crime borne by person i, $\rho_i = p_i X_i(S)/(k_i \phi_i)$ 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 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 $\eta_i = X_i'(S) S / X_i(S) < 0$ is the elasticity of precaution for person i with respect to policing. The five terms in equation (5) correspond to the five different mechanisms described above.[22]

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

[19] 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 $X_i(X_{-i}, S) = (1/\alpha)\left(\ln(\alpha k_i/p_i) - \beta X_{-i} - \gamma S\right)$ for $i = 1, 2$. This formulation thus echoes the traditional textbook treatment of Cournot duopoly with linear demand (e.g., Tirole 1988, Chapter 5).

[20] Nash equilibrium requires that beliefs be mutually consistent, which implies a set of restrictions leading 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) and corresponds to the intersection of the reaction functions in Figure 1.

[21] This is given by $V_i(S) = y_i(X_i(S), S) - k_i \phi_i(X_1(S), X_2(S), \ldots, X_n(S), S)$.

[22] If individuals are taking optimal precautions, then the second and third mechanisms exactly offset (the envelope theorem). We do not impose the envelope theorem at the outset, because it complicates somewhat the analysis below connecting $\widetilde{\theta}$ and θ to social welfare. We invoke it below when it simplifies the analysis.

reduction in crime due to increased police. While the distinction between these estimands has not garnered much attention in the empirical literature to date, it is important for drawing conclusions about social welfare.

Our study focuses on changes in crime associated with year-to-year fluctuations in policing. Whether this is more likely to identify a partial or total police elasticity depends on the nature of precautions. 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, then our study is more likely to identify a partial elasticity. If most precautions are able to adjust within a relatively short time frame, such as a year, then our study is more likely to identify a total elasticity. We show how to calibrate our empirical findings to our social welfare framework under the extreme cases that (1) precautions fully adjust and (2) entirely fail to adjust, as well as under the hybrid case that (3) some precautions adjust and others do not.

To begin, note that the total and partial elasticities are given by

$$\widetilde{\theta} = \frac{S}{C}\frac{1}{n}\sum_{i=1}^{n} k_i\{\phi_{ii}X_i'(S) + \phi_{iS} + \sum_{\ell\neq i}\phi_{i\ell}X_\ell'(S)\} = \sum_{i=1}^{n}\omega_i\big(\varepsilon_{ii}\eta_i + \varepsilon_{iS} + \sum_{\ell\neq i}\varepsilon_{i\ell}\eta_\ell\big) \tag{6}$$

$$\text{and } \theta = \frac{S}{C}\frac{1}{n}\sum_{i=1}^{n} k_i\phi_{iS} = \sum_{i=1}^{n}\omega_i\varepsilon_{iS} \ , \tag{7}$$

respectively.[23] Next, combining equations (5), (6), and (7), we have

$$\mathcal{V}'(S)\frac{S}{C} = -\frac{wS}{nC} - \sum_{i=1}^{n}\omega_i\rho_i\eta_i - \widetilde{\theta} \equiv -\frac{wS}{nC} - r - \widetilde{\theta} \tag{8}$$

$$= -\frac{wS}{nC} - \sum_{i=1}^{n}\omega_i\sum_{\ell\neq i}\varepsilon_{i\ell}\eta_\ell - \theta \equiv -\frac{wS}{nC} - e - \theta \tag{9}$$

where $r = \sum_{i=1}^{n}\omega_i\rho_i\eta_i$ is the *crowd-out 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 signs of the crowd-out and externality effects will be important for some of our reasoning. Consider first the crowd-out effect. While we can imagine that a given individual might perversely increase precaution with increased policing, we assume that, at least on average in the population, policing crowds out precautions.[24] Since $\rho_i \geq 0$, this implies $r \leq 0$.

[23]Note that in equation (7), we invoke the envelope theorem.

[24]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.

The sign of the externality effect is somewhat more ambiguous. On the one hand, we suspect 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, some precautions have aspects of *both* positive and negative externalities.[25] Consequently, although we have a prior view, when we calibrate this framework to our empirical analysis, in Section VIII, below, we allow for both positive and negative externality effects.

As noted, equations (8) and (9) are both proportional to the state's first order condition. 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 crowd-out effect. Our next task is to sign the terms that we are unable to estimate (r and e) and then derive inequalities characterizing the social optimum.

Consider first the possibility that our empirical analysis identifies the total elasticity, $\widetilde{\theta}$, i.e., that precautions adjust quickly, and define $\widetilde{\kappa} = |\widetilde{\theta}|/(wS/nC)$. Rearranging equation (8) shows that

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

Suppose that increasing police is worthwhile in the provisional cost-benefit sense that $\widetilde{\kappa} > 1$. Since r and $\widetilde{\theta}$ share sign, the adjustment term $1 + r/\widetilde{\theta}$ is bigger than one, and 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, and define $\kappa = |\theta|/(wS/nC)$, where we note that $|\theta| > |\widetilde{\theta}|$ so that $\kappa > \widetilde{\kappa}$. Rearranging equation (9) shows that

$$\mathcal{V}'(S) > 0 \quad \Longleftrightarrow \quad \theta\left(1 + e/\theta\right) < -\frac{wS}{nC} \tag{11}$$

Suppose now that increasing police is worthwhile in the provisional cost-benefit sense that $\kappa > 1$. 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

[25] 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.

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 $e < |\theta|$.[26] We then have the bounds $0 < 1 + e/\theta < 1$, and the conclusion that

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

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

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

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.

In the more realistic scenario in which some precautions are quick to respond and others are slow, our study identifies a weighted average of these two elasticities. A simple way to make this point in the framework above is to allow a fraction q of individuals to have precautions that adjust slowly and a fraction $1-q$ of individuals to have precautions that adjust quickly. Then it is not hard to show that a study like ours identifies a combination parameter $\overline{\theta} \equiv (1-q)\widetilde{\theta} + q\theta$. The first order condition for maximizing social welfare is then proportional to

$$\mathcal{V}'(S)\frac{S}{C} = -\frac{wS}{nC} - a - \overline{\theta} \tag{14}$$

where $a \equiv qe + (1-q)r$ is an adjustment term that reflects both externality and crowd-out effects in accordance with their prevalence in the population. The analysis above then applies with minor modifications.

III. The Extent of Measurement Error in the Number of Police

We begin our discussion of the nature and extent of measurement errors in police personnel data using as an example the case of New York City in 2003. The UCR data for New York show 28,614 sworn police officers in 2003. Relative to the 37,240 and 35,513 sworn officers employed in 2002 and 2004, respectively, this is a remarkably low number. If the UCR figures are to be believed, New York lost a quarter of their sworn officers in 2003 and then hired most of them back the next year.[27]

[26] Since θ is negative and $wS/(nC)$ is positive, the second inequality in (11) 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.

[27] The UCR data also indicate that New York lost a fifth of their civilian police employees in 2003 and then gained them all back in 2004, arguing against confusion over sworn officers versus civilian employees.

An alternative interpretion is that the 2003 number is a mistake. Internal documents from New York are available that shed light of the UCR records.[28] Setting aside the data for 2003, the UCR series and the internal documents series track reasonably well; after discarding the data for 2003 and 2004, the correlation is 0.92 in levels and 0.56 in growth rates. The internal documents show that the number of sworn officers in 2003 was 36,700, not 28,614, indicating that the UCR data is incorrect.

Administrative data on police such as these are difficult to obtain. Some departmental annual reports are available, but they are not practical for econometric research. Annual reports do not circulate widely and even for cities and years where they are available, they do not always report the number of officers.[29] Trading off the accuracy of administrative data for the coverage of survey data, we now present a comparison of the UCR series on the number of sworn officers with a series based on a separate survey collecting information on police officers, the ASG. These data are collected by the U.S. Census Bureau, rather than the FBI, and are filled out by officials in city-wide government, rather than by the police department specifically.[30] 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 more background on the ASG data in Section V, below.

Figure 2 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 (Cameron and Trivedi 2005, Section 26.2.5).

Many people are suprised that there are errors in measuring the number of police officers. After all, a great deal of ink has been spilled on the topic of errors in the measurement of crime, but nearly nothing has been written on the subject of errors in the measurement of police.[31] Aside from obvious problems with transcription

[28] Thanks to Franklin Zimring, the internal documents of the New York City Police Department cited are available at `http://www.oup.com/us/companion.websites/9780199844425`.

[29] See the working paper version for some limited comparisons of the UCR with administrative data and data from annual reports. An interesting and econometrically problematic pattern in annual reports is the tendency to omit police numbers when other sources indicate declining police force size.

[30] The ASG collects information on all city government employees, while the UCR collects information only on police officers.

[31] Extensive references to the large literature on measurement errors in crime data are given in Mosher, Miethe and Hart (2011). 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) and by King, Cihan and Heinonen (2011). However, these papers do not

errors or computer programming errors, errors in measuring police could arise due to (1) transitory movements within the year in the number of sworn officers, (2) conceptual confusion, or (3) organizational confusion.

Regarding the first source of error, we are not aware of any public-use data sets containing information on within-year fluctuations in police. However, during the period 1979-1997, a unique non-public dataset on sworn officers in Chicago is available that allows the construction of monthly counts.[32] In that data set, a regression of the year-over-year growth rate in sworn officers on year indicators yields an R^2 of 0.71, suggesting that maybe a third of the movement in police growth rates is transitory.[33] This point is particularly relevant, as different data sources ask for a count of officers as of different snapshots in time, or are ambiguous about the relevant date.

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").[34] Using auxiliary data from the Law Enforcement Management and Administrative Survey (LEMAS), described below in Section V, we collected measures of the number of authorized and deployed sworn officers for selected recent years. These data indicate that the number of deployed sworn officers ranges from 62 to 128 percent of authorized strength.[35]

Finally, the UCR measure of sworn police has errors that may be the product of organizational confusion. For example, the internal documents for New York discussed above list the total number of sworn officers in the department as well as the number of officers assigned to one of the six largest bureaus.[36] For 2003, that latter figure was 26,367, which is notably below the average daily total staffing of 36,700, but close to the 28,614 reported to the UCR system. Alternatively, the 2003 number may have reflected ongoing

discuss the potential for measurement errors at the city level to bias estimates of the police elasticity derived from panel data.

[32] These data are discussed in Siskin and Griffin (1997) and were previously used in McCrary (2007).

[33] This does not reflect seasonality, as monthly indicators raise the R^2 by only 0.0001.

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

[35] The population weighted mean and standard deviation of the ratio are 97 percent and 5 percent, respectively. Numbers refer to a pooled analysis of all available years of the LEMAS data. The LEMAS data also allow us to discount the possibility that there is error due to different rules for accounting for full- or part-time workers, as they show that at most 2 percent of sworn officers work part-time.

[36] These are patrol (71 percent of total), detective (9 percent), transit (8 percent), housing (7 percent), narcotics (4 percent), and vice (1 percent). Numbers taken from 2009 data, but other years are similar.

confusion over the 1995 consolidation of the New York Police Department with the police departments of the New York City Transit Authority (April 1995) and the New York City Housing Authority (May 1995), which in 2003 together comprised approximately 5,550 officers.[37]

Since there is little hope of obtaining perfect data, it is reasonable to propose simple models of the measurement process and ask what they might imply about the econometric quantities being measured in the literature. The workhorse model in this context is the classical measurement error model, which we now introduce.

Suppose that in growth rates the two observed series on police (UCR and ASG) are related to true police as

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

and suppose crime growth rates, Y_i, are given by

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

Here, S_i is the UCR measure in a given city and year, Z_i is the ASG measure, S_i^* is the "true" police growth rate 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 together constitute the classical measurement error model (Fuller 1987).

A famous result from the econometric literature on measurement errors (see, for example, Wooldridge (2002, Section 4.4.2) or Cameron and Trivedi (2005, Section 26.2.4)) relates the probability limit of the OLS estimate of θ, based on using the covariates X_i and the proxy S_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:

$$\text{plim}_{n\to\infty} \ \widehat{\theta}_{OLS} = \theta \ \frac{\sigma_*^2(1-R^2)}{\sigma_*^2(1-R^2)+\sigma_u^2} \equiv \theta \ \pi \tag{18}$$

where σ_*^2 is the variance of the signal, 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.

This formula stands for three ideas. First, since the reliability ratio, π, is positive but smaller than 1, OLS will be correctly signed, but too small in magnitude, or attenuated. Second, while it is a staple of

[37] That is, the individual filling out the form in 2003 may have thought transit and housing officers were not supposed to be included in the department total. Based on the 2003 internal document (see above), we compute a total of 3,986 officers uniquely assigned to transit or housing, and applying a department-wide adjustment factor of $36,700/26,367 = 1.39$ leads to an estimated 5,548 transit and housing sworn officers in 2003. Adding that figure back in to the UCR figure of 28,614 yields 34,162 officers, which again is in the ballpark of the correct figure.

empirical work to see whether a regression estimate is robust to the inclusion of various control variables, equation (18) indicates that the cure of additional covariates may be worse than the disease of omitted variables bias. Adding more controls increases the R^2, which exacerbates any attenuation bias. This is intuitive, since controls will explain the signal but fail to explain the measurement error. 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. This also implies that when more than one variable is measured with error, the probability limit of OLS may no longer be of the correct sign.

Now return to equation (17) and suppose that X_i is measured without error. It is straightforward to show that under the assumptions given, the coefficient on S_i in a regression of Z_i on S_i and X_i is consistent for π. The slope in panel B of Figure 2 is thus an estimate of π controlling for no covariates. The indirect least squares interpretation of IV then shows that IV is consistent for θ, as we discuss in the next section.[38]

IV. Econometric Approach

The three equation model introduced in Section III.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

$$Y_i = \theta S_i + X_i'\gamma + e_i \tag{19}$$
$$S_i = \pi Z_i + X_i'\phi + \eta_i \tag{20}$$

where Y_i is the year-over-year change in log crime in a given city and year, S_i is the year-over-year difference in observed log police, and X_i is a vector of controls such as the year-over-year change in log revenues per capita, log population, the demographic structure of the population, as well as year effects or state-year effects. In this model, $e_i = \epsilon_i - \theta u_i$, and η_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 (19) and (20) proceeds straightforwardly by IV, and we weight observations by 2010 city population to obtain a police elasticity estimate representative of the typical person.[39] Sufficient conditions for excluding Z_i from equation (19) are

[38]The indirect least squares interpretation of IV is the familiar result that IV is the ratio of two OLS estimates, namely the reduced form and first stage coefficients. An alternative to IV that is suggested in the panel data literature is to take long differences (Griliches and Hausman 1986). This approach assumes that long differences are just as likely to be exogenous as short differences, which is unlikely in this context. In particular, in the medium- to long-term, it is possible that cities may be able to respond to perceptions of lawlessness by adjusing the size of the police force. The scope for this form of endogeneity is likely to be much weaker in a short-run context, which is one rationale for the literature's focus on the short-run police elasticity of crime.

[39]We are aware of the econometric critique of regression weighting (Deaton 1997, Solon, Haider and Wooldridge 2012). See Section VI.B, below, for discussion.

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

where u_i and v_i are the measurement errors from equations (15) and (16), ϵ_i is the structural error term from equation (17), and $\mathbb{C}[\cdot, \cdot]$ is covariance.[40]

Assumptions ($A1$) through ($A3$) assert that measurement errors 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 ($A4$) would justify running a regression of crime growth rates on police growth rates and controls X_i, were police growth rates observed without error.[41]

Under the classical measurement error model, the exact same steps we used to motivate the simultaneous equations model in equations (19) and (20) can be used to motivate a second simultaneous model with the roles of S_i and Z_i reversed and identical parameters in equation (19). This result is apparently new.[42] 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.*

Since, under the classical measurement error model, both IV estimates are consistent for the police elasticity of crime, it is possible to pool the estimates, which increases efficiency. To do so, we stack the orthogonality conditions for the forward and reflected IV programs into the broader set of moments

$$
g_i(\beta) = W_i \begin{pmatrix} Z_i(Y_i - \theta_1 S_i - X_i'\gamma_1) \\ X_i(Y_i - \theta_1 S_i - X_i'\gamma_1) \\ S_i(Y_i - \theta_2 Z_i - X_i'\gamma_2) \\ X_i(Y_i - \theta_2 Z_i - X_i'\gamma_2) \end{pmatrix} \tag{21}
$$

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

[40] Assumptions ($A1$) through ($A4$) together imply that $\mathbb{E}[Z_i \epsilon_i] = \mathbb{E}[Z_i u_i] = 0$, which implies that $\mathbb{E}[Z_i e_i] = 0$, where $\mathbb{E}[\cdot]$ is expectation. Assumptions ($A2$) and ($A4$) imply that $\mathbb{E}[X_i e_i] = 0$. Of course, $\mathbb{E}[(Z_i, X_i')' e_i] = 0$ is the exclusion restriction justifiying the use of IV with Z_i as an excluded instrument and X_i as an included instrument.

[41] As our aim in this paper is to recover estimates of the police elasticity of crime that correct for measurement error, we maintain the assumption of ($A4$), conditional on state-year effects. However, we emphasize that ($A4$) is not guaranteed to hold and particularly will not hold in the context of simultaneity bias.

[42] For example, it is discussed in neither the classic monograph by Fuller (1987) nor in the recent review paper by Chen, Hong and Nekipelov (2011).

system imposing the restriction $\theta_1 = \theta_2 = \theta$.[43] A further benefit of pooling the two IV estimates using GMM is that the standard GMM test of overidentifying restrictions (Hansen's J) then provides a test of the classical measurement error model.

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 noted 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 are an important consideration in our view. We follow an approach suggested by Lubotsky and Wittenberg (2006) and include as controls growth rates in both the UCR and ASG population measures.[44]

V. 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 and contains at least one city in 44 of the U.S. states as well as the District of Columbia.[45] 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 additional information on sworn police officers from the ASG and from the LEMAS data mentioned above. These data series form the core of our analysis, but we also collect auxiliary data on city 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 Internal Revenue Service; and proxies for social disorganization from the Centers for Disease Control and the National Center for Educational 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 online Data Appendix.

[43] If we additionally seek to impose the restriction that $\gamma_1 = \gamma_2 = \gamma$, then redundancy of moments can become an issue. In a working paper version of this paper, we discuss how empirical likelihood (EL, Owen 2001) is a natural solution to this problem. Estimates using EL and two-step GMM differ in at most the third decimal place, and we suppress those results here in the interest of space.

[44] In the interest of simplicity, we refer to this as "controlling for population" throughout the paper. In a working paper version of this paper, we present evidence that this procedure is sufficient to avoid bias from failure to control for city population growth. This additional evidence is based on annual information on city and county births, which are measured with negligible error. The other obvious approach—-instrumenting both police and population growth with the other measure of police and population growth—results in nearly identical results.

[45] Alaska, Idaho, North Dakota, Vermont and Wyoming are unrepresented in our sample. Due to extensive missing data and various data quality issues, our sample of 242 cities excludes approximately 30 cities meeting the population criterion. In addition, there are 10 states for which our sample contains only a single city, which is relevant for understanding parameter estimates that condition on state-year effects.

The UCR crime data we collect are the standard measure used in the empirical literature. These data are collected monthly by the FBI and, following the literature, are aggregated to the annual level in our analysis. 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, to assess the extent of measurement errors in personnel data we augment data from the UCR with data from the employment files of the ASG. The ASG data provide annual employment counts for a large number of municipal functions, including police protection. 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. As with the UCR system, the ASG reports a point-in-time measure of police. For 1960-1995, the ASG reference period is the pay period including October 12, but beginning with 1997 the reference date has been March 12.[46] For selected analyses we also draw upon a third measure of police, as noted. This measure is drawn from two 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 simplyas the LEMAS series, have been collected at regular intervals from 1987-2008.[47]

The measure of city population used in the majority of crime research is from the FBI's Return A

[46] 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. The UCR data provide the number of full-time sworn police officers and the total number of police officers 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. The ASG FTE numbers before 1977 were then multiplied by the estimated proportion. 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.

[47] Data are available for 1987, 1990, 1992, 1993, 1996, 1997, 1999, 2000, 2003, 2004, 2007, and 2008.

file.[48] While this series contains observations for nearly all city-years, it is potentially contaminated by measurement error, particularly in the years leading up to the decennial Census. The population entries are contemporaneous; the FBI does not retroactively correct any of the population figures. As noted, we additionally use the annual population estimate recorded in the ASG. We further smooth both series over time using local linear regression.[49]

In subsequent analyses in the paper, we report the elasticity of cost-weighted crimes with respect to police. Likewise, in Section VIII, we use data on the cost of police and the cost of crime to derive approximate benefit-cost ratios, both nationally and for specific cities. We pause here to describe briefly data on the cost of crime to society, with further detail provided in the online Data Appendix. This index of cost-weighted crimes provides a single aggregate measure of the cost of criminal victimization to society. It is constructed as the weighted sum across crime types, where the weight assigned to a crime is an estimate of that crime's seriousness. Ideally, the weight used would measure 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. Unfortunately, estimates of the *ex ante* cost of crime are either unavailable or implausibly large, except for the crime of murder, where we can take advantage of the rich literature on the value of a statistical life (VSL) (Viscusi and Aldy 2003, Kniesner, Viscusi, Woock and Ziliak 2012).[50] For other crimes, *ex ante* measures are not available in the literature, and we instead 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. The values we employ are those reported in Cohen and Piquero (2009), currently the standard reference on this topic in the literature, in 2010 dollars.

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.

[48]Note that historically the Census released population estimates for cities only in Census years. Even now with the advent of the American Community Survey, estimates are not available for all years and when available are smoothed across survey years.

[49]The working paper version of this paper shows that smoothing these series increases the population elasticity of crime from roughly 0.25 to roughly 1—almost surely closer to the correct parameter.

[50]For murder, we use a VSL of $7,000,000. This is the mean value in 2010 dollars of VSL estimates derived from sixty-four estimates in the extant literature and is typical of the figure used by U.S. government agencies in 2010. For details on these figures, please see the working paper version of this manuscript. As noted above, we prefer to avoid using the cost of crime estimates from Cohen et al. (2004) because they are implausibly large. Since our primary conclusion in this paper is that U.S. cities are underpoliced, using their costs of crime would only amplify our conclusion.

Descriptive statistics are reported for a sample of 10,589 observations, the universe of data for which measures of crime, police and population growth rates are nonmissing. The left-hand panel gives statistics for levels per 100 thousand population and the right-hand panel gives statistics for growth rates.

Several features of the data are worth noting. First, a typical city employs approximately 250 police officers per 100 thousand 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. Similarly, 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. The cost-weighted crime index is a more robust measure, as it gives more weight to more costly crimes, and the three measures most accurately measured in the UCR system are murder, robbery, and motor vehicle theft (Blumstein 2000, Tibbetts 2012). 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.[51]

To assess the extent to which our sample of cities is representative of broader trends in crime and policing in the country, Figure 3 displays long-run trends in crime and police for our sample of 242 cities and for all cities from 1960 to 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.[52] Panel B presents the same time series evidence for property crimes while Panel C presents the time series for total sworn officers. While the levels of crime and police are higher for our sample of large cities than for all cities, the trends are generally similar.

[51]Moreover, in results not shown, the first difference of a log per capita measure exhibits essentially no cross-sectional heterogeneity.

[52]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.

Focusing on the trends among our sample of cities, it is clear that regardless of whether crimes are cost-weighted, the series show a 30 year rise in criminality from 1960 to 1990, followed by a 20 year decline from 1990 to 2010. The magnitude of these swings is spectacular. For violent crime, costs in 2010 dollars per person quadrupled from 1960 to 1990 and then fell by half from 1990 to 2010. For property crime, costs in 2010 dollars per person tripled from 1960 to 1990 and then returned to just above the 1960 cost by 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.

Our sample of cities also closely tracks trends in policing. 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 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 for New York. While the financial crisis has led to declining police forces in most of our cities, most of the decline occurs after 2010.

VI. Results

A. *Main Results*

The central results of this paper are contained in Table 3. This table presents OLS, IV, and GMM estimates of the police elasticity of crime for the seven major index crimes and for the three crime aggregates—violent crimes, property crimes and the cost-weighted crime index. Columns (1)-(4) correspond to OLS 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-year effects. Columns (5)-(8) correspond to IV models, and column (9) corresponds to GMM models.[53,54] In addition to OLS, IV, and GMM estimates of the police

[53] 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 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 and hence conservative.

[54] An alternative to IV and GMM is OLS applied to the subsample excluding 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). Applying OLS to all but those observations (the "restricted sample") improves upon OLS applied to all the data, but is far from the IV estimate. For example, the OLS estimate for the police elasticity of murder is -0.204 in the primary sample and -0.359 in the restricted sample. 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-year effects). If we perform the same analysis with the ASG measure

elasticity, the bottom panel of Table 3 presents first stage coefficients.

Consistent with the scatterplots presented in Figure 2, the first stage coefficients 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 (5) where the ASG measure is the instrument and the UCR measure is the endogenous regressor, we observe that conditional on population growth a 10 percent increase in the ASG measure is associated with only a 1.8 percent increase in the UCR measure. Column (6) shows that this result is robust to the inclusion of state-year effects with the coefficient value falling by roughly 10 percent from 0.18 to 0.16. Turning to columns (7) and (8), 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 (5) and (6). 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 the variance of the predicting variable. As with the forward first stage regressions, results differ only slightly when the state-year effects are added.[55]

The F-statistic on the excluded police measure is reported below the coefficient estimates. Since the F-statistics we report are all above 140, standard asymptotic approximations will be highly accurate in the context of our application. That is, weak instruments are not a concern in this context (for references to this literature, see for example Cruz and Moreira (2005)).

We turn now to our estimated police elasticities, presented in the top panel of Table 3. Beginning in column (1), we see that using the UCR measure of police officers, the OLS estimate of 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). Reflecting the large weight on murder, the cost-weighted crime elasticity is -0.21 indicating that a 10 percent increase in police is associated with a 2 percent decline in the cost of crime to victims.

The OLS estimates in column (2) control for state-year effects rather than just year effects. The explanatory power of the state-year effects indicates substantial unobserved heterogeneity in crime growth rates. For most crime types, models including state-year effects have an R^2 of about 0.60, which is remarkably high for a model specified in growth rates. However, the unobserved heterogeneity in crime growth rates appears to be largely unrelated to changes in police staffing. Comparing the OLS elasticity

as the endogenous regressor, the analogous three estimates are -0.143, -0.171, and -0.572.

[55] First stage results are similar when we condition additionally on a large number of local-level control variables, as in Table 5, below.

estimates in columns (1) and (2) reveals generally minor differences, with the possible exception of murder and motor vehicle theft, which fall by close to one-third. Conditioning on state-year effects, the largest elasticities are for murder (-0.20), robbery (-0.20), and motor vehicle theft (-0.13). Violent, property, and cost-weighted crimes have elasticities of -0.12, -0.06, and -0.14, respectively.

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. The coefficients in columns (3) and (4) are all of the same sign as those in columns (1) and (2). To the best of our knowledge, this is the first time a city-level panel data regression of crime on the ASG measure of police has been run.[56] Reassuringly, the results confirm the sign of the estimates based on the UCR data. While the estimates in columns (3) and (4) are generally smaller in magnitude, this is expected since the ASG measure exhibits more variance than the UCR measure.[57] The greater variance of the ASG measure also explains the somewhat greater precision of the ASG estimates. The OLS police elasticities are largely similar when the full set of state-year effects are included in column (4), with the exception of motor vehicle theft which falls by roughly half. Taken as a whole, OLS estimates of the elasticity of crime with respect to police point to a persistent but modest relationship between changes in police and criminal activity.

We turn now to the IV estimates in columns (5)-(8). These estimates are typically five times larger in magnitude than the OLS estimates. Referring to column (5), the largest elasticities are those for murder (-0.80), motor vehicle theft (-0.59), robbery (-0.46) and burglary (-0.22). Violent, property, and cost-weighted crimes have elasticities of -0.29, -0.15, and -0.61, respectively. 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, and estimates for the crime aggregates are similar.

In columns (6) and (8), we present IV results that condition on state-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 -0.61 and -0.40.[58] With regard to the individual crimes, elasticities are largest for murder (between

[56] 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.

[57] Abstracting from covariates and under the classical measurement error model, the probability limit of the OLS police elasticity based on S_i is given by $\theta\mathbb{V}[S_i^*]/\mathbb{V}[S_i]$ and that based on Z_i is given by $\theta\mathbb{V}[S_i^*]/\mathbb{V}[Z_i]$.

[58] We mentioned above that our estimates are essentially the same whether we include both population growth measures, instrument the one with the other, or instrument the other with the one. For example, the forward IV estimate of the police elasticity of cost-weighted crimes of -0.614 controls for both population measures. If we instead instrument the UCR measures of both police and population with the corresponding ASG measures, we obtain an elasticity estimate of -0.613; reversing the roles of the endogenous regressor and instrument for population yields -0.629. The reflected estimate is similarly robust to how the two population measures are used. The estimate in Table 3 is -0.403. Instrumenting both ASG measures with both UCR measures yields an estimate of -0.403 and reversing the roles of the population measures yields an estimate of -0.410.

-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-year effects, coefficients on motor vehicle theft are approximately 30 to 50 percent smaller with the inclusion of state-year effects, indicating some correlation between police growth rates and unobserved heterogeneity in the growth rate of motor vehicle theft.

In column (9), we present GMM estimates of the police elasticity of crime.[59] These estimates efficiently combine the information from the forward and reflected IV estimates presented in columns (6) and (8). 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.

We also calculate Hansen's J-test of overidentifying restrictions corresponding to the GMM estimates in column (9). This test provides a measure of the discrepancy between the two parameter estimates.[60] Under the null hypothesis of classical measurement error, the test statistic has a χ^2 distribution with 1 degree of freedom. These test statistics are suppressed in the interest of space but are generally quite small and are uniformly below the 95 percent critical value of 3.84. This implies that the diffferences between the forward and reflected IV estimates in columns (5) and (7) and (6) and (8) are consistent with sampling variability.

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

$$\frac{\mathbb{C}[Y_i, Z_i]}{\mathbb{C}[S_i, Z_i]} = \frac{\mathbb{C}[Y_i, S_i]}{\mathbb{C}[Z_i, S_i]} \tag{22}$$

Since the denominators for the two ratios in equation (22) 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

[59] These estimates correspond to the GMM framework in equation (21).

[60] In principle, the test statistic is available using several different estimators. We compute the test statistic using two-step GMM. 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.

hypothesis for Hansen's J as

$$0 \;=\; \frac{\mathbb{C}[Y_i,Z_i]}{\mathbb{C}[S_i,Z_i]}-\frac{\mathbb{C}[Y_i,S_i]}{\mathbb{C}[Z_i,S_i]}=\frac{\mathbb{C}[Y_i,Z_i-S_i]}{\mathbb{C}[S_i,Z_i]} \quad\Longleftrightarrow\quad 0 \;=\; \frac{\mathbb{C}[Z_i-S_i,Y_i]}{\mathbb{V}[Y_i]} \tag{23}$$

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 (this is assumption ($A2$)), and with the structural error term (this 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 restriction that the measurement errors themselves are uncorrelated (this is assumption ($A3$)). This follows from the fact that if ($A1$) and ($A2$) hold, but ($A3$) does not, both IV estimators measure the incorrect quantity of $\mathbb{C}[Y_i,S_i^*]/\left(\mathbb{V}[S_i^*]+\mathbb{C}[u_i,v_i]\right)$.

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 4. Each column of Table 4 pertains to regressions of the difference between the growth rates for two police measures on a particular covariate, controlling for population growth and state-year effects. 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.[61] Each column of Table 4 presents coefficients from a joint regression of the difference

[61] As noted above, the LEMAS data are available in 1987, 1990, 1992, 1993, 1996, 1997, 1999, 2000, 2003, 2004, 2007, and

in police growth rates on the growth rate in each of the seven index crimes (Panel A), the growth rate in the other police measure (Panel B) and the growth rate in each of our two population measures (Panel C).[62] Under the classical measurement error hypothesis, all the coefficients in the table should be zero.

Referring to Panel A, using the full sample in column (1), we find little evidence of a relationship between the difference in police growth rates 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 21 tests of this hypothesis; none provide evidence against the restrictions of the classical measurement error model. As noted, these tests amount to joint tests of Assumptions ($A1$) and ($A2$), because crime growth rates reflect the signal, the covariates, and the structural error.

Panel B of Table 4 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 $\widetilde{Z}_i$). As noted, these tests amount to joint tests of Assumptions ($A2$) and ($A3$). The results in this panel may contain some slight evidence against the classical measurement error model. Specifically, 1 of the 3 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 any covariance seems to be 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 relationship between measurement errors and population growth rates.

Finally, in the bottom panel of Table 4, 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 4 as furnishing little evidence against the assumptions of the classical measurement error model. There are 39 total tests presented in Table 3; only one of these tests rejects at the 5 percent level, and no joint test is significant at the 5 percent level.

2008, so for columns (2)-(4) we are using growth rates taken from 1987 to 1990, from 1990 to 1992, from 1992 to 1993, and so on.

[62] Note that the standard errors on the two population growth rates are quite large due to the high correlation between them. Joint tests of significance are also presented and are not affected by this multicollinearity. Regressions using only one population measure have much smaller standard errors are show similar results.

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 designed to mimic the properties our sample. The study shows that the tests utilized in Table 3 generally have good power. Details of the simulation study are provided in an Appendix.

B. Robustness

Before turning to a discussion of the results presented above, we consider several robustness checks. The estimates in Table 3 assume the exogeneity of police conditional on population growth and state-year effects. While state-year effects soak up important time-varying state-level variation, results will nevertheless be inconsistent if there are time-varying city-level factors which are correlated with both police and crime growth rates. In Table 5, we explore the extent to which the elasticities reported in Table 3 are robust to the inclusion of a variety of city-level covariates. Unfortunately, these covariates are not available for our entire sample period and we thus restrict attention to the 1970-2002 subsample.[63]

We begin in column (1) by replicating the GMM parameters estimates presented in column (9) of Table 2 for the 1970-2002 subsample of our data. These estimates condition on population growth and state-year effects. For the 1970-2002 subsample, the violent crime elasticity is -0.22 and the property crime elasticity is -0.18. The largest elasticities are for murder, robbery, and motor vehicle theft (-0.62, -0.59, and -0.34, respectively). Interestingly, the coefficients on rape and assault for this time period are perversely signed. The police elasticity of cost-weighted crimes is -0.43.

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 4 leading industrial sectors (construction, manufacturing, wholesale trade and retail trade). We also include city public expenditures exclusive of police to capture the impact of other municipal programs which might impinge on crime. 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 squares and cross-products for each of the demographic subgroups. Finally, in column (6) we add city-specific linear time trends to capture long-standing crime trends that are independent of growth in police.

[63]In a working paper version of this paper, we additionally report on the relationship between police growth rates and these city-level covariates for all available years, documenting small associations in each instance.

Looking across the columns of Table 5, it is apparent that the estimated elasticities change very little with the inclusion of the controls. Referring, for example, to the cost-weighted crime index, the estimated elasticity moves from -0.43 when conditioning only on population and state-year effects to -0.41 when economic covariates are included. Conditioning also on the lags of the economic covariates brings the estimated elasticity up to -0.42 while controlling extensively for demographics brings the elasticity back to -0.39. When time trends are included, the elasticity increases to -0.41, just 2.5 percent lower 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 murder, which falls by roughly 5 percent. The results in Table 5 thus imply a relatively minor role for the confounding of police growth rates with unobserved determinants of crime growth rates, at least conditional on population growth rates and state-year effects.[64]

We conducted three final robustness checks that are worth mentioning. First we assess the sensitivity of the estimates reported in column (9) of Table 3 to the exclusion of several theoretically-motivated groups of cities—the two largest cities in the sample (New York and Los Angeles), cities that 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 3. Second, we address the possibility of displacement—the idea that an increase in policing in one jurisdiction might displace crime to a nearby jurisdiction—by aggregating the data to the MSA level. The issue of displacement is potentially important because there is little social value in reshuffling criminal activity. Estimates from the higher level of aggregation are, if anything, somewhat larger than those from the city level analysis, suggesting displacement is an unlikely explanation for our results unless crime is as likely to be displaced between MSAs as it is between cities.[65] Third, following Solon et al. (2012), we examine interactions of the police growth rates with population to assess whether our weighted least squares procedure recovers the average partial effect. These results support the interpretation of our main results as the police elasticity of crime for a typical person in our sample of cities.[66]

[64] In unreported results, we document a similar lack of confounding when state-year effects are replaced by the more parsimonious year effects.

[65] The MSA results are presented in Appendix Table 2 of the working paper version of this paper.

[66] As noted by Solon et al. (2012), 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. We re-estimated the population-weighted OLS estimates in Table 3 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

VII. 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. Our preferred estimates are the GMM estimates from column (9) of Table 3. In this section, we compare our reported elasticities to those in the prior literature.

Table 6 presents selected police elasticities from eight recent papers utilizing U.S. data.[67] Each of these papers explicitly seeks to correct for simultaneity bias, for which our estimates do not adjust. The dominant methodologies are quasi-experimental, but two papers focus on regression strategies with high quality control variables (Marvell and Moody 1996, Corman and Mocan 2005). None of these papers discuss the possibility of measurement errors in police. However, IV and difference-in-difference strategies will correct for both simultaneity bias and measurement error bias under the classical measurement error hypothesis. The two papers using regression strategies may nonetheless suffer from measurement error bias.[68]

An examination of the estimates in these papers reveals four evident tendencies. First, the estimates are generally negative. Some of the estimates are zero—Levitt (1997) for property crime and Klick and Tabarrok (2005) for violent crime—but almost none are positive.

Second, the estimates from the quasi-experimental U.S. literature tend to be similar to, or perhaps slightly larger in magnitude than, our own estimates. It is difficult to know how to interpret this pattern. For example, there is a good deal of sampling variability associated with each estimate, indicating that not too much stock should be placed in any differences. Bracketing the issue of sampling variability, however, one obvious interpretation for the discrepancy is that the papers cited in Table 6 correct for simultaneity bias, whereas our estimates do not. On the other hand, the samples involve quite different time periods. For example, when we restrict our analysis to the years analyzed by the very careful study 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 reported by Evans and Owens (-0.99 for violent crimes and -0.26 for property crimes), suggesting that perhaps the discrepancies between studies have somewhat less to do with method and somewhat more to do with the cities and years being studied.

---

represents the average partial effect. The estimates we obtain are similar to those reported in Table 3. For example, we obtain OLS estimates (standard errors) for the UCR and ASG measure of 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 2 provide little evidence in favor of important unmodeled heterogeneity.

[67] We note that our focus on U.S. data here excludes several interesting and well-executed papers utilizing data from other countries, including Di Tella and Schargrodsky (2004), Draca et al. (2011), Machin and Marie (2011), and Vollaard and Hamed (2012).

[68] On the other hand, Corman and Mocan (2005) use administrative data on the number of police which are presumably free from measurement error.

Third, there is considerable variability in the prior literature with regard to whether police have a larger protective effect on violent crimes than on property crimes. In particular, two recent papers find violent crime elasticities that are larger than property crime elasticities (Levitt 1997, Evans and Owens 2007); two find violent and property crime elasticities that are roughly equal (Levitt 2002, Corman and Mocan 2005); and two find property crime elasticities that are larger than violent crime elasticities (Klick and Tabarrok 2005, Lin 2009).[69] Our estimates are closest to those of Evans and Owens.

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 elasticities (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 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 7, 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.

The cross-crime pattern of the police elasticity estimates could reflect non-classical measurement error,

[69] We note that Corman and Mocan (2005) employ monthly data on New York, whereas the other papers in the table employ annual data on multiple cities. This may imply that the estimand from the Corman and Mocan study differs from that of other studies.

relative deterrence effects, or relative incapacitation effects. The variety of non-classical measurement error that could lead to the cross-crime patterns we observe is simple: if reporting to police is increasing in police, then *measured* crimes could be increasing in police, even if crimes overall are decreasing in police.[70] In a particularly well-known recent episode, the Sacramento police department responded to deep budget cuts by announcing they would no longer respond to "burglary, misdemeanors, or minor traffic accidents" and would only conduct follow-up investigations for murder and sexual assault (Goode 2012). Similar policy changes have been reported for Camden (Goldstein 2011), Chicago (Spielman 2013), Oakland (Preuitt and Sanchez 2010), and Stockton (DuHain 2012). Such policies seem likely to result in a reduced reporting rate. If reporting is affected by changes in police staffing, this should amplify our central policy conclusion. We show below that the police elasticity of *reported* crime is negative and large enough in magnitude to suggest underpolicing. Correcting for reporting bias would lead to police elasticities that are more negative than those we document and strengthen this conclusion.

The cross-crime pattern of police elasticities may also reflect differences in deterrence and incapacitation effects across crimes. 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 or prison time for the convicted (McCrary 2009). It is clear that deterrence effects could differ across crime types. To see why incapacitation effects could lead to differences across crimes in elasticity estimates, note that in an efficient criminal justice system the most serious recidivists are those most likely to be placed in pre-trial detention and are most likely to be incarcerated if convicted. To the extent that police focus on the population of serious recidivists, we might indeed expect strong cross-crime differences.

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. If simultaneity bias is a lingering issue, our approach underestimates the magnitude of the policing elasticity (Nagin 1978, Nagin 1998, Levitt and Miles 2007).

[70] Levitt (1998a) finds little evidence for this effect in U.S. data, but Vollaard and Hamed (2012) do in British data.

VIII. Social Welfare Analysis

A. Simple Estimates of the BCR

The results presented in Table 3 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. Likewise, the benefit-cost ratio provides intuition regarding whether U.S. cities are optimally policed. In Section II, we established that even with heterogeneous individuals making precautionary investments with externalities where policing crowds out private precautions, it is possible to make strong statements about welfare using the reduced form police elasticity of crime. In particular, the rule-of-thumb outlined in Section II is that hiring police improves welfare when

$$|\overline{\theta}|/\frac{wS}{nC} \equiv \overline{\kappa} > 1 \tag{24}$$

In this section, we use the GMM approach used to estimate the elasticity of crime with respect to police to estimate the ratio of the benefits (as proxied by averted costs to potential victims) to the costs of police.

Using a VSL of \$7 million as the cost of a murder, we estimate a police elasticity of the cost of crime of -0.47±0.34. This elasticity estimate is based on a model including state-by-year effects and controls for population (Table 3, column (9)). The cost-weighted elasticity is powerfully affected by the assumed cost of murder. For example, varying the VSL from \$1 to \$28 million, our GMM estimate of the police elasticity of the cost of crime ranges from -0.32±0.18 to -0.55±0.52.

Scaling the cost-weighted elasticity by the ratio of the expected cost of crime to the cost of police produces an estimate of the 2010 social dollars saved from increasing spending on police by one dollar, or the benefit-cost ratio (BCR). As with the cost-weighted elasticity, the BCR is sensitive to the monetized value of an averted murder.[71] Figure 4 demonstrates this visually, plotting the BCR on the vertical axis as a function of possible VSL values on the horizontal axis. The BCR ranges from approximately 0.4 at a VSL of \$1 million to approximately 6.0 at a VSL of \$28 million.

To give some sense of what is a reasonable value for the VSL, we superimpose a kernel density estimate of the density of 64 VSL estimates drawn from the recent literature or currently in use by the federal government.[72]

[71] 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.

[72] See the working paper version of this manuscript for greater detail.

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 would save $1.63 in costs to crime victims. This would be consistent with classical notions of the underprovision of public goods. 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 overpolicing.[73]

B. Police Incapacitation Effects

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 analysis, treating the expected short-term 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 note that the long run effect on incarceration would be smaller than the effect in the first year, as the prison population would be smaller due to the lower prevalence of crime (McCrary 2009, equation (34)).

We begin with an estimate of the number of arrests per officer. Using our sample of 242 cities, an officer in the typical city made 19.45 arrests in 2010.[74] Next, we employ an estimate of the conditional probability of a conviction given an arrest. In 2006, the most recent year available for convictions, there were 14,380,370 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.[75] Dividing convictions by arrests yields an estimated conditional probability of a conviction of 8.4 percent. This implies about 1.64 convictions per police officers.

Of defendants sentenced in state courts, 40 percent were sentenced to state prison (with a mean sentence length of 4.92 years) 28 percent were sentenced to a term in local jail (with a mean sentence of 0.5 years) and the remaining 32 percent were sentenced to a term of probation or an alternate penalty that did not involve incarceration. This implies an expected sentence length given conviction of 2.11 years. Using the National

[73] We note that estimates pertaining to individuals' labor market behavior tend to yield larger VSL values (on average, $9.5 million), while estimates from non-labor market behaviors tend to yield much smaller VSL values (on average, $4 million).

[74] 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. We obtain 18.7 (20.2) arrests using the UCR (ASG) officer count, which we average to obtain 19.45.

[75] The national arrest figure is from FBI (2006), and the conviction and court processing figures here and in the remainder of this passage are from BJS (2009).

Corrections Reporting Program data for 2006, we estimate that for the seven index offenses, individuals serve 47.5 percent of their nominal sentence, for an effective sentence per conviction of 0.98 years.[76] Thus, a typical officer is associated with 19.45 arrests, 1.64 convictions, and 0.98 incarceration-years.[77] At an incarceration cost of $33,089 per year, each new officer is thus associated with $32,344 in additional incarceration costs.[78] Augmenting the salary figure with this estimate yields a benefit-cost estimate of $1.31 using the $7 million estimate of the value of a statistical life.

IX. Conclusion

Federal and state governments routinely decide whether to make safety investments (e.g., guardrails or rumble strips on highways) with the intent of reducing fatalities. In deciding which investments to make, agencies justify investments by appealing to reduced fatality risk and quantify the tradeoff between dollars and risk by, for example, being willing to pay $7,000 per capita for a reduction of 1/1,000 in the fatality risk and $70,000 per capita for a reduction of 1/100 in the fatality risk. This is fundamentally the same as the decision a city confronts in determining how much to invest in police and suggests, as we have argued, that the appropriate figure for the elimination of a single murder is $7 million, which is vastly more than that of even the most expensive property crime, motor vehicle theft. Even factoring in the greater prevalence of motor vehicle theft relative to murder, the expected cost of murder in 2010 was 27 times that of motor vehicle theft. As a consequence of these stark discrepancies in costs, the optimal number of police depends critically on how effective they are at reducing murder.

However, the recent crime literature has focused more on establishing that police reduce crime generally, rather than the extent to which police reduce violent crimes, or specific crimes such as murder. While these papers exhibit extraordinary creativity in grapping with the simultaneity bias problem, there nonetheless remains substantial ambiguity about the effect of police on the most costly crimes. As noted above, some studies find larger effects for violent crime than for property crime, some studies find the opposite, and some studies find similar sized effects for violent and property crimes.

In addition to the differing conclusions of different studies, there is ambiguity regarding the findings of each individual quasi-experimental study due to simple parameter uncertainty, and this is particularly true

[76] See `http://www.bjs.gov/index.cfm?ty=pbdetail&iid=2045`, particularly Table 0610.

[77] Note that we ignore the possible role of pre-trial detention, since time served is typically taken off of any sentence received.

[78] This figure is based on the ratio of state corrections expenditures in 2007 to the number of prisoners, adjusted to 2010 dollars. See `http://www.bjs.gov/index.cfm?ty=pbdetail&iid=4332`.

for murder. Almost none of the papers in the prior literature documents an effect of police on murder that is distinguishable statistically from no effect at all. However, several contain suggestive evidence that police may have quite large effects on violent crimes and on murder in particular. The imprecision of the findings in this literature arises due to two basic factors. First, by design and as is widely appreciated, quasi-experimental approaches ignore most of the variation in police growth rates. Second, and compounding this challenge, the most costly crimes are also quite rare, and rare events are difficult to model econometrically.[79]

We have argued that the central empirical challenge for the literature is not simultaneity bias, as has been supposed in the recent literature, but instead measurement error bias. This claim has two components. First, as we show empirically, police elasticities of crime are affected in minor ways by including the controls routinely cited in the quasi-experimental literature as sources of potential bias: demographic factors, the local economy, city budgets, social disorganization, and any possible state-level policy changes that have the same effect across cities within the state. While a fully optimizing city would hire police in anticipation of crime waves, cities operate under a large number of informational, institutional, and labor market constraints that prevent these types of adjustments from taking place in the short run.[80]

The second component to this claim is that there is in fact a high degree of measurement error in the basic data set on police used in the U.S., the Uniform Crime Reports (UCR). Using a second measure of police from the Annual Survey of Government (ASG) in conjunction with the classical measurement error model, we estimate that measurement errors have attenuated prior panel data estimates of the police elasticity of crime by a factor of four or five. Measurement error bias is thus an alternative explanation to simultaneity bias for the small police elasticity magnitudes that motivated the quasi-experimental literature in the first place.

Combining the information in the UCR and the ASG, we present a series of measurement error corrected estimates of the police elasticity of crime. While the ASG data have similar coverage to the UCR data, they have not been used in panel data regressions on city crime and police rates prior to this work. Our preferred estimates point to powerful effects of police on crime, including violent crime categories such as murder. These estimated effects maintain the hypotheses of the classical measurement error model.

[79] To appreciate the difficulty, consider that the literature generally analyzes city crime and policing rates in growth rates rather than levels or logs to eliminate time invariant differences across cities. In a typical murder regression, the standard deviation of residualized growth rates in murder is roughly 8 times that of police, whereas in a standard Mincer regression in 2010, the standard deviation of residualized log earnings is about 1.2 times that of residualized education. This implies that to obtain the same level of precision as in a Mincer regression, a crime regression for murder would require roughly 6.7 times as much data—when in fact nearly every crime data set has fewer than 20 thousand observations, i.e., far fewer observations than are available for a standard Mincer regression.

[80] The working paper version of this manuscript provides some detail regarding these constraints.

In a methodological contribution, we show how the fundamental symmetry of the classical measurement error model implies that there are two consistent estimators for the parameter of interest, rather than one, indicating a role for a more efficient generalized method of moments (GMM) procedure. The GMM framework then also suggests a series of specification tests which, as we demonstrate, have good power to reject violations of the classical measurement error model. We find little evidence against the classical measurement error model in these data, ratifying the validity of these methods.

Applying cost-benefit analysis to our estimated police elasticities of crime suggests that U.S. cities were underpoliced as of 2010, with an estimated return of $1.63 to every dollar invested in policing. The conclusion that U.S. cities are underpoliced holds in a broad range of circumstances and indeed is valid unless (1) public policing crowds out private investments in precaution, (2) private precautions have *positive* externalities on average, *and* (3) the externality effect is fully 39 percent as big as the direct effect of police. If instead private precautions have beggar-thy-neighbor effects on average, the return to a dollar invested in policing *exceeds* $1.63.

Several considerations suggest that this estimate of the social return to investing in police is, in fact, conservative. The cost of crime that we have used in quantifying the return to investing in police is limited to what a victim would be willing to pay to obtain a reduced probability of victimization for murder, rape, robbery, assault, burglary, larceny, and motor vehicle theft. While these are the only police outputs that are reliably estimated for our sample of cities over the sample period, these are not the only crimes police are tasked with preventing and solving. Presumably police activities in these other arenas (e.g., domestic violence or driving under the influence) are also socially beneficial. Moreover, crime has an extraordinary ripple effect on economic life, leading to population decline locally (Cullen and Levitt 1999) and to expensive population and economic reorganization within space (Rosenthal and Ross 2010) as individuals relocate activities to keep safe. These considerations underscore the primary conclusion of this paper that U.S. cities are underpoliced.

This conclusion is particularly compelling in the wake of the financial crisis, as U.S. cities grapple with increasingly tight budgets. From 2008 to 2011, nearly three quarters of our sample of 242 cities saw a shrinking number of officers per capita. Half saw reductions of 5 percent or more, a third saw reductions of 10 percent or more, and a tenth saw reductions of 20 percent or more. Our analysis suggests that while these staffing reductions certainly save money in a budgetary sense, they are more costly than they might seem.

Appendix: Power of Tests

In this Appendix, we report on the results of a small simulation study designed to approximate the power of the tests presented in Table 4. We generate simulated data $(Y_i, S_i, Z_i, \widetilde{Z}_i)$ as

$$Y_i = \theta S_i^* + \epsilon_i \tag{A.1}$$
$$S_i = \lambda_1 S_i^* + u_i \tag{A.2}$$
$$Z_i = \lambda_2 S_i^* + v_i \tag{A.3}$$
$$\widetilde{Z}_i = S_i^* + \widetilde{v}_i \tag{A.4}$$

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.[81] In the simulations, we allow 4 parameters of the data generating process (DGP) to vary: ρ_1, λ_1, λ_2, 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 4 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 $\lambda_1 = \lambda_2 = 1$, 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 and λ_2 index the extent to which Assumption ($A2$) is violated; and ρ_3 indexes the extent to which Assumption ($A3$) is violated.[82] We maintain Assumption ($A4$) throughout. For each of 10,000 simulated data sets, we construct the tests performed in Table 4 and record whether the null hypothesis was rejected.[83] This allows us to examine the power of these tests against specific alternatives.

The results of this analysis are shown in Appendix Figure 1, 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 4, 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 4, Panel B). The four panels in the figure 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 4, 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

[81] We set $\sigma_* = 0.044$, $\sigma_\epsilon = 0.260$, $\sigma_u = 0.047$, and $\sigma_v = 0.070$, and $\sigma_{\widetilde{v}} = 0.055$. These values roughly match the root mean square error from IV models for the cost-weighted sum of crimes, the first stage coefficients for the state-year effects, and the standard deviation of the various police measures after demeaning by state-year.

[82] 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 $\widetilde{Z}_i$. This is the standard form in the literature for discussing violations of ($A2$) (see, for example, Kim and Solon 2005).

[83] To match our tests from Table 4, 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 $n = 1,752$.

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 4, 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).

Of course, as with any specification test, there will be a lack of power in specific directions. We have examined the power to 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 the importance of validation studies in this area.

References

Ashenfelter, Orley 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.

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.

BJS, "Felony Sentences in State Courts, 2006," *Statistical Tables*, December 2009, `http://bjs.gov/content/pub/pdf/fssc06st.pdf` Last accessed May 26, 2013.

Blumstein, Alfred, "Disaggregating the Violence Trends," in Afred Blumstein and Joel Wallman, eds., *The Crime Drop in America*, New York: Cambridge University Press, 2000, pp. 13–44.

Bound, John, 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.

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, A. Colin and Pravin K. Trivedi, *Microeconometrics*, New York: Cambridge University Press, 2005.

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.

Chen, Xiaohong, Han Hong, and Denis Nekipelov, "Nonlinear Models of Measurement Errors," *Journal of Economic Literature*, December 2011, *49* (4), 901–937.

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

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

——, **Roland T. Rust, Sara Steen, and Simon T. Tidd**, "Willingness-to-Pay for Crime Control Programs," *Criminology*, March 2004, *42* (1), 89–109.

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

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.

Cruz, Luiz and Marcelo J. Moreira, "On the Validity of Econometric Techniques with Weak Instruments: Inference on Returns to Education Using Compulsory School Attendance Laws," *Journal of Human Resources*, Spring 2005, *40* (2), 393–410.

Cullen, Julie Berry and Steven D. Levitt, "Crime, Urban Flight, and the Consequences for Cities," *Review of Economics and Statistics*, May 1999, *81* (2), 159–169.

Deaton, Angus, *The Analysis of Household Surveys : A Microeconomic Approach to Development Policy*, Washington, D.C.: World Bank, 1997.

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.

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

DuHain, Tom, "Stockton Police to Focus on Violent Crime," *KCRA*, June 1, 2012, `http://www.kcra.com/Stockton-police-to-focus-on-violent-crime/-/11798090/14419974/-/2mcw07/-/index.html` Last accessed May 26, 2013.

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.

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

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

FBI, "Arrests," *Crime in the United States*, 2006, `http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2006` Last accessed May 26, 2013.

Fudenberg, Drew and Jean Tirole, *Game Theory*, Cambridge: MIT Press, 1991.

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

Goldstein, Joseph, "Police Force Nearly Halved, Camden Feels Impact," *New York Times*, March 6, 2011.

Goode, Erica, "Crimes Increases in Sacramento After Deep Cuts to Police Force," *New York Times*, November 3, 2012.

Griliches, Zvi and Jerry A. Hausman, "Errors in Variables in Panel Data," *Journal of Econometrics*, 1986, *31* (1), 93–118.

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.

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.

King, William R., Abdullah Cihan, and Justin A. Heinonen, "The Reliability of Police Employee Counts: Comparing FBI and ICMA Data, 1954-2008," *Journal of Criminal Justice*, August 2011, *39*, 445–451.

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.

Kniesner, Thomas J., W. Kip Viscusi, Christopher Woock, and James P. Ziliak, "The Value of a Statistical Life: Evidence from Panel Data," *Review of Economics and Statistics*, February 2012, *94* (1), 74–87.

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.

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

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

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

____ **and** ____ , "Empirical Study of Criminal Punishment," in A. Mitchell Polinsky and Steve Shavell, eds., *Handbook of Law and Economics*, Vol. 1, North-Holland, 2007, chapter 7, pp. 455–498.

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

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*, August 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.

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.

____ , "Dynamic Perspectives on Crime," in Bruce Benson and Paul R. Zimmerman, eds., *Handbook of the Economics of Crime*, Northampton, MA: Edward Elgar, 2009.

Mosher, Clayton J., Terance D. Miethe, and Timothy C. Hart, *The Mismeasure of Crime*, 2nd ed., Thousand Oaks, California: Sage Publications, 2011.

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.

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

Owen, Art B., *Empirical Likelihood*, New York: Chapman and Hall/CRC, 2001.

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.

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

Preuitt, Lori and Kris Sanchez, "Suffer These Crimes in Oakland? Don't Call the Cops," *NBC Bay Area*, July 13, 2010, `http://www.nbcbayarea.com/news/local/Suffer-These-Crimes-in-Oakland-Dont-Call-the-Cops-98266509.html` Last accessed May 26, 2013.

Rosenthal, Stuart S. and Amanda Ross, "Violent Crime, Entrepreneurship, and Cities," *Journal of Urban Economics*, January 2010, *67* (1).

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.

Spielman, Fran, "City Implements 911 Dispatch Changes Freeing Up Officers for Response," *Chicago Sun-Times*, February 2, 2013.

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

Tibbetts, Stephen G., *Criminological Theory: The Essentials*, Thousand Oaks: Sage Publications, 2012.

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

Viscusi, W. Kip and Joseph E. Aldy, "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World," *Journal of Risk and Uncertainty*, 2003, *27* (1), 5–76.

Vollaard, Ben and Joseph Hamed, "Why the Police Have an Effect on Violent Crime After All: Evidence from the British Crime Survey," *The Journal of Law and Economics*, Forthcoming 2012.

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, 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.

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

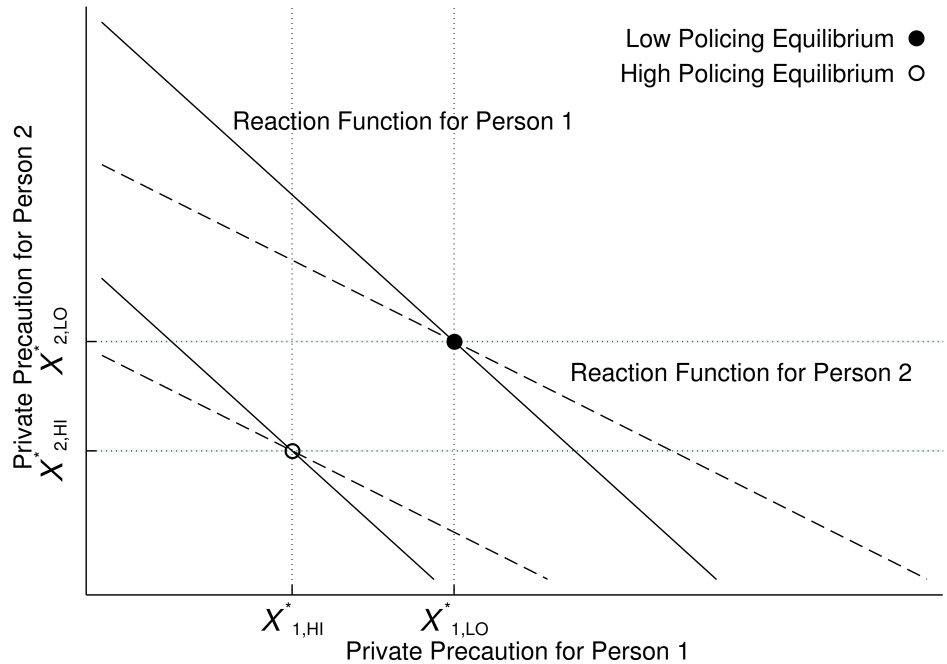
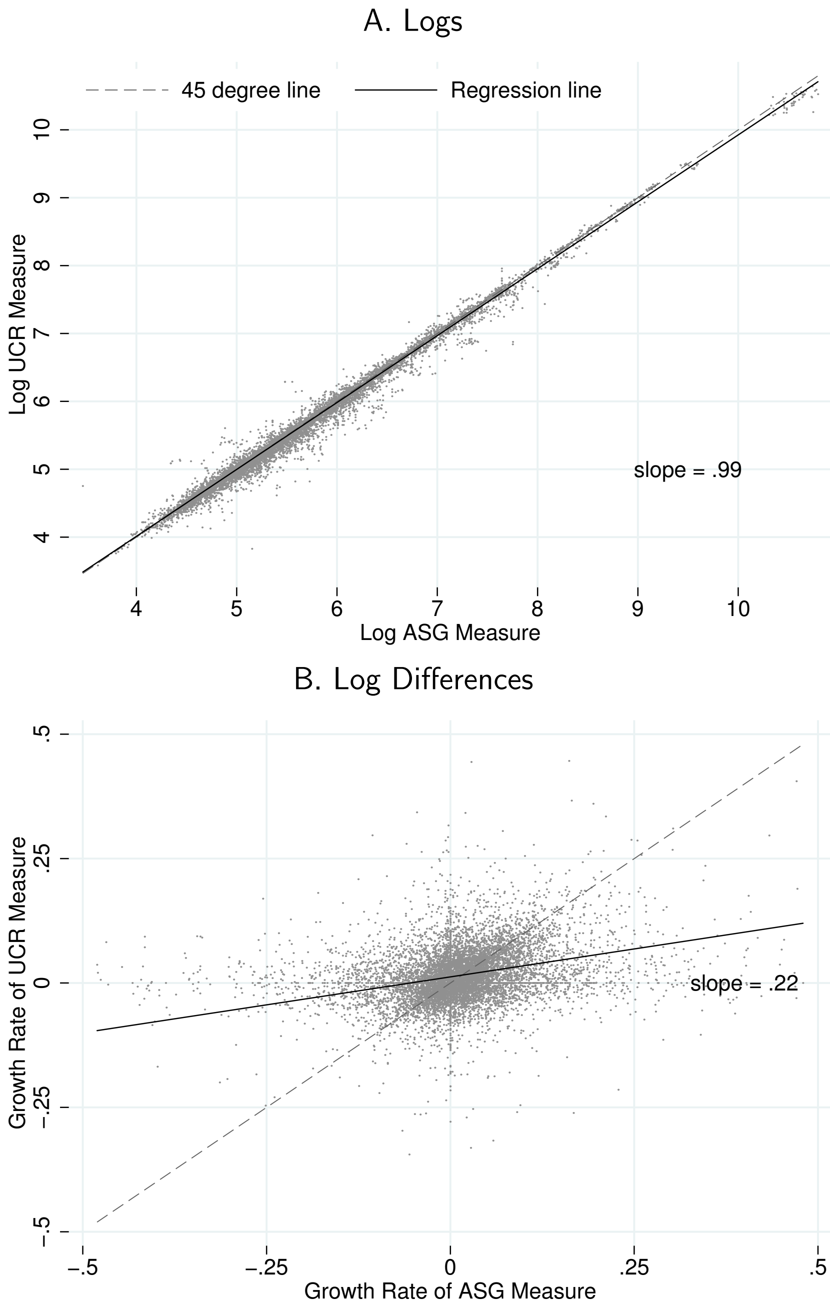
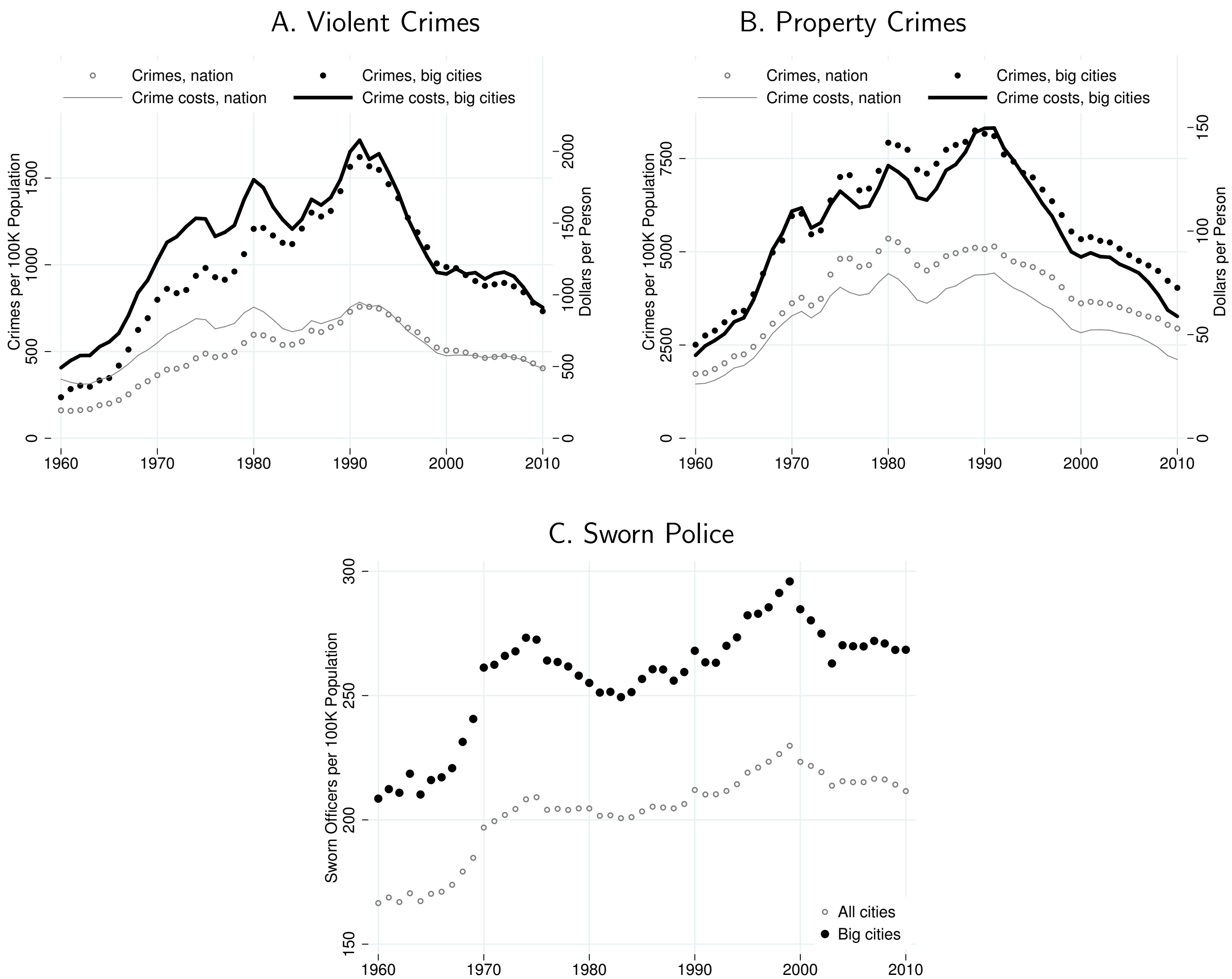


Figure 2. Two Leading Measures of Sworn Officers: the Uniform Crime Reports and the Annual Survey of Government



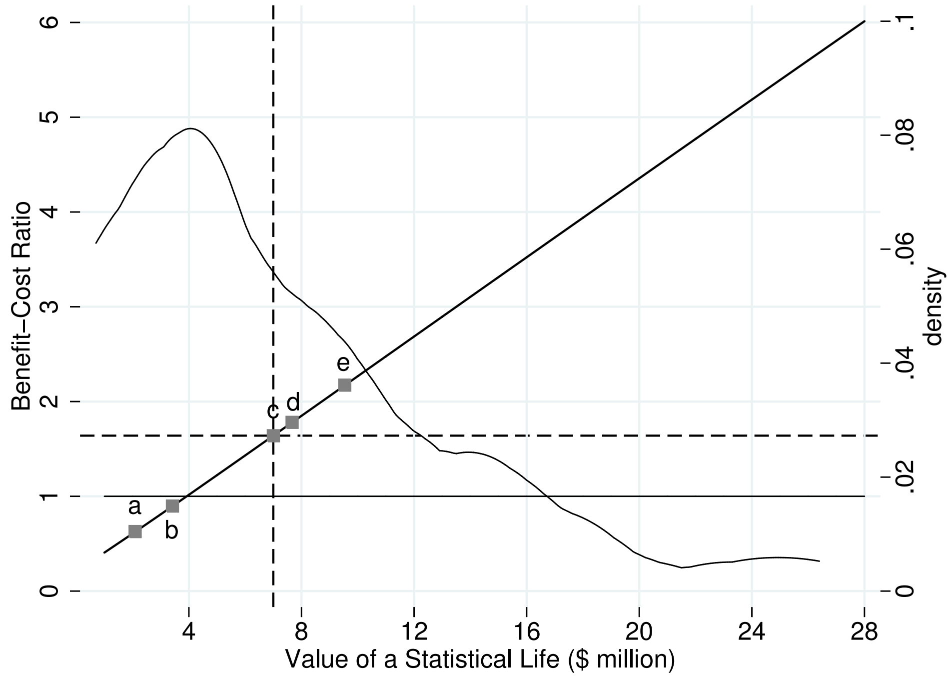
Note: Panel A plots the relationship between the UCR and ASG measures of police in logs. Panel B plots the relationship in log differences ("growth rates"). For ease of visual comparison, in Panel B, we have focused on data points for which the growth rates are smaller than 50 percent in magnitude. The vast majority (99.9 percent) of the data are in this space. The regression slope (0.22) is drawn through the entirety of the data. See text and online Data Appendix for details.

Figure 3. Aggregate Trends in Violent and Property Crime and Police: Evidence from the Uniform Crime Reports



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. In all panels, "big cities" refer to the sample of 242 cities analyzed in the paper. See text and online Data Appendix for details.

Figure 4. 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 (2009). 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 corresponds 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 among all studies in the literature ("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. Costs of Police and Crime

	Cost per Officer	Officers per 100K Population	Annual Cost per Capita
Sworn police	\$130,000	262.7	\$341

	Cost per Crime	Crimes per 100K Population	Annual Expected Cost per Capita
Murder	\$7,000,000	9.9	\$693
Rape	\$142,020	30.9	\$44
Robbery	\$12,624	286.4	\$36
Assault	\$38,924	418.9	\$163
Burglary	\$2,104	976.2	\$21
Larceny	\$473	2,623.3	\$12
Motor vehicle theft	\$5,786	454.3	\$26
		Grand Total:	\$995
		Income per capita:	\$26,267

Note: Numbers pertain to a sample of 242 large U.S. cities in 2010, which have a collective population of 73,820,297. Data on crimes from the Uniform Crime Reports. Data on income per capita are taken from the American Communities Survey five-year estimates (2007-2011). Data on costs of police and crime taken from the literature. See text for details.

Table 2. Summary Statistics on Police and Crime

		Levels				Log Differences			
Variable		**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. Estimates of the Effect of Police on Crime

	Least Squares Estimates				2SLS Estimates				GMM Estimates
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	UCR Measure		ASG Measure		*Forward Models* UCR Measure		*Reflected Models* ASG Measure		*Pooled Models*
Violent crimes	-0.117	-0.120	-0.053	-0.058	-0.289	-0.361	-0.321	-0.336	-0.344
	(0.037)	(0.040)	(0.024)	(0.023)	(0.128)	(0.143)	(0.100)	(0.106)	(0.096)
Murder	-0.270	-0.204	-0.148	-0.143	-0.804	-0.889	-0.742	-0.572	-0.666
	(0.071)	(0.097)	(0.047)	(0.059)	(0.260)	(0.364)	(0.198)	(0.262)	(0.238)
Rape	-0.066	-0.074	-0.038	-0.054	-0.208	-0.339	-0.181	-0.208	-0.255
	(0.069)	(0.092)	(0.043)	(0.050)	(0.234)	(0.301)	(0.188)	(0.248)	(0.219)
Robbery	-0.180	-0.204	-0.085	-0.084	-0.459	-0.521	-0.493	-0.572	-0.559
	(0.048)	(0.047)	(0.032)	(0.029)	(0.176)	(0.177)	(0.128)	(0.125)	(0.117)
Assault	-0.052	-0.037	-0.010	-0.013	-0.052	-0.079	-0.143	-0.104	-0.099
	(0.044)	(0.050)	(0.030)	(0.035)	(0.164)	(0.209)	(0.120)	(0.136)	(0.127)
Property crimes	-0.071	-0.059	-0.028	-0.030	-0.152	-0.189	-0.195	-0.167	-0.174
	(0.028)	(0.026)	(0.020)	(0.015)	(0.109)	(0.090)	(0.077)	(0.068)	(0.062)
Burglary	-0.061	-0.062	-0.041	-0.054	-0.222	-0.339	-0.166	-0.174	-0.225
	(0.043)	(0.037)	(0.027)	(0.021)	(0.144)	(0.128)	(0.118)	(0.098)	(0.089)
Larceny	-0.038	-0.025	-0.002	-0.018	-0.012	-0.113	-0.103	-0.070	-0.083
	(0.031)	(0.027)	(0.021)	(0.017)	(0.115)	(0.103)	(0.085)	(0.074)	(0.067)
Motor vehicle theft	-0.187	-0.131	-0.109	-0.047	-0.592	-0.292	-0.514	-0.367	-0.343
	(0.049)	(0.043)	(0.031)	(0.025)	(0.169)	(0.151)	(0.130)	(0.115)	(0.101)
Cost-Weighted crime	-0.213	-0.144	-0.112	-0.099	-0.605	-0.614	-0.583	-0.403	-0.473
	(0.054)	(0.071)	(0.034)	(0.041)	(0.184)	(0.250)	(0.147)	(0.192)	(0.171)
Instrument:	—	—	—	—	ASG		UCR		
First stage coefficient	—	—	—	—	0.184	0.161	0.364	0.356	—
					(0.014)	(0.013)	(0.029)	(0.029)	—
F-statistic on excluded instrument	—	—	—	—	169.1	144.7	154.2	146.4	—
year effects	yes	—	yes	—	yes	—	yes	—	—
state-year effects	no	yes	no	yes	no	yes	no	yes	yes

Note: Columns (1)-(4) report 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 ten crime rates on the first lag of the growth rate in the number of per capita sworn police officers. Finally, column (9) 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. Below estimates of the effect of police on crime, we report coefficient estimates and standard errors from 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. We also report the F-statistic on the excluded police measure. 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 4. Further Tests of the Classical Measurement Error Model

	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 5. Robustness of Results to the Inclusion of Covariates: 1970-2002 Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Violent crimes	-0.221	-0.216	-0.216	-0.199	-0.208	-0.225
	(0.093)	(0.094)	(0.094)	(0.094)	(0.096)	(0.097)
Murder	-0.617	-0.570	-0.588	-0.562	-0.565	-0.583
	(0.238)	(0.240)	(0.240)	(0.241)	(0.246)	(0.248)
Rape	0.020	0.025	0.013	0.035	0.060	0.041
	(0.168)	(0.170)	(0.170)	(0.172)	(0.176)	(0.177)
Robbery	-0.595	-0.590	-0.588	-0.589	-0.594	-0.607
	(0.114)	(0.115)	(0.114)	(0.115)	(0.117)	(0.118)
Assault	0.149	0.144	0.148	0.180	0.162	0.151
	(0.122)	(0.123)	(0.125)	(0.126)	(0.127)	(0.128)
Property crimes	-0.184	-0.190	-0.187	-0.180	-0.185	-0.198
	(0.070)	(0.070)	(0.070)	(0.070)	(0.071)	(0.070)
Burglary	-0.224	-0.223	-0.219	-0.212	-0.202	-0.213
	(0.096)	(0.097)	(0.098)	(0.100)	(0.102)	(0.100)
Larceny	-0.092	-0.102	-0.099	-0.095	-0.109	-0.116
	(0.078)	(0.077)	(0.077)	(0.077)	(0.078)	(0.077)
Motor vehicle	-0.342	-0.325	-0.339	-0.325	-0.329	-0.345
	(0.115)	(0.114)	(0.112)	(0.112)	(0.114)	(0.116)
Cost-Weighted Crimes	-0.434	-0.408	-0.415	-0.389	-0.392	-0.411
	(0.159)	(0.160)	(0.160)	(0.161)	(0.164)	(0.165)
state-by-year effects	yes	yes	yes	yes	yes	yes
economic covariates	no	yes	yes	yes	yes	yes
lagged economic covariates	no	no	yes	yes	yes	yes
demographic variables	no	no	no	yes	yes	yes
polynomials and interactions	no	no	no	no	yes	yes
linear time trends	no	no	no	no	no	yes

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, unrestricted state-by-year dummies and various sets of control variables for the 1970-2002 subsample of our data. The first column reports GMM parameter estimates, conditional on both the UCR and the ASG measures of the growth rate in the city's population and unrestricted state-by-year dummies. The second column adds a series 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 6. 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	OLS: lags as control variables		-0.24	-0.22		-0.15	-0.30
					(0.09)	(0.06)		(0.04)	(0.07)
Levitt (1997)	1970-1992	59 cities	IV: 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	IV: 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	IV: 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	OLS: 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	DiD: high terrorism alert days	0.00				-0.30	-0.84
				(na)				(0.18)	(0.25)
Evans and Owens (2007)	1990-2001	2,074 cities	IV: 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	IV: 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)

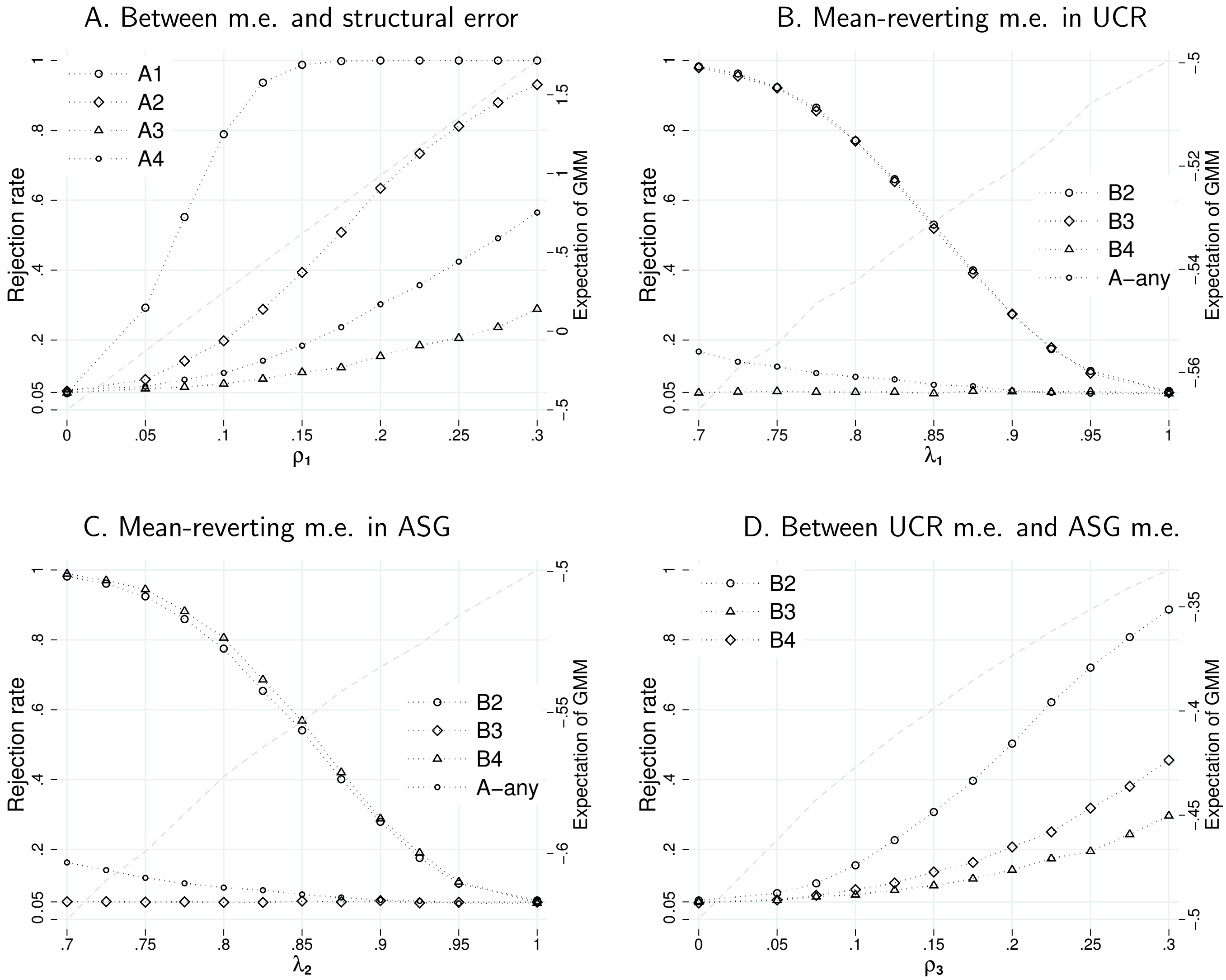
Note: Table reports implied elasticities and standard errors from recent articles employing a novel identification strategy to estimate the effect of police on crime. Under research design, studies are classified as either using least squares ("OLS"), instrumental variables ("IV") or differences-in-differences ("DiD") to identify 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 7. Tests of the Equality of Cross-Crime Elasticities

Type	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft	Violent Crimes	Property Crimes
Murder	-	0.213	0.649	0.036	0.058	0.015	0.181	-	.035
Rape	-	-	0.181	0.485	0.917	0.452	0.689	-	0.731
Robbery	-	-	-	0.002	0.008	0.001	0.120	-	0.001
Assault	-	-	-	-	0.382	0.922	0.114	-	0.554
Burglary	-	-	-	-	-	0.109	0.287	0.295	-
Larceny	-	-	-	-	-	-	0.010	0.010	-
Motor vehicle theft	-	-	-	-	-	-	-	0.997	-
Violent crimes	-	-	-	-	-	-	-	-	0.075

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 3 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.

Appendix Figure 1. 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 4. 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 4. 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 4. There are four such tests examined corresponding to the columns of Table 4. For example, "A1" corresponds to the t-ratio tests of Panel A of Table 4 for the first column, i.e., the full sample, whereas "B2" corresponds to the t-ratio tests of Panel B of Table 4 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.