THE IMPACT OF VICTIMIZATION STUDIES ON ECONOMETRIC EXPLANATIONS OF CRIME*

Wm C. Birdsall, The University of Michigan

A. Leslie Robb, McMaster University

The intention of this paper is to briefly review some econometric models of crime, describe the potential problem that erroneous crime data implies for these models, and use the data from the National Crime Panel Surveys to clarify how serious the problem is.

Although single equation models to explain crime have been quite common in crime research by political scientists and sociologists,[1] they are relatively rare in the economics literature on crime.[2] Most economists who have written on the subject---especially since Becker's (1) pioneer theoretical paper---are convinced that criminals respond to at least some variables that are policy instruments from the viewpoint of society, the levels of which are set in response to crime itself. Put most simply, they hypothesize that observed crimes or offenses are jointly determined by an offenses function, describing the determinants of criminal behavior, and by a defenses function, describing the determinants of the amount of resources society devotes to protecting citizens from crime and to punishing offenders. Two equation models of this type are quite common.[3] In Orsagh (13) one such model is estimated using cross section analysis of the cities/counties of California; that paper describes quite clearly the identification problem that one equation models are faced with and shows the sensitivity of some of the parameters to an appropriate two equation estimation procedure.

Two contributions, Erhlich (6) and Carr-Hill and Stern (4), posit and estimate more complex models. If offenses are not determined directly by the amount of resources that society devotes to defense, but are responsive to the probability of punishment, then the model has a third equation, akin to a production function, describing how the probability of punishment depends negatively on the total number of offenses in the jurisdiction and depends positively on resources devoted to defense (or resources devoted to detection and prosecution at least).

In cross-section analyses of their simultaneous equation models, both Erhlich and Carr-Hill and Stern found that society's policy instruments, the probability of punishment and the severity of punishment, deter crime. This in itself tends to weigh the burden of proof on those who might wish to assume that criminal behavior does not depend on society's punitive response to crime, and it casts doubt on inferences based on single equation estimates. However, our focus is on what the offenses functions in these models have in common with more naive models; namely that low social status (Carr-Hill and Stern and many predecessors) or poverty (Erhlich and many predecessors) and nonwhiteness (Erhlich and most studies using U.S. data[4]) are positive determinants of crime. An important question, especially given the apparent robustness of these results, is whether or not they are sensitive to the fact that all estimates have been made using data on crimes-reported-to-the-police.

Most of the contributors are quite frank about the inadequacies of the crime data used in the models. Erhlich (6) in Appendix 1 of Section III, analyzes the implications about bias and consistency of estimates when the errors in reporting are random shocks; Carr-Hill and Stern (4) expose the problem implied if the errors are not random; it is this latter point that is important for our purpose. The basic question in this regard is whether reporting error is systematic in a way that will interfere with our ability to make inferences about parameters estimated by using reported-crime data. An example will help. As we have noted, a common finding is that crime rates are positively related to income poverty, however measured. The usual economic interpretation is that the percent poor is a proxy variable for wages available in legal activity and that the positive coefficient on this variable is an estimate of the responsiveness of offenders to changes in alternative legal opportunities. To anyone well versed in econometrics, it is well known that if reporting error is systematically related to any exogenous variable, in our example percent poor, then the estimate derived from regressions using reported crime is not an estimate of the coefficient on poverty in the offenses function, but an estimate of a complex term which includes the true parameter in that function and a parameter in an equation relating reported crime to true crime. Put most simply, the effect of poverty on actual crime cannot be separated from the effect of poverty on the reporting of crime.

It is important to understand how general this analysis is. The problem that reporting error implies about interpreting empirical coefficients is completely independent of whether the true model of crime is a simple one-equation model or a simultaneous-equations model. To repeat this conclusion: If reporting error is systematically related to any exogenous variable in the offenses equation then no inferences can be made about the effect of that exogenous variable on true crime. (Of course, in a multiple-equation model at least this serious a problem occurs.)

The Carr-Hill and Stern insight takes us this far but no further; with the advent of victimization surveys however, we can go further, though, as we shall see, not nearly as far as we would like. Basically, insofar as a victimization survey is successful in gathering true crime statistics, it would allow us to avoid the problem entirely. In fact by examining equations with both true and reported crime, we might even gain insight into whether offenders and/or authorities appear to know actual crime and the true probability of punishment better than reported data would

allow them to.

All this would be possible if we had victimization data by the various jurisdiction levels on which crime models have been estimated (census tracts, precincts, cities, states) and compare estimates based on victimization data with estimates based on the FBI Uniform Crime Reports' (UCR) reported-crimes.

There are two reasons why this cannot be done at present. The first is that sample surveys of the approximate size and sampling methodology of the Current Population Survey (and we think the National Crime Panel Surveys (NCPS) are of that class) are not designed to give reliable estimates for many jurisdictions as small as even the state or city level. The second reason is that insofar as the Bureau of the Census will release files at all, to our knowledge it only releases them with locational identifiers for states and subdivisions within states removed. This decision is based in part on the Census' view of how finely the data can be reliably divided, and in part on the Census' view of its confidentiality responsibilities.

The NCPS, in fact, gives us a little more on jurisdictions than is usual. In published reports they give detailed data on the five largest cities and on eight other cities, not randomly chosen to be sure, for a total of 13 jurisdictions. Thirteen non-randomly selected observations is hardly a data set from which to estimate anything reliably. It may however, be a sufficient set to examplify a problem and point up the serious need for similar data for other jurisdictions. In that spirit we proceed.

The two NCPS's for cities give us victimizations by type of crime and reporting (to the police) rates for definitions of crimes that can be made to conform quite closely to UCR definitions. The first question to be asked of these data in relationship to econometric models is whether the reporting rates are statistically independent of the other variables used in those models. One test of this independence is simple correlation coefficients which should not be significantly different from zero if two variables are statistically independent. In Table 1 we give the simple r's between reporting rates and several exogenous variables commonly used to explain crime in simultaneous-equation and in more naive models. Reporting of crime, particularly of assault and burglary is significantly related to some favorite exogenous variables. (If we set a 5% chance as the maximum we would accept as the chance of concluding a significant relationship where none exists, then we would expect to see 1 or 2 coefficients with single asterisks among the 30 coefficients in this table; in fact we find 8.)

The strong correlation between some favorite predictors of crime, such as poverty and race, and reporting rates for crimes implies that poverty- and race-effects on true crime are confounded with reporting-effects. The direction of the bias is upward (in absolute value) if the parameter estimate of the effect of an independent variable is of the same sign as the correlation coefficient; it is a downward bias if they are opposite in sign. Percent poor has been a good predictor of many types of crime with a positive coefficient in studies using URC data. Since, for assault and burglary at least, it is also related positively to the likelihood that crimes are reported, the size of those estimates, and possibly their statistical significance, is exaggerated. In more sophisticated models the direction of the bias may depend upon one or more other parameter values. For example, in Erhlich's model the direction of bias of the estimate of the responsiveness of crime to the percent poor and the percent nonwhite depends upon whether the coefficient on probability of punishment is greater than, equal to, or less than -1. If it is negative but less than 1 in absolute value then Erhlich's estimates of the positive responsiveness of crime to percent poor and percent nonwhite are overestimates of the true parameters.

The small size of our sample of cities is taken into consideration in determining the significance of the r's on Table 1, but the non-random nature of this set of cities is not. The latter leads to considerable modesty in drawing conclusions from the table. On the other hand, the frequency of significant correlation in a sample of such limited size requires one to take the results quite seriously.

A final word is due on why we leave these results in terms of simple correlations. Our answer is that we have no reasonable a priori theory or model to explain the reporting rates. Our own crude hypotheses were that percent nonwhite, percent poor, and percent female-headed households, would be negatively related to reporting rates for most, if not all, types of crime. We thought we may have arrived at these ideas from published data or discussions of the earlier NORC pilot victimization survey, (2) and (3), but have not found any basis for those views. Perhaps our ideas were based on the assumption that in neighborhoods so characterized crime is more commonplace, and there is more hostility to police; thus reporting rates might be expected to be low. Since we had that naive neighborhood model in mind, the opposite correlation surprised us. It may be quite important that the unit of observation on which our correlations are based is the city as a whole. More detailed analyses may give us some insight into why these correlations exist for cities, but the data in their published form tell us little more. The correlations certainly wet our appetites to understand more about the determinants of crime-reporting rates.

We do not wish to overstress results based on such a small non-random sample, but at a minimum these significant correlation coefficients should make us very cautious about drawing facile conclusions about the "causes" of crime from estimates of offenses functions based on UCR data.

Table 1.

Simple Correlation Coefficients Between "Percent of Victimizations[a] as Reported to the Police" and Commonly Used Explanatory Variables in Crime Models, by Type of Victimization[b]

	Robbery	Rape	Assault	Burglary	Larceny
Density	-.10	.57*	.50**	.63*	.28
Nonwhite	.65*	-.05	.70*	.40	.00
Median Education	-.41	-.07	-.80*	-.51**	-.42
% Female Headed Households	.36	-.03	.85*	.55**	.22
% Poor	.47	-.08	.68*	.52**	.17
Police Employment Per Capita	.27	.29	.75*	.50**	.35

a Victimizations categorized to approximate U.C.R. definitions. See (9) for the methodology. The victimizations are for the year 1972 (approximately).

b The observations are for the 13 cities for which both the victimization and socioeconomic data are available: Los Angeles, Denver, Atlanta, Chicago, Baltimore, Detroit, St. Louis, Newark, New York, Cleveland, Portland, Philadelphia, and Dallas.

* (**) indicates significance on a two-tail test at the 5% (10%) level.

Sources: The socioeconomic data is from the 1970 U.S. Census; the victimization data is from Crime in the Nations's Five Largest Cities, April 1974, and Crime in Eight American Cities, July 1974; Table 8 in both; both are Advance Reports, U.S. Department of Justice, LEAA, NCJISS, Washington D.C.

FOOTNOTES

* This research was supported by Grant 1R03MH 25565-01, awarded by NIMH, DHEW.

1. See, for example, (5) and (15); the latter surveys a great number of examples.

2. But see (7) and (14); Mehay (12) discusses a more complex model, but his estimates are of a single-equation model.

3. See (8) and (11).

4. Ref. (11) is an exception. The race variable is insignificant in his offenses function.

REFERENCES

1. Becker, Gary S., "Crime and Punishment: An Economic Approach," in Essays in the Economics of Crime and Punishment, Gary S. Becker and William Landes, (editors) NBER, 1974, pp. 1-54. (Reprinted from the Journal of Political Economy, 1967).

2. Biderman, Albert D., "Surveys of Population Samples for Estimating Crime Incidence," Thorsten Sellin (editor) The Annals of the American Academy of Political Science, Vol. 374, November 1967, pp. 16-33.

3. Biderman, Albert D., and Albert J. Reiss, Jr., "On Exploring the 'Dark Figure' of Crime," Thorsten Sellin (editor) The Annals of the American Academy of Political Science, Vol. 374, November 1967, pp. 1-15.

4. Carr-Hill, R.S., and N.H. Stern, "An Econometric Model of the Supply and Control of Recorded Offenses in England and Wales," Journal of Public Economics, 1973, pp. 289-318.

5. Cho, Yong Hyo, "A Multiple Regression Model for Measurement of the Public Policy Impact on Big City Crime," Proceedings of the 1972 meeting of the Operations Research Society of America, Santa Monica, California.

6. Ehrlich, Isaac, "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," Essays in the Economics of Crime and Punishment, Gary S. Becker and William M. Landes (editors), NBER, 1974, pp. 68-132. (Reprinted from Journal of Political Economy, 1973 with supplementary material.)

7. Fleisher, Belton M., The Economics of Delinquency, Quadrangle Books, 1966.

8. Grieson, Ronald E., "The Determinants of Juvenile Arrests," Working Paper #87, July 1972, Department of Economics, MIT.

9. Hindelang, Michael, Criminal Victimization in Eight American Cities, Unpublished manuscript. 1975.

10. Hoch, Irving, "Factors in Urban Crime," Journal of Public Economics, 1974, pp. 184-229.

11. McPheters, Lee R., and William B. Strange, "Law Enforcement Expenditures and Urban Crime," National Tax Journal, December 1974 (XXVII, No. 4), pp. 633-644.

12. Mehay, Stephen L., (1974), "The Production of Crime Deterrence by Urban Police Departments," Concordia U. Working Paper #1974-32.

13. Orsagh, Thomas, "Crime, Sanctions and Scientific Explanation," Journal of Criminal Law and Criminology, 1973 (Vol. 64. No. 3), pp. 354-361.

14. Phillips, Llad and Harold L. Votey, Jr., "Crime, Youth, and the Labor Market," Journal of Political Economy, Vol. 80, No. 3, Part I, May/June, 1972.

15. Wilks, Judith A., "Ecological Correlates of Crime and Delinquency," Appendix A of Task Force Report: Crime and Its Impact--An Assessment, President's Commission on Law Enforcement and Administration of Justice, United States Government Printing Office, Washington, D.C., 1967.