Why Is There More Crime in Cities?

Edward L. Glaeser

Harvard University and National Bureau of Economic Research

Bruce Sacerdote

Dartmouth College

Crime rates are much higher in big cities than in either small cities or rural areas. This paper explains this connection by using victimization data, evidence from the NLSY on criminal behavior, and the Uniform Crime Reports. Higher pecuniary benefits for crime in large cities can explain at most one-quarter of the connection between city size and crime rates. Lower probabilities of arrest and a lower probability of recognition are features of urban life, but these factors seem to explain at most one-fifth of the urban crime effect. Between one-third and one-half of the urban effect on crime can be explained by the presence of more female-headed households in cities.

I. Introduction

According to the 1994 *Statistical Abstract of the United States,* metropolitan areas have 79 percent more violent crimes than other American cities and 300 percent more violence than rural areas. New York and Los Angeles have crime rates that are approximately four times higher than the crime rates of metropolitan areas as a whole and have violent crime rates that are more than 2.5 times the violent

This paper was written for the Fraser Institute conference for Gary Becker's 65th birthday. We both thank the National Science Foundation for financial support. Sacerdote thanks the Jacob K. Javits Fellowship program. David Cutler, Benjamin Friedman, Claudia Goldin, Larry Katz, Steven Levitt, Ann Piehl, and Richard Posner provided helpful comments.

[*Journal of Political Economy*, 1999, vol. 107, no. 6, pt. 2]

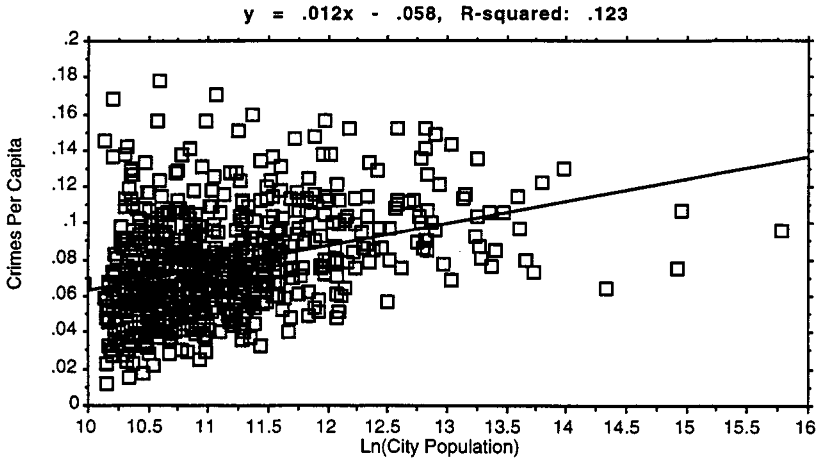

FIG. 1.—Crime and city population: relationship between crime and city population taken from the 1982 Uniform Crime Reports and the 1980 census. *t*-statistic is 9.44.

crime rates of all metropolitan areas. Figures 1 and 2 show the positive correlation between city size and crime rates per capita (fig. 1) and murders per capita (fig. 2). Victimization results from 1989 show that the probability that an individual has been victimized (i.e., has had any crime perpetrated against him or her over a six-month period) is 21.7 percent if that individual lives in a city of more than 1 million people. The comparable figure for cities with populations between 1,000 and 10,000 is 9.4 percent.[1] This paper asks why crime rates are so much higher in cities.[2]

The connection between crime and city size is not a new fact. Criminologists have discussed the urban tendency toward crime for decades (see, e.g., Flango and Sherbenou [1976]; Schichor, Decker, and O'Brien [1979]; Larson [1984]; or two separate articles in Radzinowicz and Wolfgang [1977]). Wirth (1938) discusses the observed connection between crime and urbanization and argues that this connection is evidence for his theory of "urbanism as a way of life." Social observers (such as Thomas Jefferson and Jean-Jacques Rousseau) have long argued that there exists a connection between cities and immoral behavior. Lane (1979) documents that in the nine-

[1] These results are calculated from the National Crime Victimization data.

[2] Urban density can occasionally lead to safety rather than to crime. After all, medieval cities were built to protect their residents (Pirenne 1929). Archer and Gartner (1984) find that in six out of 24 countries they survey, homicide rates are lower in the largest city (Tokyo is the prime example).

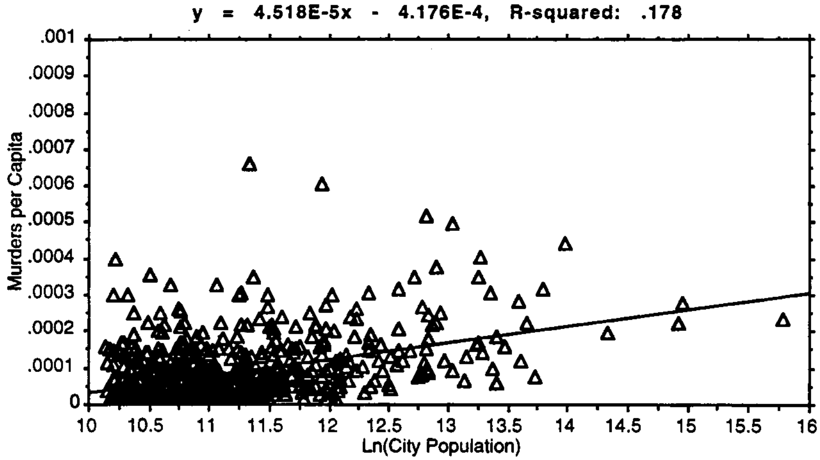

FIG. 2.—Murder and city population: relationship between crime and city population taken from the 1982 Uniform Crime Reports and the 1980 census. *t*-statistic is 11.27.

teenth century, the crime rate in at least one city was high enough to justify these comments.

This paper tests a variety of theories about the correlation between crime and urban size.[3] We decompose the observed connection between cities and crime into three main categories: (1) higher pecuniary returns to crime in urban areas, (2) lower probability of arrest in urban areas, and (3) urban areas' attraction (or creation) of crime-prone individuals. We use data from the National Crime Victimization Survey (NCVS), the National Longitudinal Survey of Youth (NLSY), and the Uniform Crime Reports (UCR) to decompose the urban crime premium into these different factors.

Cities may create greater returns to crime because criminals may have greater access to the wealthy and face a greater density of victims in urban areas. The connection between city size and value per crime is large (.13 in the victimization records and .09 in the UCR data). Depending on the elasticity of crime with respect to deterrence, greater returns to crime in cities can explain anywhere from 13.33 percent to one-third of the urban crime effect.

[3] Across cities, increases in population are often uncorrelated with changes in crime (see Archer and Gartner [1984] for a discussion). The natural explanation of this seemingly anomalous fact is that there is a reverse causality problem in which increases in crime lead to declines in population growth. The existence of safe, highly urbanized societies (such as Japan or Holland) can certainly be explained as a result of strong social cohesion, which makes it possible to urbanize despite the problems that large cities bring.

If we control for underreporting of crimes in cities, it appears that committing crimes in cities is less likely to lead to an arrest (the elasticity is −.10). One explanation for this phenomenon is that for any given crime, dense urban areas have a much larger number of suspects. The importance of this phenomenon depends on the elasticity of crime with respect to deterrence. If one accepts relatively modest deterrence elasticities (such as those put forth by Levitt [1998*b*]), then lower arrest rates in cities explain around 10 percent of the urban crime effect. If one believes the significantly higher estimates found in the traditional deterrence literature (e.g., Ehrlich 1973, 1975), the importance of this effect may be many times larger.

Between 30 and 50 percent of the urban crime premium in both UCR data and NCVS data can be explained by observable characteristics of urban residents. In the cross-city UCR data, the most important variable is the percentage of female-headed households in the city. This result withstands instrumenting for the number of female-headed households with the lagged value of the state-level generosity of payments from Aid to Families with Dependent Children (AFDC).

The residual urban crime effect can be explained by other omitted social variables, which are often endogenous, and possibly by the weakness of social sanctions in dense urban areas. For example, in the NLSY, variables such as marijuana usage, tendencies toward noncriminal violence, church attendance, and variables related to patience all are significant in their ability to explain crime. Some of these variables are related to urban residence, and controlling for these forces can reduce the urban crime effect substantially. We do not have these variables in the UCR cross-city data, on which our decompositions are based, so the remaining city crime effect may in part be the result of omitting these variables.

This paper makes four points. First, pecuniary returns and classic deterrence are important in explaining the urban crime premium. Second, while these classic forces are important, together they probably explain not more than 50 percent of the connection between crime and city size, so that we must look elsewhere to understand one-half of the urban crime premium. Third, a sizable component of the urban crime premium is related to cities attracting crime-prone individuals. Understanding why cities attract these individuals is a pressing topic for further research. Finally, our findings suggest that cities create both positive and negative agglomeration economies, and similar forces lie behind both the positive and the negative attributes of cities. For example, the same larger market size that makes cities appealing to firms, especially to those with fixed costs, will also make cities more appealing to drug dealers and thieves trying to pawn stolen merchandise.

Section II of the paper presents a simple decomposition for examining the prevalence of crime. The remainder of the paper discusses the estimation of parameters used in the decomposition.

II. A Decomposition

This decomposition attempts to separate the potential causes of the relationship between cities and the level of criminal behavior.[4] As in Becker (1968), individuals commit crimes when the benefits of crime exceed the costs. In general, the rule is to commit a crime whenever benefits (denoted B) exceed total costs (denoted $\theta + PC$), or $B > \theta + PC$. We divide total costs into costs associated with the crime itself (lost time, pangs of conscience, etc., which are denoted θ) and costs associated with the probability of arrest (P) times the costs of punishment (C).[5]

The variable θ is a function of a vector $\mathbf{X}$ of individual attributes. These attributes are correlated with and determined by location. The benefits of crime, the probability of arrest, and the probability of facing social reproach are also a function of location. The cost of punishment (C) is assumed to be a constant over space and across individuals.[6] The benefits per crime will also be a declining function of the total number of crimes committed (denoted Q). For simplicity, we assume that all potential criminals in the same location have the same $\mathbf{X}$ variables.[7] Thus the criminal equilibrium will be defined by the following condition:

$$B(Y, Q) = \theta(\mathbf{X}) + P(Y)C. \tag{1}$$

[4] In this model, we shall intentionally ignore the intensive margin of criminality. While sources on crimes per criminal do exist (Chaikin 1978; Blumstein and Cohen 1979), we do not currently have the ability to link the number of crimes per criminal with city size.

[5] In a previous draft (Glaeser and Sacerdote 1996), we also emphasized the role of community sanctions in stemming crime. While in principle this effect may be important, we were able to measure only dim proxies for it, such as the probability of knowing one's offender. This probability is higher in cities, but even given this effect, our work estimated that community sanctions explain no more than 10 percent of the urban crime effect.

[6] In fact, the opportunity costs of time lost in jail appear to be higher in urban areas, which should create less crime. In our NLSY sample, among individuals who had stolen in the past year, earnings are about 10 percent higher in metropolitan areas.

[7] It is relatively straightforward to allow for heterogeneity across individuals. In that case, eq. (1) must hold only for the marginal criminal.

We use the convenient notation that the partial elasticity of a variable A with respect to another variable B is $\epsilon_B^A = (B/A)(\partial A/\partial B)$; differentiation implies that

$$\epsilon_P^Q = \frac{P}{Q}\frac{\partial Q}{\partial P} = \frac{PC}{QB_Q}.$$

We can then differentiate (1) with respect to city size (denoted N) to find

$$\epsilon_N^Q = \epsilon_P^Q \epsilon_N^P - \frac{B}{PC}\epsilon_P^Q \epsilon_N^B + \sum_x \epsilon_{\mathbf{X}}^Q \epsilon_N^{\mathbf{X}}. \tag{2}$$

Location-specific attributes, such as city size, might (1) change the probability of apprehension (the $\epsilon_P^Q \epsilon_N^P$ term) or (2) affect the returns from crime (the $[B/PC][\epsilon_P^Q \epsilon_N^B]$ term); alternatively, (3) the community attributes might change the level of location-based individual characteristics that could affect the apprehension-invariant costs of crime (the $\sum_x \epsilon_{\mathbf{X}}^Q \epsilon_N^{\mathbf{X}}$ term). Our objective is to estimate these parameters and to decompose the connection between crime and city size.

The purpose of this decomposition is to determine why cities are so much more crime-prone than small towns and suburbs. Is there more crime in cities because urban density makes it harder for the police to track criminals? Does urban crime result from higher pecuniary returns due to lower transport costs and greater market size? Or are cities filled with crime because cities are filled with crime-prone individuals?

III. Methodology and Data

The decomposition suggested three primary ways in which cities might affect the crime rate. In order to implement this, we need to determine the magnitude of three sets of parameters listed immediately above. When possible, we shall take parameters from elsewhere in the literature. For example, the connection between the probability of arrest and the level of crime is a much-studied issue, and we are surely not going to improve on that literature here. However, for some parameters—in particular, those relating crime with city size—we use our own data.

Data Sources

All three data sets are discussed at length in the Appendix. The means of our primary variables from all three data sources are in table 1. Here we present a brief overview of them. Our first data source is the National Crime Victimization Survey administered by the Department of Justice's Bureau of Justice Statistics. The NCVS asks respondents a battery of background information questions, and it asks them whether or not they were the victims of a crime within the prior six months. If the respondent was victimized one or more times, the NCVS asks dozens of further questions about each incident.[8] We use two different samples from the NCVS; we limit both samples to heads of households to increase the accuracy of answers. Our first sample is the full sample from 1989 third-quarter interviews. This sample contains people who have and who have not been victimized in the past six months. Our second sample is taken from the third-quarter interviews of the 1980 incident-level file. This sample includes only people who were victimized. We believe that the costs of using older (1980) data are made up for by the fact that the 1980 data have a much larger sample of victims. Unfortunately, neither sample tells us the exact city in which the respondent lives. We are forced to use a range of population values for the respondent's place of residence.

Our second data source is the Federal Bureau of Investigation's (FBI's) Uniform Crime Reporting program. These data give us crimes, arrests, and value of property taken by city by year. We link the UCR data with city demographics from the *County and City Data Book* (CCDB), and these demographics have generally been drawn ultimately from the U.S. census. Our sample includes those cities with more than 25,000 people and for which we have both CCDB and UCR data. We used 1982 data for comparability for our NCVS and NLSY data sets.

The UCR data are based on law enforcement agencies' reports to the FBI regarding the number and type of crimes reported in the local jurisdiction. There is an extensive literature discussing the variety of flaws with the UCR data.[9] A well-known issue is the potential for UCR numbers to understate the extent of criminal activity if citizens underreport crimes to police. We employ a simple solution to the reporting problem. For this paper, we are primarily interested in underreporting, as far as it is related to city size. The NCVS en-

[8] The NCV surveys are not administered to a random sample of the U.S. population, and so our calculated victimization rates will not be the same as the U.S. victimization rates. However, the nonrandomness of the sample should not in any way affect our estimates of the correlates of victimization.

[9] For example, Levitt (1998*a*) shows both that there is substantial underreporting and that the reporting propensities increase with increases in the police force. Donohue and Siegelman (1998) provide an extensive discussion of the reporting biases inherent in the UCR.

TABLE 1

MEANS BY SMSA AND NON-SMSA

A. 1989 NCVS DATA*

	All (N = 8,328)	Non-SMSA (N = 1,929)	SMSA (N = 6,399)
Victim of a crime	.13	.08	.14
	(.33)	(.28)	(.35)
SMSA	.77	0	1
	(.4219)	(.00)	(.00)
Black	.10	.07	.11
	(.30)	(.25)	(.31)
Hispanic	.07	.04	.07
	(.25)	(.19)	(.26)
Reported crime to police	.41	.38	.41
	(.49)	(.49)	(.49)
Value of property taken	489.11	378.80	508.80
	(1,153.25)	(965.45)	(1,182.19)
Knew offender (or one of multiple offenders)	.46	.58	.44
	(.50)	(.51)	(.50)

B. NLSY DATA*

	All (N = 9,145)	Non-SMSA (N = 2,872)	SMSA (N = 6,273)
Age	20.36	20.29	20.40
	(2.22)	(2.21)	(2.23)
SMSA	.69	0	1
	(.46)	(.00)	(.00)
Stole something <$50 in past year	.18	.16	.20
	(.39)	(.37)	(.40)
Shoplifted in past year	.27	.22	.29
	(.44)	(.42)	(.45)

C. CROSS-CITY DATA: UCR AND CCDB

	N = 633
Serious crimes per capita:	
Unadjusted	.08
	(.03)
Adjusted for reporting bias	.17
	(.07)
Value taken per crime	$543
	($304)
Arrests per crime	.24
	(.08)
Population 1982	116,852
	(334,666)

NOTE.—Standard deviations are in parentheses.

* Sample sizes for arrest made, report to police, value taken, and knew offender are 346, 1,048, 709, and 145.

ables us to estimate the ratio of reported crimes to actual crimes by city size. We use this ratio to adjust the UCR crime rates differentially by city size. Since larger cities appear to have more underreporting, larger cities get a larger adjustment.

Our final data source is the National Longitudinal Survey of Youth, which provides self-reported evidence on criminal behavior (see also Freeman 1992; Grogger 1995). Our sample takes all respondents who answered questions in 1980 about whether or not they had engaged in various criminal activities and for whom we know basic family and individual characteristics (e.g., race, income, and census region). We are skeptical about relying on self-reported criminality measures, but the high degree to which individuals seem to freely admit to petty crimes makes us somewhat more reassured.[10]

Because of significant problems with other measures of urban status in the NLSY, we use whether the individuals inhabit a standard metropolitan statistical area (SMSA) as our measure of urban status. This measure eliminates any of the variation between large cities and suburbs with metropolitan areas. A large majority of the sample (69 percent) lives in SMSAs.

How Big Is the Connection between Crime and City Size?

We begin with our estimates of the elasticity of reported crime to city size. Two of our data sets allow us to estimate such a connection. Table 2 shows results from the different data sets. The first regression uses the UCR data to regress the serious crimes per capita on city size. We find an overall elasticity of .16, controlling for the region of the country.

If there is greater underreporting in cities, this finding may actually understate the true elasticity. To adjust for underreporting, we use the NCVS, which tells us whether the victim reports the crime. Using this survey, we can determine the mean level of underreporting by city size category. Figure 3 shows the amount of underreporting by city size. We have then adjusted the reported crime rate by the level of underreporting in regression 2. As there is somewhat less reporting in larger cities, this adjustment causes the link between city size and crime to rise: the adjusted elasticity is .24.

Regression 3 shows that the link between crime and city size is not a new phenomenon. In 1970, the elasticity of crime with respect to city size was higher than it was in 1986. Regression 4 shows that the

[10] DiIulio and Piehl (1991) extensively discuss self-reported data on crime. Spelman (1994) suggests that self-reported crime data of prisoners appear to match police records.

TABLE 2

CONNECTION BETWEEN CRIME AND CITY SIZE

	Dependent Variable					
	OLS on Cross-City UCR Data					
	Log Serious Crimes per Capita (1)	Log Serious Crimes Adjusted* (2)	Log Serious Crimes per Capita 1970 (3)	Log Murders per Capita (4)	Probit on NCVS Micro Data: Victim of a Crime (5)	Probit on NLSY Micro Data: Stolen Property <$50 in Past Year (6)
Intercept	−4.41	−4.50	−6.48	−13.43		
	(.20)	(.20)	(.23)	(.42)		
Log(city population)	.16	.24	.28	.32		
	(.02)	(.02)	(.02)	(.04)		
Implied elasticity of "victim" with respect to city size					.12	
SMSA						.03
						(.02)
City population:						
1,000–9,999					.01	
					(.01)	
10,000–24,999					.04	
					(.01)	
25,000–49,999					.06	
					(.02)	
50,000–99,999					.09	
					(.02)	
100,000–249,999					.11	
					(.02)	
250,000–499,999					.09	
					(.02)	
500,000–999,999					.15	
					(.03)	
$\geq$ 1,000,000					.16	
					(.03)	
Regional dummies included?	yes	yes	yes	yes		
R^2	.20	.29	.29	.25		
$1 - [\ln(L)/\ln(L_0)]$					.02	.01
Observations	634	634	634	541	8,328	8,910

NOTE.—Standard errors are in parentheses. For ease of interpretation, probits show partial derivatives of Y with respect to **X** rather than the actual probit coefficient itself. The NCVS provides city size categories as shown.

* Adjusted crime numbers are adjusted *by city size* for underreporting of crimes. Adjustment methodology uses crime and reporting data from the NCVS.

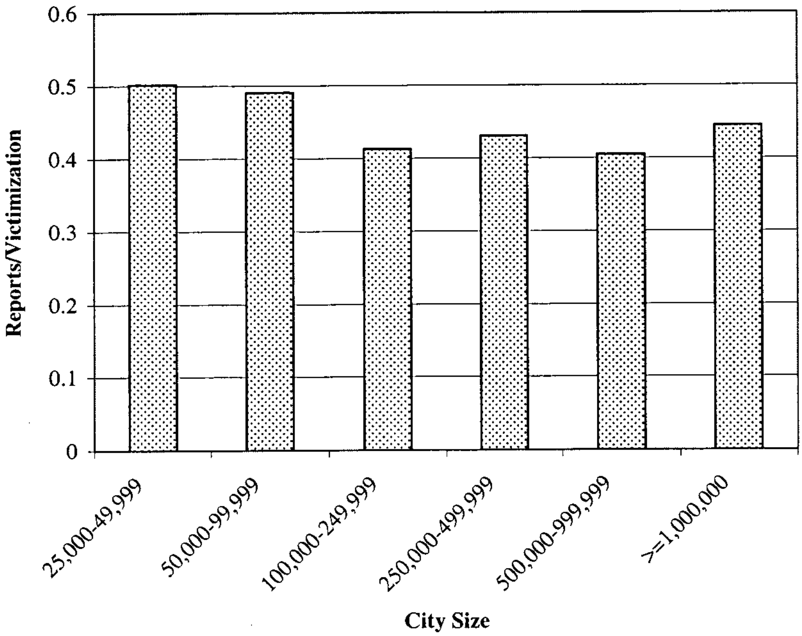

FIG. 3.—Reporting of crimes by city size. Source: National Crime Victimization Survey.

connection between crime and city size is even stronger when we examine the most serious crime that is least subject to reporting error—murder. The elasticity of murders with respect to city size is .32.

Regression 5 uses the victimization surveys and also finds that crime is higher in large cities. For the NCVS we do not observe the actual city size, but instead a range of city size groups that describes the city in question. We have used the estimates from these ranges of city sizes to calculate an overall elasticity of .12, which is reported in the table as the implied elasticity. There is a significant gap between the victimization elasticity with respect to city size and the elasticity between crime and city size implied by both the base UCR data and the adjusted UCR data. One explanation for this difference is that victimization reports will exclude most crimes against businesses and that cities are centers of business activity. Regression 6 uses the self-reported criminality data in the NLSY and shows that the self-reported probability of having committed a petty theft in the past year for the sample of youths is higher in metropolitan areas.[11]

[11] In Glaeser and Sacerdote (1996), we reproduce all these results using self-reported shoplifting as well. In unreported regressions we have also used as the measure of criminal behavior being incarcerated during one of one's NLSY interviews. We have restricted our attention to this minor crime because we believe that

IV. The Importance of Deterrence—$\epsilon_P^Q \epsilon_N^P$

Cities might lower the costs of crime by lowering the probability of arrest and the probability of punishment conditional on arrest. A natural hypothesis is that police are more effective in small towns because they are more likely to know the residents of their community in a relatively stable, small-town environment. One version of this hypothesis is that police solve crimes by considering the full set of possible suspects in a crime and eliminating all but one of the potential criminals.[12] In small towns, police are able to narrow the range of suspects in any particular crime to a much smaller set and apprehend criminals much more easily. As cities grow, if the number of crimes and the number of police grow at the same rate, then apprehension will still be harder in larger cities because the pool of potential suspects is larger. Alternatively, the connection between policing and urban areas may come about because cities choose to spend fewer dollars on crime prevention and to acquire fewer police officers.

Estimating the Parameters

We are now interested in the elasticity of crime with respect to the probability of arrest times the elasticity of the probability of arrest with respect to city size. Throughout this section, we ignore issues that might come from a lower probability of incarceration conditional on arrest in cities and a lower length of sentence conditional on incarceration in cities. The data on these issues, which are available, suggest that there are small differences associated with highly urbanized areas that tend to suggest that there is less deterrence in urbanized areas along these dimensions as well.[13]

The elasticity of the quantity of crime with respect to the probability of arrest is perhaps the most important single elasticity in the empirical literature on crime. As the amount of crime may also influence the probability of arrest through congestion of law enforcement, and as other factors may determine both variables, ordinary

the probability of dishonest answers will be lower with a less serious offense, and with this offense a much higher share (18 percent) of the population admits to having committed the crime.

[12] Indeed, basic police procedure textbooks (e.g., Weston and Wells 1994) do recommend just such an investigation process in many instances.

[13] We have examined incarceration rates for big counties (which are more likely to be heavily urban) relative to the United States as a whole. While the big counties have slightly higher incarceration rates, the difference is not huge. For example, 47 percent of arrests for murder end in prison in large counties and 52 percent of arrests for murder end in prison in the rest of the country.

least squares (OLS) estimates have long been thought to be biased (see Taylor 1978). However, many of the early two-stage least-squares estimates rely on identification restrictions that seem difficult to accept. Unfortunately, one cannot even convincingly sign the bias facing OLS estimates.

In three separate papers, Levitt (1996, 1997, 1998*b*) provides the best evidence on the elasticity of crime with respect to different measures of deterrence (probability of arrest, number of police officers, and size of prison population), which collectively strongly suggests that an estimate between $-.1$ and $-.4$ is reasonable. His work uses sophisticated instrumentation strategies or use of timing to estimate the effects of deterrence.[14] Other authors (such as Grogger [1991]) find similar estimates, but older researchers occasionally find higher elasticities. For example, the classic OLS estimate of ϵ_P^Q by Ehrlich (1973) is approximately -5. Early two-stage least-squares estimates are generally higher. For example, the Ehrlich estimate is $-.991$, and an estimate by Mathieson and Passell (1976) is -2.96. These estimates rely on dubious instrumentation strategies and seem implausibly large relative to most work. We shall present results for parameter estimates of $-.2$ and $-.5$, which seem like reasonable bounds on the magnitude of ϵ_P^Q.

While we can take estimates of ϵ_P^Q from the existing literature, we need to determine our own estimates of the elasticity of the probability of arrest with respect to city size. Our methodology estimates

$$\log\left(\frac{\text{arrests}_{i,j}}{\text{crime}_{i,j}}\right) = \alpha_i + \beta\log(N_j) + \mathbf{X}_j'\Delta + \mu_j + \epsilon_{i,j}, \qquad (3)$$

where i indexes across crime and j indexes across city, and N_j is city population, $\mathbf{X}_j$ are other city-specific characteristics, and μ_j is a city-specific random effect. We have a different observation for each type of crime, and allowing for different levels of α_i enables us to correct for the possibility that there is a different probability of arrest associated with each different type of crime. We assume that city-level characteristics affect the probability of arrest for each type of crime in the same way.[15]

[14] As long as the probability of arrest is a linear function of the number of police and the amount of imprisonment, then all Levitt's elasticities should be close and all serve as estimates of ϵ_P^Q.

[15] We have investigated this assumption, and it appears that the elasticity of the probability of arrest with respect to city size does not change significantly across the different types of crimes.

Regression 1 in table 3 uses the UCR to show the most basic estimate of the coefficient β: $-.05$. Regression 2 in table 3 shows that this elasticity falls to $-.03$ when we have controlled for a variety of other city-level controls. If we control for the city's crime rate itself (not shown, because of the extreme endogeneity problems involved in that regression), our estimate falls to $-.02$.

These different specifications have advantages and costs, and since we could not possibly find exogenous instruments for each variable, no specification is perfect. The decomposition model suggests that we should control for city-level characteristics and the average crime rate. However, these variables are often endogenous with respect to the arrest rate. As the parameter estimates do not change terribly much, our approach will be to show results in several specifications and to ensure that they do not change much depending on the particular specification in question. Under some conditions, our two specifications give bounds on the size of the connection between city size and the arrest rate.[16]

Reporting problems matter if individuals in cities are less likely to report crime than individuals in small towns. In this case, while the number of arrests per reported crime might be the same in big cities as in small towns, the number of arrests per actual crime may be much lower in big cities. To correct for this problem, we again multiplied the crime rate in large cities by the extent of underreporting in each city class in the NCVS (as described above). In this way, we should have a measure of crimes per city for which underreporting is not correlated with city size.[17] In table 3, regression 3, we regress this adjusted crime rate on city size. In regression 4, we regress the adjusted crime rate on city size and the full set of other controls. Again the elasticity estimates are quite close, ranging from $-.13$ to $-.10$. If we run regressions with the actual crime rate, we get an elasticity of $-.09$.[18]

[16] In the case of a true model in which $Y = b\mathbf{X} + cZ$, failure to control for Z will mean that the estimated value of b equals $b + c \cdot [\mathrm{cov}(\mathbf{X}, Z)/\mathrm{var}(\mathbf{X})]$. Controlling for Z when c is misestimated will mean that the estimated value of b equals $b + (c - c') \cdot [\mathrm{cov}(\mathbf{X}, Z)/\mathrm{var}(\mathbf{X})]$. In the case in which Y reflects crime, $\mathbf{X}$ reflects city size and Z reflects the crime rate, and both b and c are truly negative, the OLS estimate of b with no other controls is biased away from zero (because $c \cdot [\mathrm{cov}(\mathbf{X}, Z)/\mathrm{var}(\mathbf{X})]$ is negative), but the estimate of b is biased toward zero when Z is controlled for and when c' overstates the extent to which higher crime rates cause lower arrest rates because of reverse causality.

[17] Of course, since we have not adjusted by a city-specific reporting ratio, the number of crimes will still be mismeasured, but at least this mismeasurement will not be correlated with city size.

[18] Unfortunately, we are not able to estimate the elasticity of arrest rates with respect to city size for our NCVS sample. We have only 33 observations in which crimes led to an arrest.

TABLE 3

Cities and the Probability of Arrest

	Dependent Variable					
	UCR Data					
	Log Arrests per Crime		Log Adjusted Arrests per Crime		Log Police Officers per Capita	NCVS Data: Police Visit Crime Scene
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	.43	−.16	.53	.02	−7.10	
	(.21)	(.34)	(.13)	(.34)	(.19)	
Log(city population 1982)	−.05	−.03	−.13	−.10	.05	
	(.02)	(.02)	(.02)	(.02)	(.01)	
Implied elasticity with respect to city size						−.01
Percentage population below poverty, 1979		1.32		1.41	−1.82	−.03
		(.55)		(.55)	(.31)	(.03)
Percentage housing owner occupied		.14		.13	−.08	−.01
		(.18)		(.18)	(.10)	(.04)
Percentage nonwhite		−.38		−.36	−.26	.04
		(.17)		(.18)	(.10)	(.04)
Percentage with 4 years high school		.49		.40	−.71	−.01
		(.32)		(.32)	(.18)	(.01)
Percentage with 4 years college		−.34		−.30	.81	
		(.30)		(.30)	(.17)	
Unemployment rate		−1.10		−1.20	.34	.02
		(.70)		(.70)	(.39)	(.03)
Percentage female-headed household		−.69		−.83	4.10	.00
		(.65)		(.66)	(.36)	(.03)
Regional dummies included?	yes	yes	yes	yes	yes	no
Number of cities × 6 different types of crime*	3,663	3,663	3,663	3,663	. . .	
Number of cities or people	634	634	634	634	634	1,054

Note.—Standard errors are in parentheses. Regressions 1–4 are estimated using a panel approach: random city effects with fixed effects for type of crime. Regression 6 is a probit of "police visit crime scene" if a crime was committed. For interpretation, elasticities are shown rather than probit coefficients. For this regression, the following other variables are used: for percentage high school, highest grade completed; for percentage nonwhite, a dummy variable for black; for unemployment rate, had job; for percentage housing owner occupied, owns home; for percentage population below poverty, a dummy variable for low income; and for percentage female-headed household, married. The elasticity in regression 6 is calculated from coefficients on eight city size dummies.

* Log(arrest rate) is not available for all cities and all crimes.

In regression 5, we test the subhypothesis that lower crime rates in cities are related to greater levels of spending on police. We find that big cities spend considerably more per capita on the number of police. This result is true even if we control for the average crime rate per capita (not in the table), although in that case the elasticity of police officers per capita with respect to city size falls to .02. Regression 6 uses NCVS data to show that police are no less likely to come to the scene of a crime in larger cities. Thus we believe that the probability of being arrested in cities is indeed lower, but this reduction is not the result of less spending on police in cities or a lower likelihood that police investigate crimes. The lower probability of arrest apparently stems from a greater difficulty of catching criminals in large urban areas.

Magnitude of effect.—It appears that the probability of arrest is indeed lower in cities. The elasticity of the arrest rate with respect to city size plausibly ranges between −.03 and −.13, with −.08 as a reasonable midpoint. If we accept the −.2 estimate of the effect of deterrence on crime, the role of less deterrence in explaining why cities have more crime cannot be more than −.02, or approximately 8.33 percent of the .24 estimate when adjusted crime rate data are used. (If we use the upper bound for the arrest rate elasticity, it seems reasonable to compare this with the upper bound elasticity connecting city size with crime.) However, if we use the −.5 estimates of the effect of deterrence on crime, then as much as −.05 or 20.8 percent of the effect can be explained by less deterrence in cities.

Informal community sanctions may be as important as formal, legal sanctions in eliminating crime. Urban anonymity makes it hard to enforce sanctions (Wilson and Herrnstein 1985). Likewise, the communities that would enforce any sanctions may be weaker in cities because people in the cities are more transient or more anonymous and therefore more likely to be free riders (see Wirth 1938; Putnam 1993).[19] Glaeser (1998) formalizes some of these arguments and presents evidence suggesting that social cohesion appears to decline in large cities. In a previous version of this paper (Glaeser and Sacerdote 1996), we estimate the relationship between knowing one's offender and city size as −.11. This finding suggests that urban anonymity is a fact, but it does not suggest how important this fact may be in crime prevention.

[19] Milgram (1970) claims that cities create an informational overload that leads bystanders to avoid involvement in crimes against their neighbor. Jacobs (1961) argues that cities abet crimes only when urban neighborhoods lose their traditional social structures.

V. The Returns to Crime and City Size—$\epsilon_P^Q \epsilon_N^B B/PC$

A natural explanation for why cities have a high return from crime is that costs of transport for crime are extremely high. Indeed, criminological work strongly suggests that criminals do not travel long distances to perform crimes. These high transport costs may stem from the need to leave the scene of the crime quickly or the difficulties inherent in carrying stolen merchandise over long distances.[20] A particular form of the advantages of density is that cities create crime by creating proximity between wealthy potential victims and poor potential criminals. Urban density should lower transport costs, increase the returns per crime, and increase the overall crime level.

Density may play a particular role in street crime, where the method of street criminals is essentially to sit and wait for prospective victims to come within their range of sight. A dense area will have a much larger stream of potential victims than an empty area, and the returns from this type of crime should be higher in urban areas. Either the returns per crime will rise with density as criminals choose only the more promising victims or criminals will select more victims and the returns per hour of criminal activity will rise with density.

Dense urban areas may also help criminals become better informed about a wider range of victims, and thus criminals in big cities will be able to choose the most lucrative crime among a greater range of crimes. Greater information flows in cities also make it possible for individuals to acquire information that will itself reduce the costs of crime to the criminal, for example, learning easier ways to break into apartments. Finally, cities may raise the returns to criminal activity because urban areas create scale economies that ease the resale of stolen goods or the purchase of criminal implements (e.g., guns). It is ironic that the same urban advantages, lower transport costs, faster urban information flows, and the same scale economies that help to make cities more productive also increase the level of crime in the city.

Estimating the Parameters

The effect of values per crime is $\epsilon_P^Q \epsilon_N^B B/PC$, where ϵ_P^Q is the same deterrence elasticity discussed above. The ratio B/PC (cash benefits of crime over law-related expected costs) is necessary so that we can transform the elasticity of crime with respect to the probability of

[20] These costs of distance become particularly high if the criminal belongs to a socioeconomic or ethnic group that is visibly different from the norm of the victim's neighborhood.

arrest into the elasticity of crime with respect to pecuniary benefits. The average financial loss reported in the NCVS per crime is $543. The average financial loss in the UCR data is $489 per crime in property crimes. We shall therefore use $500 as our estimate of *B*.[21]

The probability of reporting a crime is 41 percent according to the NCVS data. From the UCR, the probability of arrest per reported crime across the United States as a whole is 24 percent. We shall borrow from Levitt (1998*b*) for the expected amount of time lost conditional on arrest.[22] He calculates an average loss of 70 days per arrest over property crimes. Thus the total amount of time lost per crime is 6.89 days. If we calculate each day as being worth $50 (10 hours times $5 per hour), then the ratio of pecuniary benefits over costs is 1.45. While this estimate is presumably noisy, it seems like a reasonable benchmark.[23]

In table 4, we estimate the connection between returns to crime and city size. Regression 1 shows the connection between value per crime and city size. We again treat each crime category in each city as a separate observation and control both for crime category–specific fixed effects and city-specific random effects. We consider only crimes that have a pecuniary reward (e.g., robberies, larcenies, burglaries, and auto theft). In regression 2, when we control for other city-level characteristics, our estimated elasticity is .09. If we reproduce these results controlling for the number of crimes per capita (in each crime category), the elasticity is .10. While there are advantages and disadvantages of each specification, as all three specifications produce similar parameter estimates, we are comfortable with the complete picture from the regressions.[24]

The next two regressions in table 4 show results from the NCVS. Regression 3 gives the elasticity with respect to city size when region and crime category are controlled for: .14. Regression 4 shows that the elasticity is .13 when victim-level characteristics are controlled for. Regression 5 uses the NCVS and examines only those assaults (including rape) that were not accompanied by any sort of theft. Regression 6 examines the number of rapes per capita in the UCR data. In both cases, the effect of city size on crime is greater with

[21] It is certainly likely that the loss to the victim is less than the gain to the criminal. However, this number might also be biased downward by underreporting.

[22] Levitt considers time lost only within two years of arrest, but for many crimes and in most cases, this is probably a good approximation, especially if discount rates are high.

[23] One way of judging the reasonableness of this number is that it implies that 69 percent of the costs of crime are related to police and 31 percent of the costs of crime come from other sources.

[24] Furthermore, our results might be biased if individuals in either location were attempting more difficult crimes.

these nonpecuniary crimes. We must therefore conclude that in many cases the connection between crime and city size has nothing to do with higher pecuniary returns in cities.

Magnitude of effect.—We shall consider the benchmark elasticity of .11 for the elasticity of benefits with respect to city size. However, the magnitude of this effect hinges critically on one's view of the estimate of the elasticity of crime with respect to deterrence. If one accepts the estimate of $-.2$ for the deterrence elasticity and 1.45 as the value of B/PC, then the size of this effect should be approximately .032, which is more than 10 percent of the connection between crime and city size. If one accepts the higher estimate of the deterrence elasticity, then this number rises to .080, which is one-third of the overall cities-crime connection when we adjust for greater underreporting in cities.

VI. Crime-Prone Individuals and Cities—$\sum_x \epsilon^Q_{\mathbf{X}} \epsilon^{\mathbf{X}}_N$

High urban crime levels may occur because people with a greater propensity toward crime may choose to live in cities. This is an old view: Genesis 10:6 claims that the first murderer (Cain) built the first city. Attributes of urban neighborhoods or labor markets or transfer programs may selectively induce crime-prone persons to come to the cities.

Alternatively, cities may actually alter their residents in a way that makes them more prone toward crime. Attributes of cities may alter people's investment in human capital or in tastes (such as patience) that are related to crime. Social interactions and neighborhood effects (as in Case and Katz [1991]) may be more important in dense urban areas, and these forces may alter preferences (Wilson and Herrnstein 1985). These peer influences are found to be more influential when there are weak families, so we should expect that family structure in cities might be an important explanatory variable as well (Glaeser, Sacerdote, and Scheinkman 1996).

Estimating the Parameters

In general, we shall just report the extent to which the relationship between city size and the level of crime is affected by the inclusion of other variables. In general,

$$\frac{d\log Q}{d\log N} = \frac{\partial \log Q}{\partial \log N} + \sum_x \frac{\partial \log Q}{\partial \mathbf{X}} \cdot \frac{\partial \mathbf{X}}{\partial \log N} = \frac{\partial \log Q}{\partial \log N} + \sum_x \epsilon^Q_{\mathbf{X}} \epsilon^{\mathbf{X}}_N, \quad (4)$$

TABLE 4

CITIES AND THE RETURNS TO CRIME

	OLS UCR: Log(Value per Crime)		OLS NCVS 1980: Log(Value Taken)		Probit NCVS: Victim of Assault	OLS UCR: Log(Rapes per Capita)
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	7.71	8.18	4.00	3.36		−12.48
	(.28)	(.44)	(.06)	(.17)		(.54)
Log(population 1982)	.05	.09				.21
	(.02)	(.02)				(.03)
Age				−.00	−.0002	
				(.00)	(.0000)	
Male				.20	.0033	
				(.06)	(.0014)	
Highest grade achieved		−.04		−.00	−.0002	2.07
		(.41)		(.01)	(.0002)	(.52)
Black		.41		.42	−.0032	−.24
		(.23)		(.11)	(.0009)	(.28)
Had job last week				−.01	.0001	1.28
				(.07)	(.0012)	(1.11)
Owns residence		−.46		.02	−.0043	−.26
		(.23)		(.08)	(.0019)	(.31)
Married				−.06	−.0036	
				(.07)	(.0016)	
Number of cars owned				.09	−.0013	
				(.03)	(.0006)	

	(1)	(2)	(3)	(4)	(5)	(6)
Lives in building with ≥10 apartments				.17	−.0013	
				(.10)	(.0010)	
Family income (15 groups)		−1.61		.04	.0027	
		(.70)		(.01)	(.0017)	
City population:						
1,000–9,999			−.12	−.05	−.0004	
			(.10)	(.10)	(.0017)	
10,000–24,999			−.06	−.05	−.0034	
			(.11)	(.11)	(.0011)	
25,000–49,999			−.08	−.07	.0033	
			(.12)	(.12)	(.0029)	
50,000–99,999			.14	.16	.0052	
			(.12)	(.12)	(.0036)	
100,000–249,999			−.03	.01	−.0003	
			(.12)	(.12)	(.0019)	
250,000–499,999			.21	.22	.0013	
			(.14)	(.14)	(.0029)	
500,000–999,999			.21	.24	.0029	
			(.14)	(.14)	(.0035)	
≥ 1,000,000			.94	.88	.0055	
			(.12)	(.13)	(.0044)	
Elasticity of value (or ‘‘report’’ or ‘‘knew’’) with respect to city size	.05	.09	.14	.13	.22	.21
	(.02)	(.02)				(.03)
$1 - [\ln(L)/\ln(L_0)]$ or R^2			.02		.15	.44
Observations	634	634	3,498	3,498	8,328	604

Note.—Regression 1 is the log(value) on city size run as a panel with fixed effects for the type of crime. Standard errors are corrected for city-level random effects. Regression 2 adds demographic controls (e.g., percentage below the poverty line and the percentage high school graduates) and four dummies for the type of crime (e.g., robberies, burglaries, larcenies, and auto thefts). Regressions 1, 2, and 6 include regional dummies. For the UCR cross-city regression, the following other variables are used: for owns residence, percentage housing owner occupied; for had job, city unemployment rate; for highest grade achieved, percentage graduated from high school; and for family income, percentage below poverty.

or the difference between the raw coefficient of crime on city size and the coefficient of crime on city size when we control for other variables is equal to $\sum_x \epsilon^Q_{\mathbf{X}}\epsilon^{\mathbf{X}}_N$. Thus the difference in the coefficients between a regression that does not control for **X** variables and a regression that does control for **X** variables provides us with an estimate of $\sum_x \epsilon^Q_{\mathbf{X}}\epsilon^{\mathbf{X}}_N$.

The model suggests that the key variables will be the characteristics of the offenders. To examine this hypothesis, we must use the NLSY, which contains micro-level evidence connecting the offender's characteristics with his or her probability of committing a crime. As mentioned earlier, the raw coefficient of SMSA status on crime from the NLSY is .03. In part A of table 5, regression 1, we find that many other variables affect the propensity to become a criminal. However, the coefficient on city size does not change at all when we control for a variety of basic individual demographics.

Regression 2 includes the SMSA status at age 14 and whether the individual migrated since age 14. Individuals who lived in SMSAs at age 14, but do not as adults, are just as prone to crime as individuals who now live in SMSAs. An extreme interpretation of this result is that since people raised in cities are just as crime-prone when they leave, we should interpret all of the urban crime premium as coming from characteristics of urban residents. For many reasons (the limits on geography in this data set or the problems with using self-reported data on petty theft), we prefer to make a more limited interpretation that the NLSY suggests that the characteristics of urban residents are very important.

In regression 3 we include a variety of other variables that are highly endogenous and may reflect omitted individual characteristics. These variables may be correlated with patience, risk aversion, or willingness to break rules, and we view them as proxies for these attributes. We include marijuana usage, which is extremely strongly correlated with property theft. We include the age at which the respondent first smoked cigarettes (which is negatively correlated with crime) and other sex-related variables, which are insignificant in this regression.[25] We also include two variables relating to violent behavior: having fought at school or work in the last year and having attacked someone with the intent to injure or kill. Both of these variables are strongly correlated with criminal activity, but including these variables reduces the SMSA effect on crime by only 10 percent.

In part B of table 5, we show the results of controlling for city- and victim-level characteristics in the NCVS and the UCR. The effect of victim-level characteristics may reflect the difficulty of the crime;

[25] These variables are almost all significant if included on their own.

if criminals generally rob their neighbors, then the victim-level characteristics may be a proxy for the criminal's characteristics. Regression 1 shows that the connection between crime and city size declines by one-third (from .12 to .08) in the NCVS when we control for victim-level characteristics.

Regression 2 presents results from the UCR and includes a variety of standard demographics. With these controls, the coefficient on city size drops to .09 from .16.[26] The only highly significant control variable in regression 2 is the percentage of female-headed households, which has an extremely positive effect on the crime rate. The inclusion of these controls reduces the urban crime coefficient by 44 percent, which is close to the 33 percent drop in the elasticity of city size found when demographics are included in the NCVS victimization regression. A greater drop would be expected in the UCR since city demographics control for characteristics of both the criminals and the victims.

Regression 3 addresses the possibility that female-headed households are being caused by crime and not the reverse. Indeed, it is quite plausible that the high levels of crime among men induce women to avoid marriage. To handle this reverse causality issue, we instrument for female-headed households by using lagged values of AFDC benefits at the state level. Our identifying assumption is that the AFDC benefits should be expected to have a differential effect in cities with and without a significant group of poorer individuals. This assumption seems quite reasonable to us since these payments should not influence the marriage-related incentives for wealthier individuals. The first-stage regression that we estimate is

$$\begin{aligned}\text{Percentage female-headed household, 1982}\\ = \underset{(.014)}{.088} - \underset{(.000)}{.000} \times \underset{\text{level, 1970}}{\text{AFDC benefit}} + \underset{(.078)}{.24} \times \text{poverty rate, 1970}\\ + \underset{(.0005)}{.003} \times \text{AFDC benefits} \times \text{poverty rate,}\end{aligned}$$

$N = 633$, $R^2 = .45$; standard errors are in parentheses.

Poverty in 1970 is not an instrument. We allow it to have an independent effect on crime by including it in the second-stage regression. The actual instruments are level of AFDC benefits and the interaction of AFDC benefits with poverty. As we expected, the effect of AFDC benefits on the number of female-headed households is much higher if the level of poverty is higher (i.e., if the number of

[26] Controlling for demographics (i.e., percentage nonwhite) alone will create 90 percent of this drop.

TABLE 5

Cities and Characteristics

A. Probits from NLSY (N = 8,910)*

	Stolen Property <$50 in Past Year (1)	(2)	(3)
SMSA	.030	.060	.027
	(.010)	(.014)	(.015)
Age	.073	.073	.084
	(.041)	(.040)	(.062)
Age squared	−.002	−.002	−.002
	(.001)	(.001)	(.001)
Local unemployment rate	−.001	−.002	−.001
	(.007)	(.007)	(.010)
Family intact (mother and father)	−.024	−.024	−.008
	(.011)	(.011)	(.016)
Highest grade achieved	−.001	−.001	−.001
	(.004)	(.004)	(.007)
Male	.126	.127	.107
	(.008)	(.008)	(.013)
Black	−.043	−.042	−.028
	(.013)	(.014)	(.022)
Hispanic	−.028	−.026	.019
	(.019)	(.019)	(.033)
Highest grade mother achieved	.004	.004	.006
	(.002)	(.002)	(.003)
Mother worked when respondent was age 14	−.001	−.001	−.020
	(.009)	(.009)	(.013)
AFQT score	.001	.001	.001
	(.000)	(.000)	(.000)
Attends church ≥ once per month	−.028	−.029	.011
	(.009)	(.009)	(.013)
SMSA at age 14 but non-SMSA now		.061	
		(.019)	
SMSA now but non-SMSA at age 14		.025	
		(.014)	
Changed city/town of residence since birth		−.009	
		(.009)	
Number of times used marijuana in past year			.035
			(.003)
Age first smoked cigarette			.005
			(.002)
Age first had sexual intercourse			−.002
			(.003)
Has sex without birth control			−.001
			(.018)
Age started drinking ≥ once per week			−.001
			(.003)
Attacked someone with intent to injure in past year			.056
			(.021)
Fought at school or work in past year			.071
			(.016)
Regional dummies included?	yes	yes	yes
$1 - [\ln(L)/\ln(L_0)]$	.04	.04	.10

TABLE 5 (*Continued*)

B. UCR and NCVS

	Dependent Variable			
	NCVS:[†] Victim of a Crime (N = 8,328)	Log Serious Crimes per Capita (N = 634)		Log Serious Crimes Adjusted[§] (N = 634)
	(1)	(2)	(3)[‡]	(4)
Intercept		−4.22	−4.23	−4.39
		(.29)	(.29)	(.29)
Log(city population)	.08	.09	.09	.17
		(.02)	(.02)	(.02)
Percentage population below poverty	.03	.61	.46	.52
	(.01)	(.46)	(.69)	(.47)
Percentage housing owner occupied	−.05	−.12	−.13	−.12
	(.01)	(.15)	(.15)	(.16)
Percentage nonwhite	−.00	−.29	−.29	−.31
	(.01)	(.15)	(.15)	(.15)
Percentage with 4 years high school	.00	.23	.23	.30
	(.00)	(.27)	(.27)	(.27)
Percentage with 4 years college		−.41	−.41	−.43
		(.25)	(.25)	(.26)
Unemployment rate	.01	.93	.91	1.02
	(.01)	(.59)	(.59)	(.60)
Percentage female-headed household		3.37	3.39	3.49
		(.55)	(.55)	(.56)
Lives in building with ≥10 apartments	.03			
	(.01)			
Regional dummies included?	no	yes	yes	yes
Lagged percentage population below poverty (IV regression)			.16	
			(.53)	
R^2 (or 1 − log likelihood ratio)	.09	.42	.42	.48

Note.—Standard errors are in parentheses.

* For east of interpretation, probits show partial derivatives of Y with respect to $\mathbf{X}$ rather than the actual probit coefficient.

[†] For NCVS probit, the elasticity of ''victim'' with respect to city size is estimated from coefficients on eight city size dummies. (The NCVS provides only city size categories.) Also included in the NCVS probit but not shown are controls for age, male, married, and Hispanic. The following other variables are used: for percentage below poverty, in bottom 25 percent of sample for family income; for unemployment rate, had a job last week; and for percentage nonwhite, a dummy variable for black. Partial derivatives are shown rather than probit coefficients.

[‡] In regression 3, percentage female-headed household is treated as endogenous, and two instruments are used to estimate the coefficient on percentage female head of household. The instruments used are AFDC payments per recipient household in 1970 and AFDC payments times percentage of population below poverty line in 1970.

[§] Adjusted crime numbers are adjusted *by city size* for underreporting of crimes. Adjustment methodology uses crime and reporting data from the NCVS.

TABLE 6

Cities and Characteristics: Relationship between Size and Characteristics ($N = 634$)

A.

	Dependent Variable				
	Elasticity with Respect to City Size (1)	Regression of Log(Serious Crime) on City Size and Demographics (2)	Regression of Log(Serious Crime Adjusted) on City Size and Demographics* (3)	Col. 1 × Col. 2 (4)	Col. 1 × Col. 3 (5)
Percentage population below poverty	.02	.61	.52	.01	.01
	(.00)	(.46)	(.47)		
Percentage housing owner occupied	−.04	−.12	−.12	.00	.00
	(.01)	(.15)	(.16)		
Percentage nonwhite	.07	−.29	−.31	−.02	−.02
	(.01)	(.15)	(.15)		
Percentage with 4 years high school	−.01	.23	.30	.00	.00
	(.01)	(.27)	(.27)		
Percentage with 4 years college	−.00	−.41	−.43	.00	.00
	(.00)	(.25)	(.26)		
Unemployment rate	.00	.93	1.02	.00	.00
	(.00)	(.59)	(.60)		
Percent family-headed household	.02	3.37	3.49	.07	.07
	(.00)	(.55)	(.56)		
Sum				.06	.06

B.

Dependent Variable	Controls Used	Elasticity with Respect to City Size	Elasticity-Adjusted Data with Respect to City Size†
1. Log serious crimes per capita	None	.18	.25
		(.02)	(.02)
2. Log serious crimes per capita‡	Region dummies	.15	.23
3. Log serious crimes per capita	Demographics and region dummies	.09	.17
		(.02)	
Difference (row 2 − row 3)		.06	.07

NOTE.—In part A, col. 1 shows the elasticity of the **X** variable with respect to city size; col. 2 shows the elasticity of crime with respect to the **X** variable; and col. 4 shows the product that represents that portion of the crime–city size elasticity that is associated with the **X** variable.

* Adjusted for underreporting of crimes by city size. The elasticity of the **X** variable with respect to city size is calculated by regressing **X** on log(population).

† Adjusted for underreporting of crimes by city size.

‡ Controls for region dummies and fixes coefficients on region dummies and intercept to be equal to coefficients in part B of table 5, regressions 2 and 4.

people who might potentially be influenced by AFDC levels is greater, then the AFDC effect is greater).

Regression 3 reports the second-stage results from our estimates and shows that with our instruments the coefficient on the percentage of female-headed households is basically unchanged. This should not be interpreted as meaning that the percentage of female-headed households can be seen as an exogenous variable. This variable, and our instruments, may be correlated with other related urban pathologies (perhaps the lack of strong social sanctions) that are actually driving the results. The instrumental variables estimates suggest only that our results do not come about because of reverse causality; they do not eliminate the possibility of related omitted variables.

Regression 4 adjusts the crime rate data for underreporting. Controlling for the demographics in this case reduces the connection between city size and the crime rate by 29 percent. One reason why the reduction is lower is that we are correcting for underreporting in larger cities, but we are not controlling for any possible extra underreporting in cities with a higher percentage of single-parent families.

Table 6 shows the separate effect of including each variable on the connection between crime and city size. The extent to which city-level demographics explain the amount of crime should not be surprising. All of the effect comes from controlling for the share of households that have a female head. The overall effect of this force appears to be .07. The net total effect of city-level controls is .06.

VII. Conclusion

Table 7 summarizes the results of this exercise. In all cases, we use data adjusted for underreporting. In the first row of the table, we begin with the raw elasticity to be explained. In the second and third rows, we investigate the importance of lower arrest rates in cities. In both cases, we use a value of $-.10$ for ϵ_N^Q. When we use the Levitt estimate of .2 for ϵ_P^Q, deterrence explains 8.33 percent of the city-crime connection. When we use the higher elasticity of .5, deterrence explains 20.8 percent of the city-crime connection.

In the fourth and fifth rows, we show the estimated importance of higher returns to crime in cities. In both cases, we use a value of .11 (midway between the UCR and the NCVS estimates) and a value of 1.45 for B/PC. When we use an estimate of .2 for ϵ_P^Q, higher returns to crime explain 13.33 percent of the city-crime connection. When we use the higher elasticity of .5, deterrence explains one-third of the city-crime connection. Individual characteristics explain

TABLE 7

Implementing the Decomposition

	Effect	Percentage of City Size–Crime Connection Explained by Effect
Initial city size–crime connection	.24	
Effect of deterrence:		
$\epsilon_P^Q = -.2$	.02	8.33
$\epsilon_P^Q = -.5$	.05	20.8
Effect of pecuniary returns:		
$\epsilon_P^Q = -.2$	.032	13.33
$\epsilon_P^Q = -.5$	.080	33.33
Effect of city composition	.07	29.2
Unexplained city size–crime connection:		
$\epsilon_P^Q = -.2$		49.14
$\epsilon_P^Q = -.5$		16.67

Source.—Row 1: table 2, regression 2; row 2: Levitt (1998*b*) and table 3; row 3: Ehrlich (1973) and table 3; rows 4 and 5: table 4 and text; row 6: table 6; rows 7 and 8: residual city size–crime effect.

29.2 percent of the city-crime effect. Using the higher estimate of ϵ_P^Q, we explain 83.33 percent of the city-crime connection. Using the more conservative estimate of ϵ_P^Q, we explain slightly more than one-half of the urban-crime premium.

One primary point of this paper is that even though classic deterrence and returns to crime explanations of the level of crime are important in explaining the urban crime premium, other variables (particularly family structure) also matter. It is hoped that future research will focus more on understanding the link between female-headed households and crime and also on understanding why cities have so many single-parent families. In particular, it would be valuable to know whether urban environments just attract these families or whether urban environments actually create more single-parent families.

Appendix A

Description of Data Sets

National Crime Survey

The National Crime Victimization Survey is an ongoing survey administered by the Department of Justice's Bureau of Justice Statistics.[27] The origins of the survey can be traced back to the President's Commission on Law

[27] Information in this section is taken from the codebooks and abstracts from U.S. Department of Justice (1991).

Enforcement and the Administration of Justice, impaneled in 1965. At that time it was recognized that existing measures of crime focused on only those crimes actually reported to the police and that more comprehensive information could be useful to law enforcement officials and social scientists. The three primary objectives of the NCVS are to gather detailed information about the victims and consequences of crime, to estimate the proportion of crimes that are not reported to the police, and to permit comparisons over time among areas.

The NCVS is a study of personal and household victimization. The survey is administered by interviewers who visit households throughout the country and interview each person in the household who is over the age of 12. (Phone interviews may also be used for follow-up surveys.) Interviewers first administer the Basic Screen Questionnaires to all household members 12 years of age or older. These questions gather information about the characteristics of the respondent and the household. The questions then request information on whether or not any crimes were committed against the respondent or any members of the household in the preceding six months. The survey covers the following types of crimes (including attempts): rape, robbery, assault, burglary, larceny, and auto theft. Other crimes such as murder, shoplifting, and illegal gambling do not fit within the survey's victimization framework and are excluded.

If the household or any individual has been victimized, the interviewer will ask the questions on the Crime Incident Report. One such report is made for each crime. The report gathers detailed information on the crime including (but not limited to) time and place of occurrence; injuries suffered; medical expenses incurred; age, race, and sex of offender(s); and relationship of offenders to victim. The Bureau of Justice Statistics has performed a number of tests and pilots to maximize the accuracy of the responses. In particular, the time frame of six months (i.e., asking whether or not the respondent was victimized in the preceding six months) was chosen following several pretests of people's ability to recall crime incidents. The households in the sample are interviewed every six months for three years and are then rotated out of the sample. To avoid declining interest and cooperation or biased responses from repeated interviewing, households are not interviewed indefinitely.

The NCVS is not administered to a random sample of the U.S. population, and hence simple averages of victimization rates within the sample are not unbiased estimates of victimization rates in the United States. The NCVS contains sample weights that are adjusted to be representative of the U.S. adult population. We draw two samples from the NCVS data. In both cases we take only those data that come from heads of households. We believe that this increases the accuracy of reporting and reduces the chance of double counting a single crime that might be reported by both the head of the household and another family member. The dates of the two samples (1980 and 1989) were chosen to be close to the NLSY data on self-reports of criminal acts (1980).

Our first sample uses the entire sample from the 1989 third-quarter interviews. This sample contains responses from heads of household that have and have not been victimized in the past six months. We had 8,328 heads of household with nonmissing data, and we use these data to run probits of ‘‘was a victim’’ on key demographics including dummies for city size. Our second sample uses the incident-level file from the 1980 third-quarter interviews. These data exist only when the head of the household reported that a member of the household was victimized by a crime. These data contain detailed information on the crime including the value of any property taken and whether or not the victim knew the offender.

Uniform Crime Reports

The Uniform Crime Reports are among the most widely used and cited sources of data on crime in the United States.[28] The FBI administers the UCR data. The UCR program is a nationwide cooperative effort of over 16,000 city, county, and state law enforcement agencies. Law enforcement agencies covering 96 percent of the U.S. population participate in the UCR program. The program was begun in the 1920s when the International Association of Police Chiefs met to develop a uniform plan for reporting of crimes.

Each participating agency fills out and submits to the FBI monthly forms detailing reports of murder, rape, robbery, assault, burglary, larceny, and auto theft. The reports include details on the crime such as any weapons used, whether or not there was forcible entry, and, in the case of auto theft, the type of vehicle stolen. The reports also detail any arrests for these categories of crime. The arrest reports include information on the person arrested such as sex and age. Finally, the participating agencies also submit data on the number of full-time law enforcement personnel.

The FBI reports the data in a variety of publications and tables. For example, there are annual tables by crime for the entire country. There are also tables by individual city and individual SMSA. The arrest reports are broken out by type of crime, gender, and age of the arrested person. The UCR data are complementary to the NCVS data in that the UCR data are crime statistics gathered from the point of view of law enforcement agencies whereas the NCVS data are gathered from the point of view of the victims. The UCR data count only crimes that are reported to the police. This has the weakness that it does not measure unreported criminal activity, and the amount of unreported activity presumably can vary widely over time, by type of crime, by size of city, and by demographics of the victim. However, the strength of the UCR approach is that gathering the data through the police might increase accuracy because of detailed police reports that are filled out at the time of the report or the ability of the police to filter out false or mistaken reports. Conversely, the NCVS seeks to avoid the nonreporting bias by asking households directly whether or not they were victimized.

Our sample of the UCR is taken from the reports aggregated by city. We have data for 634 cities, which are the U.S. cities with populations greater

[28] Information in this section is taken from various issues of the FBI’s *Crime in the United States*.

than 25,000 for which we have complete data. The year 1982 was chosen to be close to the year for which we have the self-reports of criminal activity in the NLSY (1980). We took UCR data on crimes, arrests, and value taken per crime and merged these with data from the *County and City Data Book*, aggregated from census data, containing demographic information for each city.

National Longitudinal Survey of Youth

The National Longitudinal Survey of Youth is one of the National Longitudinal Surveys sponsored by the Bureau of Labor Statistics, U.S. Department of Labor.[29] These surveys take a cohort of Americans and follow the members of the cohort over time to learn about the members' labor market experiences, places of residence, education, children, and many other items of interest. The NLSY data currently available are taken from a cohort of men and women who were aged 14–22 in 1979. There are 12,686 people in this cohort: 6,111 of the cohort were chosen to be a representative sample of young people living in the United States in 1979, and an additional 5,295 are the ''poverty oversample,'' which is designed to oversample economically disadvantaged youth. Finally, there is a military oversample of 1,280 youth who were aged 17–21 in 1979 and who were enlisted in the four branches of the military at that time.

Each time the cohort is surveyed, some questions are changed and some questions are mixed into the survey. The 1980 survey contained a large number of questions on illegal activity, including crimes committed, income from illegal activity, illegal drug use, and contact with the police. We use these self-reports of criminal activity as the dependent variables in our analysis. In particular, we run probits for whether or not the respondent reports having stolen property worth less than $50 in the past year and whether or not the respondent reports having fought at work or school in the past year. We look at small thefts for two reasons: First, we believe that people may be more likely to accurately self-report minor crimes than major ones. Second, for identification we wanted to examine a crime that a nontrivial number of respondents admit committing. We have about 4,000 observations for which we have both the response to ''stolen something less than $50'' and the regressors such as SMSA/non-SMSA, age, family intact, race, mother's education, and so forth. Nineteen percent of our sample reports having stolen something worth less than $50 in the past year.

References

Archer, Dane, and Gartner, Rosemary. *Violence and Crime in Cross-National Perspective*. New Haven, Conn.: Yale Univ. Press, 1984.

Becker, Gary S. ''Crime and Punishment: An Economic Approach.'' *J.P.E.* 76 (March/April 1968): 169–217.

[29] Information in this section is taken from the Bureau of Labor Statistics web site, stats.bls.gov.

Blumstein, Alfred, and Cohen, Jacqueline. ‘‘Estimation of Individual Crime Rates from Arrest Records.’’ *J. Criminal Law and Criminology* 70 (Winter 1979): 561–85.

Case, Anne C., and Katz, Lawrence F. ‘‘The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths.’’ Working Paper no. 3705. Cambridge, Mass.: NBER, May 1991.

Chaikin, J. ‘‘Rand Prison Inmate Survey.’’ Working Paper no. WN-10107-DOJ. Santa Monica, Calif.: Rand Corp., 1978.

DiIulio, John J., Jr., and Piehl, Anne Morrison. ‘‘Does Prison Pay? The Stormy National Debate over the Cost-Effectiveness of Imprisonment.’’ *Brookings Rev.* 9 (Fall 1991): 28–35.

Donohue, John J., III, and Siegelman, Peter. ‘‘Allocating Resources among Prisons and Social Programs in the Battle against Crime.’’ *J. Legal Studies* 27 (January 1998): 1–43.

Ehrlich, Isaac. ‘‘Participation in Illegitimate Activities: A Theoretical and Empirical Investigation.’’ *J.P.E.* 81 (May/June 1973): 521–65.

———. ‘‘The Deterrent Effect of Capital Punishment: A Question of Life and Death.’’ *A.E.R.* 65 (June 1975): 397–417.

Flango, Victor E., and Sherbenou, Edgar L. ‘‘Poverty, Urbanization, and Crime.’’ *Criminology* 14 (November 1976): 331–46.

Freeman, R. ‘‘Crime and the Employment of Disadvantaged Youths.’’ In *Drugs, Crime, and Social Isolation: Barriers to Urban Opportunity*, edited by Adele Harrell and George E. Peterson. Washington: Urban Inst. Press, 1992.

Glaeser, Edward L. ‘‘Cities and Ethics: An Essay in Honor of Jane Jacobs.’’ Discussion paper. Cambridge, Mass.: Harvard Inst. Econ. Res., 1998.

Glaeser, Edward L., and Sacerdote, Bruce. ‘‘Why Is There More Crime in Cities?’’ Manuscript. Cambridge, Mass.: Harvard Univ., 1996.

Glaeser, Edward L.; Sacerdote, Bruce; and Scheinkman, José A. ‘‘Crime and Social Interactions.’’ *Q.J.E.* 111 (May (1996): 507–48.

Grogger, Jeffrey. ‘‘Certainty vs. Severity of Punishment.’’ *Econ. Inquiry* 29 (April 1991): 297–309.

———. ‘‘The Effect of Arrests on the Employment and Earnings of Young Men.’’ *Q.J.E.* 110 (February 1995): 51–71.

Jacobs, Jane. *The Death and Life of Great American Cities*. New York: Random House, 1961.

Lane, Roger. *Violent Death in the City: Suicide, Accident, and Murder in Nineteenth Century Philadelphia*. Cambridge, Mass.: Harvard Univ. Press, 1979.

Larson, Calvin J. *Crime, Justice, and Society*. Bayside, N.Y.: General Hall, 1984.

Levitt, Steven D. ‘‘The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation.’’ *Q.J.E.* 111 (May 1996): 319–52.

———. ‘‘Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime.’’ *A.E.R.* 87 (June 1997): 270–90.

———. ‘‘The Response of Crime Reporting Behavior to Changes in the Size of the Police Force: Implications for Studies of Police Effectiveness Using Reported Crime Data.’’ *J. Quantitative Criminology* 14 (February 1998): 62–81. (*a*)

———. ‘‘Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?’’ *Econ. Inquiry* 36 (July 1998): 353–72. (*b*)

Mathieson, Donald, and Passell, Peter. ‘‘Homicide and Robbery in New York City: An Economic Model.’’ *J. Legal Studies* 5 (January 1976): 83–98.

Milgram, Stanley. ‘‘The Experience of Living in Cities.’’ *Science* 167 (March 13, 1970): 1461–68.

Pirenne, Henri. *Medieval Cities: Their Origins and the Revival of Trade*. Princeton, N.J.: Princeton Univ. Press, 1929.

Putnam, Robert D. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, N.J.: Princeton Univ. Press, 1993.

Radzinowicz, Leon, and Wolfgang, Marvin E., eds. *Crime and Justice*. 2d ed. New York: Basic Books, 1977.

Schichor, David; Decker, David L.; and O’Brien, Robert M. ‘‘Population Density and Criminal Victimization: Some Unexpected Findings in Central Cities.’’ *Criminology* 17 (August 1979): 184–97.

Spelman, William. *Criminal Incapacitation*. New York: Plenum Press, 1994.

Taylor, John B. ‘‘Econometric Models of Criminal Behavior: A Review.’’ In *Economic Models of Criminal Behavior*, edited by J. M. Heineke. Amsterdam: North-Holland, 1978.

U.S. Department of Justice. Bureau of Justice Statistics. *National Crime Surveys: National Sample, 1973–1995*. Computer file. 4th ed. Ann Arbor, Mich.: Inter-university Consortium Polit. and Soc. Res., 1991.

Weston, Paul B., and Wells, Kenneth M. *Criminal Investigation: Basic Perspectives*. Englewood Cliffs, N.J.: Prentice-Hall, 1994.

Wilson, James Q., and Herrnstein, Richard J. *Crime and Human Nature*. New York: Simon and Schuster, 1985.

Wirth, Louis. ‘‘Urbanism as a Way of Life.’’ *American J. Sociology* 44 (July 1938): 1–24.