REEXAMINING CRIMINAL BEHAVIOR: THE IMPORTANCE OF OMITTED VARIABLE BIAS

David B. Mustard*

Abstract—Recently many papers have used the arrest rate to measure punishments in crime-rate regressions. However, arrest rates account for only a portion of the criminal sanction. Conviction rates and time served are theoretically important, but rarely used, and excluding them generates omitted variable bias if they are correlated with the arrest rate. This paper uses the most complete set of conviction and sentencing data to show that arrest rates are negatively correlated with these normally excluded variables. Consequently, previous estimates of arrest-rate effects are understated by as much as 50%. Also, conviction rates, but not sentence lengths, have significant explanatory power in standard crime-rate regressions.

I. Introduction

To what extent do changes in the expected penalty alter the willingness of individuals to engage in illegal activity? Becker (1968) developed a theoretical model of criminal behavior that asserted offenders respond to incentives in much the same way as those who engage in legal activities respond. Ehrlich (1973) pioneered empirical tests of the model and concluded that there was a strong positive correlation between income inequality and property crimes, that the crime rate was positively related to relative gains and negatively related to the costs associated with criminal activity, and that law enforcement activity had a deterrent effect for all crimes.

Recently many papers have used arrest rates to measure illegal sanctions. Cornwell and Trumbull (1994), Lott and Mustard (1997), Black and Nagin (1998), Lott (1998), Bronars and Lott (1998), Dezhbakhsh and Rubin (1998), Levitt (1998), Dezhbakhsh, Rubin, and Shepherd (2001), Gould, Weinberg, and Mustard (2002), Grinols and Mustard, (2001), and Shepherd (2002a, 2002b) showed that arrest rates are important explanatory variables in crime-rate regressions. However, arrest rates account for only a portion of the penalty. Conviction rates and time served are theoretically important, but rarely used, and estimates that exclude them will suffer from omitted variable bias if the omitted variables are correlated with the arrest rate. Conviction and sentencing data are rarely included because they are very difficult to obtain. Although the FBI collects offense and arrest data for the Index I offenses,[1] there is no comprehensive source of conviction or sentence length data. By contacting the individual states directly I obtained county-level conviction and sentencing data for four states.

Although many studies have examined how arrest rates affect crime rates, few have taken account of conviction rates and sentence lengths, essential components of punishment. Only a few studies, such as Ehrlich (1973), Cornwell and Trumbull (1994), and Lott and Mustard (1997), have included conviction and sentencing data. Ehrlich (1973) used repeated cross-section data at the state level. Cornwell and Trumbull (1994) used only North Carolina data and aggregated all seven FBI Index I offenses into one measure. Lott and Mustard (1997) used the measures only in a county-level panel for Oregon.

By analyzing the most comprehensive conviction and sentencing data, this paper provides new evidence about criminal behavior and sanctions and a more complete assessment of the penalty associated with illegal activity. If arrest rates are positively correlated with the omitted variables, previous studies overstated the effect of arrest rates. If they are negatively correlated, estimates of the criminal sanction were understated. Because crime rates and responses to crime vary significantly at the local level and over time, I use annual cross-sectional time series data at the county level from 1977 to 1992 to examine this more complete model of crime.

This paper is organized as follows. Section II presents the theory. Section III explains the data and model. Section IV contains the empirical work, and Section V concludes.

II. Theory

This section briefly overviews the economic theory of crime and expands that theory to consider the omitted variable bias when conviction rates and sentence lengths are not utilized in the standard crime-rate regressions.

The illegal-sector equilibrium is determined by the supply and derived demand for crime. Individuals who decide whether to participate in the illegal sector determine the supply of crime. Their propensity to engage in criminal activity is a function of the expected penalty of criminal action, including the probability of being arrested; the conditional probability of being convicted given arrest;

Received for publication July 11, 2000. Revision accepted for publication January 11, 2002

* Terry College of Business, University of Georgia.

I thank Gary S. Becker, Christopher Cornwell, Steven Levitt, John R. Lott, Jr., Kevin M. Murphy, Derek A. Neal, two anonymous referees, and workshop participants at the University of Chicago, the University of Georgia, Clemson University, and George Mason University for their many helpful suggestions. I have received financial support from the Jacob Javits Foundation and the Bradley Foundation. The original collectors of the data, the ICPSR, the Jacob Javits Foundation, and the Bradley Foundation bear no responsibility for the use of this collection or for the interpretations or inferences based upon the data.

[1] Index I offenses include murder, rape, robbery, aggravated assault, burglary, larceny, and auto theft.

The Review of Economics and Statistics, February 2003, 85(1): 205–211

monetary fines; time served in prison, probation, and parole; and the loss of reputation, which has both social and economic components.[2] People form expectations about legal and illegal opportunities based on their information, and evaluate the expected utility given their risk preferences for entry into the illegal and legal sectors. An individual will commit a crime if his expected utility from the crime exceeds his expected utility from using his time and resources in legal activities.

Derived demand comes from both public and private expenditures for crime reduction. The government affects the derived demand by changing the resources devoted to apprehending, convicting, and incarcerating criminals through hiring police, prosecutors, and correctional officers. Private individuals and groups also affect the derived demand.[3] They may increase their security by installing security systems, purchasing cellular phones and dogs, carrying concealed weapons[4] and installing Lojack,[5] and making choices about occupation, transportation, and residential and vocational location. The standard economic theory of crime neither discusses the relationship between the arrest rates and conviction rates and sentencing, nor indicates the direction of the omitted variable bias.

Arrest rates may be either positively or negatively related to conviction rates and sentence lengths; the net effect is an empirical question. If decision-makers are concerned about being tough on crime, they may increase all these sanctions. Arrest rates may also be negatively correlated with conviction rates and sentence length. As more people are arrested for a given offense level, the marginal quality (the quality and amount of evidence against an offender) of arrestees decreases. Law enforcement officials devote more resources to arresting offenders who cause the most social damage (for example, murderers do more social harm than larcenists) and those who have the most evidence against them (and the highest probability of conviction). As agencies arrest more people, the marginal offender has done less social harm and has a lower probability of conviction and lower expected sentences. Therefore, counties that arrest relatively few will arrest, on average, those with higher probabilities of conviction and longer expected sentences, thus generating a negative correlation between arrest rates and conviction rates and sentence lengths.

III. Data and Model

The economic model of crime is frequently estimated by[6]

$$\ln\left(\frac{o}{N}\right)_{jkt} = \alpha_j + \beta_j \frac{a_{jkt}}{o_{jkt}} + \gamma_j \frac{c_{jkt}}{a_{jkt}} + \lambda_j S_{jkt} + \pi_j P_{kt} + \phi_j I_{kt} + \delta_{jk} D_{jk} + \tau_{jt} T_{jt} + \varepsilon_{kt}, \tag{1}$$

where $(o/N)_{jkt}$ is the number of per capita offenses in county k in year t for each crime category j: murder, rape, robbery, aggravated assault, burglary, larceny, and auto theft. The arrest rate for crime j, a_{jkt}/o_{jkt},[7] is the number of arrests divided by offenses reported in year t.[8] The annual conviction rate in county k is c_{jkt}/a_{jkt}, the number of convictions divided by the number of arrests. S_{jkt} is the average prison sentence for crime j in county k at time t, and P_{kt} is a vector of population, age, race and gender variables. The population is divided by race (white, black, and neither), gender, and 10-year age intervals. I_{kt} is a vector of real per capita income variables and contains personal income, insurance payments, income maintenance, and retirement payments. County fixed effects, D_{jk}, control for unobserved county-specific factors that affect the markets for crime, such as policing methods, law enforcement practices, capital punishment laws (Ehrlich, 1975; Dezhbakhsh et al., 2001),[9] and shall-issue laws (Lott & Mustard, 1997).[10] Time fixed effects, T_{jt}, control for changes in

[2] Lott (1990, 1992) argued that the reputational effects of arrest and conviction are extremely large for those of highest income.

[3] Philipson and Posner (1996) argued that U.S. private crime-reducing expenditures are \$300 billion annually, significantly greater than public expenditures.

[4] Lott and Mustard (1997) showed that states that pass laws that allow citizens to carry concealed weapons experience decreases in violent crime of between 2% and 7%.

[5] Ayres and Levitt (1998) showed that metropolitan areas that introduce Lojack, a concealed automobile tracking system, substantially lower their auto theft rates.

[6] See Dezhbakhsh et al. (2001) for a more detailed discussion of the most commonly used specifications: double-log, trans-log, and levels. I estimated the regressions for all three of these specifications, but report only the results for equation (1). The unreported results are qualitatively the same as the reported results and strongly support the conclusion that omitting conviction and sentencing data biases the arrest-rate coefficients down.

[7] For some observations the number of arrests was greater than the number of offenses. Arrests sometimes exceeded offenses because there may have been multiple offenders arrested for one crime, or offenses may have occurred in one year and the arrests in the following year. The same is true for convictions and arrests.

[8] Levitt (1998) showed that the estimated coefficient of the arrest rate in the standard crime rate regression can be biased when the offenses are reported with measurement error, and that this problem can be mitigated by including a lagged version of the arrest rate. I include in all regressions a once lagged arrest rate. When the lagged arrest rate is omitted from the regressions, the coefficient estimates on the arrest rate variables are generally more negative.

[9] New York (which reinstituted the death penalty in 1995 but hasn't executed anyone since 1963), Washington (which reinstituted it in 1975 and has only executed four people since then), and Oklahoma (which reinstituted it in 1973 and has executed 45 people since then) had no change in the legal status of the death penalty over the period of this data set. Therefore, the county fixed effects would capture the differences across counties. Oregon reinstituted capital punishment in 1978, the second year of the sample in this study. However, Oregon did not execute anyone until 1996 and has executed only two people since 1996. (The information in this footnote is from Joanna Shepherd (personal email) and from the Death Penalty Information Center, http://www.deathpenaltyinfo.org).

[10] Similar to the death penalty as discussed in the preceding footnote, concealed-carry laws did not change during this sample in New York, Oklahoma, and Washington, and therefore the effects of these laws are taken into account by county fixed effects. Oregon implemented a shall-

crime-related issues across time. The coefficient estimates for each variable are provided for each crime type j.

The most significant difference between this paper and others is that this one includes c_{jkt}/a_{jkt} and S_{jkt}. Conviction and imprisonment are important components of the total penalty of the illegal sector, and not allowing for them could bias previous estimates. To determine the extent of the omitted variable bias, I estimate equation (1) and then rerun the regression excluding the normally omitted variables. The coefficient estimates β_j, γ_j, and λ_j are expected to be negative for two reasons—higher sanctions decrease crime by incapacitating existing offenders and by deterring future offenders. Because the central purpose of this paper is to analyze the effects of the omitted variable bias, I do not discuss the income and population variables, which I use as controls rather than to estimate their effects.[11]

I analyze annual county-level panel data for seven crimes: aggravated assault, murder, forcible rape, robbery, burglary, larceny, auto theft.[12] The time series runs from 1977 to 1992, the last year in which all four states can provide the data. The Federal Bureau of Investigation (FBI) classifies the first four offenses as violent crimes and the last three as property crimes. The sum of the violent and property crimes is the number of Index I offenses. The Uniform Crime Report (UCR) provides the number of reported offenses and arrests for each FBI Index I crime for each county.[13]

To obtain the conviction and sentencing data I contacted the Departments of Corrections, Bureaus of Justice, and Statistical Analysis Centers in every state and the District of Columbia. Only four states—New York, Oklahoma, Oregon, and Washington—were able to provide county-level data.[14] These states have 62, 77, 36 and 39 counties, respectively, for a total of 214 counties. New York and Oregon provided the data for 1977–1992, Oklahoma provided the data for 1980–1992, and Washington provided the data for 1985–1992. New York was unable to provide conviction or sentence information for auto theft. The Oregon sentencing data for 1990–1992 are inconsistent with prior years, because the state instituted determinate sentencing that required offenders who committed a crime after November 1, 1989, to serve at least 85% of their sentence. Consequently, the Oregon data show large decreases in the sentence lengths after 1989, because the judges decreased the sentences to compensate for the new minimum-stay requirements.[15] New York and Oklahoma had indeterminate sentencing for the entire period of this study.[16] Washington changed from an indeterminate to a determinate system in 1985, the first year for which that state could provide data.

The data are summarized in table 1.

IV. Empirical

This section presents the regression results corresponding to equation (1) and shows that the standard practice of excluding conviction and sentencing data biases the coefficient on the arrest rate towards zero.

A. *County-Level Regressions*

Table 2 shows the weighted least squares[17] regression results of equation (1). The control variables are included in the regression but not reported and include the percentages of the population in various age groups (age 10–19, age 20–29, age 30–39, age 40–49, age 50–64, and over 65 years), percentages of the population in various demographic groups (male, black, and neither white nor black), population per square mile, real per capita measures (income, unemployment compensation, income maintenance, and retirement transfers), and county and year fixed effects.

As expected, the coefficient estimates of the arrest rates are negative and statistically significant for all seven crimes.[18] These estimates are greater for the property offenses of burglary, larceny, and auto theft than for the violent crimes, and the arrest-rate estimates vary from -0.0016 to -0.012; a one-percentage-point increase in the arrest rate is associated with a drop between 0.16% and 1.2% in the different crime rates. A one-standard-deviation change in the arrest rates accounts for between 6.9% and 50.6% of a one-standard-deviation change in the respective crime rates, with assault accounting for the smallest change and auto theft for the largest. The lagged arrest rates are also negative and statistically significant in five of the seven examples and are essentially 0 for murder and auto theft.

The coefficient estimates on the conviction rates are all negative. Five of the seven are statistically significant at the 0.10 level, and the other two (robbery and burglary) are marginally significant at 0.13 and 0.12, respectively. The coefficient estimates on the conviction rate vary from

issue law for the last two years of this sample. This paper's results are robust to allowing for Oregon's shall-issue law.

[11] Lott and Mustard (1997) thoroughly discussed the expected and estimated effects of these variables.

[12] Although arson was included as an Index I crime in 1979, it is excluded from this study because many agencies do not report it and there are many inconsistencies in the data even when they are reported.

[13] Property crime accounts for approximately 90% of all Index I crimes. Larceny and burglary account for approximately 55% and 25%, respectively.

[14] Cornwell and Trumbull (1994) used North Carolina county-level data. North Carolina officials initially agreed to give me the data, but in subsequent calls indicated the data do not exist.

[15] To address this I interact the sentence length with a dummy variable for the years after the sentencing reform. I also ran regressions that omitted those years and obtained the same qualitative results.

[16] On June 10, 1995, New York made effective the Sentencing Reform Act of 1995, which incorporated many determinate sentencing components into New York's sentencing laws.

[17] I estimated equation (1) by weighted least squares, with the county population as the weight. Unweighted estimates would produce heteroskedasticity, because the magnitude of the error terms is negatively correlated with the population size. The crime, arrest, and conviction rates fluctuate more in the low-population counties because these counties have relatively small crime rates, and small changes in the numbers of offenses, arrests, and convictions generate large changes in the ratios.

[18] The standard errors reported in the tables allow for correlations across counties within states.

TABLE 1.—SUMMARY STATISTICS

Variable	Number	Mean	St. Dev.
Offense rates (offenses per 100,000 people)			
Violent crimes	2881	301.52	335.72
Murder	2881	5.22	7.07
Rape	2881	23.51	22.33
Robbery	2881	71.82	196.18
Aggravated assault	2881	200.97	158.87
Property crimes	2881	3481.21	1589.32
Burglary	2881	1032.97	476.75
Larceny	2881	2213.76	1062.74
Auto theft	2881	234.48	256.51
Log offense rates			
Violent crimes	2881	5.35	0.97
Murder	2881	0.40	1.98
Rape	2881	2.38	1.83
Robbery	2881	2.75	2.19
Aggravated assault	2881	4.96	1.11
Property crimes	2881	8.05	0.47
Burglary	2881	6.84	0.46
Larceny	2881	7.57	0.54
Auto theft	2881	5.12	0.88
Arrest rates			
Violent crimes	2864	56.92	33.33
Murder	1962	96.86	88.06
Rape	2557	52.72	50.84
Robbery	2526	61.14	154.35
Aggravated assault	2845	59.67	40.54
Property crimes	2881	18.79	8.562
Burglary	2881	17.63	11.10
Larceny	2881	18.91	9.68
Auto theft	2871	26.70	85.69
Conviction rates			
Violent crimes	2826	15.80	26.37
Murder	1872	68.92	82.22
Rape	2345	47.77	114.19
Robbery	2208	32.23	46.38
Aggravated assault	2805	6.62	17.31
Property crimes	1867	8.32	17.36
Burglary	2833	16.12	32.96
Larceny	2837	3.42	13.22
Auto theft	1758	20.21	40.01
Mean sentence length (mo.)			
Violent crimes			
Murder	1620	198.40	248.77
Rape	1566	125.55	164.21
Robbery	1720	107.20	178.84
Aggravated assault	1738	80.33	230.57
Property crimes			
Burglary	2487	50.68	40.44
Larceny	1904	41.36	54.35
Auto theft	1232	53.40	51.10
Population variables			
County population	2881	141406.79	320000.87
Age $<$ 9 (%)	2881	14.68	1.53
Age 10–19 (%)	2881	15.73	2.02
Age 20–29 (%)	2881	14.90	3.28
Age 30–39 (%)	2881	14.61	1.88
Age 40–49 (%)	2881	11.24	1.52
Age 50–64 (%)	2881	14.62	1.88
Age $>$ 65 (%)	2881	14.32	3.47
Black (%)	2881	3.27	5.37
White (%)	2881	92.07	8.14
Other (%)	2881	4.68	5.66
Male (%)	2881	49.25	1.22
Density (per sq. mi.)	2881	908.59	5137.74
Area (sq. mi.)	2881	1290.10	1406.71

TABLE 1.—(*Continued*)

Variable	Number	Mean	St. Dev.
Income variables (real/1983 $)			
Real PC income	2881	11221.70	2674.54
Real PC un. comp.	2881	79.63	47.91
Real PC income maint.	2881	165.44	82.41
Real retirement PCO	2881	13185.78	2641.16

When counties have a greater number of arrests than reported offenses for a given year, an arrest rate of 100% is assigned. When counties have a greater number of convictions than arrests for a given year, a conviction rate of 100% is assigned. The data include the following states and years: NY (1977–1992), OK (1980–1992), OR (1977–1992), WA (1985–1992).

−0.0006 for burglary to −0.0076 for larceny. A one-percentage-point increase in the conviction rate is associated with a drop between 0.06% and 0.76% in the crime rates. A corresponding change in the conviction rates accounts for between 4.3% (burglary) and 18.6% (larceny) of a one-standard-deviation change in the respective crime rates.

The sentence length presents less convincing results—it appears to have no effect on crime rates. All the sentence-length coefficients are positive, close to 0, and statistically insignificant. A one-standard-deviation change in the sentence length accounts for less than 8.5% of a one-standard-deviation change in all the crime rates. One interpretation of the sentencing results is that the sentence length has no effect on the crime rate. A second interpretation is that the sentencing data contain measurement error that biases the results towards zero. Ideally the sentence length would measure the degree to which a county becomes more or less punitive over time. Sentence lengths are measured with imprecision for several reasons. First, sentence length does not measure time served.[19] Although sentence length is correlated with time served, it overstates it because of early release and parole. Therefore, the actual time served will be overestimated and its regression coefficient will be biased downward.

More important, however, Mustard (2001) showed that even for a given offense type, the sentence length is substantially determined by the severities of the offenses committed and the offenders' criminal histories, neither of which are available at the state or county level. For example, if county A has a greater percentage of repeat offenders than county B, then county A may have a longer observed average sentence length even though it may punish criminals less severely for a given offense. Because the sentencing data do not control for offense severity and criminal

[19] Sentence length and time served, the two most common measures for the cost of conviction, have relative strengths and weaknesses. One advantage of sentence length is that it measures the sentence at the time of the flow into prison, while time served measures it at the time of flow out of prison. Therefore, the determination of the sentence length is closer to the time individuals choose whether to participate in the illegal sector. Time served is biased to the extent that the penalty assessed on those who enter the system differs from the penalty assessed on those who leave prison. The bias is least for crimes with short periods of incarceration (burglary, auto theft, and larceny), and most severe for murder, which has the longest sentence length.

TABLE 2.—LOG OF CRIME RATES ON REGRESSORS, INCLUDING CONVICTION AND SENTENCING DATA

Variable	ln (Murder Rate)	ln (Rape Rate)	ln (Robbery Rate)	ln (Assault Rate)	ln (Burglary Rate)	ln (Larceny Rate)	ln (Auto Theft Rate)
Arrest rate	−0.0035*	−0.0026†	−0.0016*	−0.0019#	−0.0123#	−0.0072#	−0.0052#
	(0.00025)	(0.00091)	(0.00036)	(0.00049)	(0.00267)	(0.00178)	(0.00108)
	[15.6%]	[7.2%]	[11.3%]	[6.9%]	[29.7%]	[12.9%]	[50.6%]
Lagged arrest rate	0.0000	−0.0031#	−0.0035#	−0.0038*	−0.0102#	−0.0046†	0.0003
	(0.00034)	(0.00075)	(0.00076)	(0.00049)	(0.00304)	(0.00166)	(0.00080)
Conviction rate	−0.0028*	−0.0009†	−0.0025	−0.0061†	−0.0006	−0.0076#	−0.0023†
	(0.00023)	(0.00042)	(0.00119)	(0.0025)	(0.00027)	(0.00198)	(0.00068)
	[11.6%]	[5.6%]	[5.3%]	[9.5%]	[4.3%]	[18.6%]	[10.5%]
Sentence length	0.00002	0.00050	0.00064	0.00020	0.00116	0.00036	0.00007
	(0.00004)	(0.00028)	(0.00033)	(0.00012)	(0.00052)	(0.00030)	(0.00026)
	[0.3%]	[4.5%]	[5.2%]	[2.1%]	[8.3%]	[0.9%]	[0.4%]
Intercept	−0.46134	−1.03543	−11.59003	7.94230*	11.21289*	16.45943*	13.82592*
	(4.74427)	(3.32142)	(7.01062)	(2.90417)	(1.31315)	(1.31190)	(3.01113)
Number of obs.	1115	1385	1535	1692	2412	1853	1181
Adjusted R^2	0.9295	0.8243	0.9539	0.9078	0.8309	0.7528	0.8972

* $p < 0.01$. # $p < 0.05$. † $p < 0.10$.
Standard errors are in parentheses and allow for correlation across counties within a state.
The numbers in brackets represent the change in the endogenous variable (in percent of standard deviation) that can be explained by a one-standard deviation change in the exogenous variable.
Control variables included in the regression but not reported are: percentages of the population in various age groups (age 10–19, age 20–29, age 30–39, age 40–49, age 50–64, and over 65 years), percentages of the population in various demographic groups (male, black, and neither white nor black), population per square mile, real per capita measures (income, unemployment compensation, income maintenance, and retirement transfers), and county and year fixed effects.
All regressions are weighted by the county population.

history, the coefficient on the sentencing variable is biased towards 0, and the result that the sentence length does not affect crime is not necessarily a behavioral implication, but may be a result of a mismeasured variable.

B. Omitted Variable Bias

Because conviction rates and sentencing data are rarely studied, it is important to examine how the estimates of the arrest rates are affected by the inclusion of these data. The bias depends on the correlation between the arrest rate and the omitted variables. Parts A and B of table 3 present the correlation coefficients of sanctions for violent and property crimes, respectively. The left third of both tables indicates that arrest rates for different types of crimes are positively correlated with one another and negatively correlated with conviction rates and sentence lengths. Eight of ten violent-crime and all six property-crime arrest rates are positively correlated with each other. The only two negative arrest-rate correlations (murder-robbery and robbery-assault) are the smallest two of all 16 correlations. Of the 41 correlations between arrest and conviction rates (25 violent and 16 property), 34 are negative, and of 32 correlations between arrest rates and sentence lengths, 29 are negative. All 16 of the conviction rates are positively correlated with each other, and all 9 of the sentence-length correlations are positively correlated with each other.

The negative correlations between the arrest rates and the conviction rates and sentence lengths have two important implications. First, counties with high arrest rates tend to have low conviction rates, and vice versa, which presents some evidence that arrest rates are substitutes for conviction rates and sentences. Second, the negative correlation implies that estimates that exclude the conviction rates and sentence lengths will understate the true effect of the arrest rate on the crime rate.

Table 4 illustrates the second implication by running the same regression reported in table 2 but excluding the conviction rates and sentence lengths. Table 4 shows that when the conviction rates and sentence lengths are excluded, the arrest-rate coefficients are still statistically significant but are generally less negative. The new murder arrest-rate coefficient is 45.7% smaller than the one in table 2. The other coefficients are 23.1% smaller for rape, 0.6% smaller for robbery, 5.9% smaller for burglary, 9.7% smaller for larceny, and 48.1% smaller for auto theft. The one exception is assault, for which the new coefficient estimate is 0.0001 higher than in table 2.

Two robustness checks indicate that these reported results likely understate the bias—omitting the lagged arrest rate and using top-coded arrest and conviction rates both generally produce larger decreases than I document above. The crime literature has documented advantages and disadvantages of using raw rates (which, as discussed earlier, are sometimes larger than 100%) and rates top-coded at 100%. Using rates that exceed 100% can provide valuable information about arrest and conviction that we would not want to discard arbitrarily. However, using the raw data also creates a few outliers, especially in small counties with low crime rates, that have five or ten times as many arrests as offenses, or convictions as arrests. Using top-coded arrest and conviction rates produces larger decreases in the estimate of the arrest-rate coefficient (between the table 2 and table 4 specifications) in five of the seven cases, including robbery and assault, that previously showed the smallest

TABLE 3.—CORRELATIONS BETWEEN ARREST RATES, CONVICTION RATES, AND SENTENCE LENGTH

A. Violent Crimes

	Arrest Rate Viol.	Arrest Rate Murd.	Arrest Rate Rape	Arrest Rate Robb.	Arrest Rate Asslt.	Conv. Rate Viol.	Conv. Rate Murd.	Conv. Rate Rape	Conv. Rate Robb.	Conv. Rate Asslt.	Sent. Lenth Murd.	Sent. Lenth Rape	Sent. Lenth Robb.	Sent. Lenth Asslt.
AR Viol.	1.000													
AR Murd.	0.192	1.000												
AR Rap	0.183	0.087	1.000											
AR Robb.	0.146	−0.014	0.026	1.000										
AR Asslt.	0.800	0.095	0.344	−0.008	1.000									
CR Viol.	−0.255	−0.027	−0.161	−0.047	−0.250	1.000								
CR Murd.	0.055	−0.177	−0.119	−0.055	0.030	0.280	1.000							
CR Asslt.	0.024	0.023	−0.173	−0.025	0.007	0.607	0.134	1.000						
CR Robb.	−0.095	−0.060	−0.110	−0.064	−0.107	0.353	0.192	0.099	1.000					
CR Asslt.	−0.258	−0.062	−0.116	−0.052	−0.273	0.397	0.187	0.063	0.338	1.000				
SL Murd.	−0.097	−0.059	−0.141	−0.039	−0.074	0.138	0.182	0.060	0.095	0.203	1.000			
SL Rape	−0.170	−0.073	−0.128	−0.046	−0.148	0.092	0.122	−0.026	0.122	0.344	0.305	1.000		
SL Robb.	−0.117	−0.099	−0.191	−0.068	−0.101	0.128	0.177	−0.006	0.312	0.405	0.170	0.496	1.000	
SL Asslt.	0.073	−0.046	−0.213	−0.050	0.068	0.315	0.225	0.196	0.201	0.072	0.190	0.164	0.271	1.000

B. Property Crimes

	Arrest Rate Prop.	Arrest Rate Burg.	Arrest Rate Larc.	Arrest Rate Auto	Conv. Rate Prop.	Conv. Rate Burg.	Conv. Rate Larc.	Conv. Rate Auto	Sent. Length Burg.	Sent. Length Larc.	Sent. Length Auto
AR Prop.	1.000										
AR Burg.	0.660	1.000									
AR Larc.	0.918	0.376	1.000								
AR Auto	0.608	0.455	0.394	1.000							
CR Prop.	−0.266	0.022	−0.358	−0.037	1.000						
CR Burg.	−0.226	−0.186	−0.189	−0.138	0.409	1.000					
CR Larc.	−0.173	0.075	−0.334	0.180	0.515	0.333	1.000				
CR Auto	−0.201	−0.048	−0.189	−0.259	0.336	0.640	0.306	1.000			
SL Burg.	−0.177	−0.118	−0.132	−0.163	0.113	−0.091	−0.100	−0.090	1.000		
SL Larc.	−0.002	−0.002	0.008	−0.018	0.028	−0.040	0.057	−0.035	0.087	1.000	
SL Auto	−0.115	−0.104	−0.061	−0.152	0.058	0.147	0.040	0.147	0.117	0.088	1.000

change. The robbery results decrease only 0.5% with the raw rates, but decrease 36.9% with the top-coded rates. The assault results increase slightly with the raw rates, but decrease 7.7% with the top-coded rates.

Simultaneity may bias the coefficients on the arrest rate, conviction rate, and sentence length towards zero if these variables are increased in response to growing crime rates. Because it is extremely difficult to obtain variables that meet the identifying restrictions for valid instruments, only a few studies have addressed this issue.[20] The key finding of all these studies is that properly controlling for simultaneity increases the magnitudes of the coefficients on the variables of interest. Utilizing the conviction and sentencing data and running the regressions for each crime category effectively preclude finding enough instruments that meet the necessary identification restrictions. However, this does not affect the fundamental relationship between the arrest rates and the conviction and sentencing data typically omitted from other studies.

V. Conclusion

This paper uses the most exhaustive data on the complete measure of criminal sanctions—arrest rates, conviction rates, and sentence lengths. Because the latter two variables are extremely difficult to obtain, crime regressions typically exclude them, and therefore suffer from omitted variable bias. I extend the traditional theory of crime to examine the relationship between the arrest rates and omitted variables. Theoretically, arrest rates may be either negatively or positively correlated with the missing variables. However, the empirical work documents a negative relationship, consistent with the explanation that for a given offense level, as more people are arrested the marginal quality of the arrestees decreases because law enforcement officials devote greater resources to arresting individuals against whom they can make stronger cases or who impose greater social costs. This negative relationship indicates that studies that exclude the conviction rate and sentence length understate the true effect of the arrest rate on crime. Estimating similar empirical specifications without the conviction and sentencing data results in estimates of the effect of arrest rates up to

[20] Ehrlich (1973) used expenditure and employment variables, and Levitt (1996) used prison overcrowding legislation, as instruments. Unfortunately, they are only available at the state level, and only four states provide information about them as well as the conviction and sentencing data. Cornwell and Trumbull (1995) could use the ratio of violent to property crimes as an instrument only because they aggregated all seven into one total crime rate and did not estimate separate regressions for each offense. Electoral cycles can be used with Levitt's (1997) city data but not with county data, because there are many municipalities and election cycles within most counties. Also, conviction and sentencing data do not exist at the city level.

TABLE 4.—LOG OF CRIME RATES ON REGRESSORS, EXCLUDING CONVICTION AND SENTENCING DATA

Variable	ln (Murder Rate)	ln (Rape Rate)	ln (Robbery Rate)	ln (Assault Rate)	ln (Burglary Rate)	ln (Larceny Rate)	ln (Auto Theft Rate)
Arrest rate	−0.0019*	−0.0020#	−0.0016*	−0.0020#	−0.0117#	−0.0065†	−0.0027*
	(0.00015)	(0.00053)	(0.00004)	(0.00036)	(0.00288)	(0.00208)	(0.00004)
	[8.5%]	[5.6%]	[11.35%]	[7.3%]	[28.2%]	[11.7%]	[26.3%]
Lagged arrest rate	−0.0002	−0.0033#	−0.0035#	−0.0040#	−0.0098#	−0.0046†	0.0016
	(0.00078)	(0.00069)	(0.00077)	(0.00062)	(0.00289)	(0.00155)	(0.00116)
Intercept	8.3926	10.0538	10.8145	10.0302*	12.7851#	13.5110*	17.3827#
	(5.0349)	(7.0203)	(5.8167)	(1.6669)	(3.6378)	(1.3630)	(4.3773)
Number of Observations	1511	2290	2277	2720	2783	2783	2766
Adjusted R^2	0.8893	0.7428	0.9383	0.8734	0.8182	0.7277	0.9243

Notes: Same as for table 2.

50% lower than when conviction rates and sentence lengths are included.

Arrest and conviction rates have significant negative effects on crime rates, but sentence length has little effect. This conclusion does not necessarily imply a behavioral result, because the sentence length is measured with error, does not control for many important criteria like the offense level and criminal history, and is therefore biased towards zero.

REFERENCES

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

Becker, Gary, "Crime and Punishment," *Journal of Political Economy* 76:2 (1968), 169–217.

Black, Daniel, and Daniel Nagin, "Do Right-to-Carry Laws Deter Violent Crime?" *Journal of Legal Studies* 27:1 (1998), 209–220.

Bronars, Stephen G., and John R. Lott, Jr., "Criminal Deterrence, Geographic Spillovers, and the Right to Carry Concealed Handguns," *The American Economic Review* 88:2 (1998), 474–479.

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

Dezhbakhsh, Hashem, and Paul H. Rubin, "Lives Saved or Lives Lost? The Effects of Concealed-Handgun Laws on Crime," *The American Economic Review* 88:2 (1998), 468–474.

Dezhbakhsh, Hashem, Paul H. Rubin, and Joanna Mehlhop Shepherd, "Does Capital Punishment Have a Deterrent Effect? New Evidence from Post-moratorium Panel Data," Emory University working paper (2001).

Ehrlich, Isaac, "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy* 81:3 (1973), 521–565.

——— "The Deterrent Effect of Capital Punishment: A Question of Life and Death," *American Economic Review* 65:3 (1975), 397–417.

Gould, Eric, Bruce Weinberg, and David B. Mustard, "Crime Rates and Local Labor Market Opportunities in the United States: 1977–1997," *Review of Economics and Statistics* 84:1 (2002), 45–61.

Grinols, Earl L., and David B. Mustard, "Casinos, Crime and Community Costs," University of Georgia working paper (2001).

Levitt, Steven D., "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," *Quarterly Journal of Economics* 111:2 (1996), 319–351.

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

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

Lott, John R. Jr., "The Effect of Conviction on the Legitimate Income of Criminals," *Economics Letters* 34:4 (1990), 381–385.

——— "Do We Punish High Income Criminals Too Heavily?" *Economic Inquiry* 30:4 (1992), 583–608.

——— "The Concealed-Handgun Debate," *Journal of Legal Studies* 27:1 (1998), 221–243.

Lott, John R. Jr., and David B. Mustard, "The Right-to-Carry Concealed Guns and the Importance of Deterrence," *Journal of Legal Studies* 26:1 (1997), 1–64.

Mustard, David B., "Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the US Federal Courts," *Journal of Law and Economics* 44:1 (2001), 285–314.

Philipson, Thomas, and Richard Posner, "The Economic Epidemiology of Crime," *Journal of Law and Economics* 39:2 (1996), 405–433.

Shepherd, Joanna Mehlhop, "Fear of the First Strike: The Full Deterrent Effect of California's Two- and Three-Strikes Legislation," *Journal of Legal Studies,* 31:1 (part 1, January 2002a), 159–201.

Shepherd, Joanna Mehlhop, "Police, Prosecutors, Criminals and Determinate Sentencing: The Truth about Truth-in-Sentencing Laws," Emory University Working Paper, (2002b).

U.S. Department of Commerce, Bureau of the Census. *Census of Population and Housing, 1990 (United States),* Summary Tape File 3C (computer file), Washington: U.S. Department of Commerce, Bureau of the Census (producer) (1992); Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor) (1993).

U.S. Department of Justice, Federal Bureau of Investigation. *Uniform Crime Reports: County Level Detailed Arrest and Offense Data,* 1977–1992 (Washington: U.S. Department of Justice, Federal Bureau of Investigation; Ann Arbor, MI: Inter-university Consortium for Political and Social Research).

——— *Uniform Crime Reports for the United States,* Washington (1977–1993).