

THE VALIDITY OF OFFICIAL CRIME STATISTICS: AN EMPIRICAL INVESTIGATION¹

WESLEY G. SKOGAN

Northwestern University

STUDIES OF CRIME, DEVIANCE, AND SOCIAL DISORGANIZATION, AND THE IMPACT of social conditions and policy decisions upon their scope, have been hampered by widespread suspicion of the validity of official measures of crime. Crime statistics, the most notable and accessible of which are those contained in the Federal Bureau of Investigation's yearly *Uniform Crime Report*, are reputedly invalid indicators even of the limited actions which they purport to measure directly. Their validity is threatened in the sense that the reported figures are not one-to-one reflections of events; the number of burglaries "known to the police" in official parlance do not equal the number of burglaries which have taken place. These statistics actually reflect the interaction between three sets of activities: things which go on "out there" in the environment (crime), the responses of those who are victims of it (reporting), and society's effort to discover and record it (policing). The frequency of these activities vary across communities and over time, and the mix reflected in any given set of crime statistics is problematic. The independent effect of these factors upon our estimates of the distribution of crime is currently unknown, but both anecdotal and systematic observation suggest that the impact of each is considerable.

Although there is a large methodological literature critical of these statistics, "crimes known" figures continue to be used in quantitative studies of variation in crime. They are readily available, they have been collected for over 40 years, and they are aggregated at a convenient level, the city. There they can be related to census, police department, and political data. But because it appears inappropriate to take crime figures at face value, the authors of these studies typically hedge their conclusions, footnote the critical literature, and suggest that their numbers *really* measure something like "acts which come to the attention of the authorities," or "the socially recognized volume of crime."

This report suggests that the data may be more useful. While error in official estimates of the incidence of crime may present serious problems for those interested in the development of social indicators, it may not be devastating for other enterprises. Official crime statistics may be quite useful as such, if we make modest demands of the data. This report compares official city crime statistics in two categories (robbery and auto theft) with survey-generated measures of the incidence of victimization in the same communities. The comparison suggests that official figures may be useful indicators of the relative distribution of crime across cities. That is, while they do not specify the exact incidence of crime in a city, they tell us with

¹ I would like to thank Paul Friesema, Herbert Jacob, and Floyd Schwartz for their suggestions and contributions.

some accuracy which cities have more crime than others. The convergence of measures indicates that quantitative studies of inter-city variations in officially measured crime may not be seriously affected by measurement error. Statistics, like means and regression coefficients, which purport to reflect the actual levels of measured quantities, will be inaccurate, but correlations and scatterplots can be used to good effect in analysing the causes and consequences of variations across jurisdictions in the underlying, victim-defined distribution of crime.

MEASURING CRIME

Regardless of the technique employed to gauge the incidence of crime in a community, the measurement process proceeds in three stages: crime is (1) uncovered, (2) classified, and (3) recorded. Events that occur in the environment which surface as recognized crimes pass through these filters, each of which affects both volume and distribution of the final figures.

The first problematic step in the official measurement process is the uncovering of events. Modern, big-city policing is largely reactive—the police respond to citizen reports of victimization.² National survey studies indicate that official statistics greatly underrepresent the universe of reportable events. The incidence of more serious offenses is estimated to range from two to five times the officially recorded volume, figures which *may* be attributable largely to citizen non-reporting.³ The largely victimless crime which is uncovered by proactive police investigations add greatly to the total of arrests each year (drunkenness is the most common), but the number which are detected is largely a function of the resources which police administrators are willing to allocate to such activities, not the underlying incidence of “criminal events.”⁴

At the classification stage, police officers are faced with technical and political considerations which shape the resulting statistics. Large numbers of citizens' complaints are not based upon legally actionable events, and police rules for “unfounding” complaints, or deciding which act in a complex series of events is the “highest” on the offense list (which statistical category it will be counted in), are variously understood and followed.⁵ It is at the classification stage where organizational pressures to reduce the apparent incidence of serious crime will first be evident. There, robberies may be “reduced” to larceny, grand larceny to petty larceny and auto theft to unauthorized use of an automobile. As one Chicago police officer recently noted, “It is impossible under the present system to write factual and

² Albert J. Reiss, Jr., *The Police and the Public* (New Haven: Yale University Press, 1971), pp. 3–19.

³ President's Commission on Law Enforcement and Administration of Justice, *Task Force Report: Crime and Its Impact—An Assessment* (Washington, D.C.: U. S. Government Printing Office, 1967), pp. 17–19.

⁴ Raymond T. Nimmer, *Two Million Unnecessary Arrests: Removing a Social Service Concern from the Criminal Justice System* (Chicago: American Bar Foundation, 1971).

⁵ Federal Bureau of Investigation, *Uniform Crime Report 1970* (Washington, D.C.: U. S. Government Printing Office, 1971), pp. 57–60.

honest official reports and stay out of the commander's office very long."⁶

The recording stage is the point at which the decision to give official status to an event is made. Black's report of an observational study of police-citizen interactions suggests this action is as problematic as any in the measurement process. The legal seriousness of an event, the deference of the parties toward the police, the preferences of the complainant, social status, and the relational distance between complainant and suspect, all affect the willingness of investigating officers to record officially that a crime has taken place.⁷

It has long been recognized that the social processes which shape the uncovering, classifying and recording of events greatly affect the volume of officially known crime in modern societies. At one extreme, this has led to Beattie's admonition not to use them for research purposes at all.⁸ Critics in this tradition argue that official crime statistics are merely measurement artifacts, largely reflecting contingencies in reporting and policing. At the other extreme, Tittle concludes:

The unreliability of crime statistics is well known, but the lack of other sources of data precludes alternative approaches. We must continue to work with these records and alter the conclusions if necessary when more reliable information becomes available.⁹

In between, there have been a number of attempts to reconceptualize the *meaning* of official statistics. To social interactionists, crime statistics identify behaviors which are "organizationally defined" as deviant.¹⁰ At the extreme end of this position, Black suggests that "crime statistics are not evaluated as inaccurate or unreliable. They are an aspect of social organization and cannot, sociologically, be wrong."¹¹

While by logical extension this perspective may lead one to conclude that there is no objectively definable event called a "crime," quantitative research grounded even in this tradition would be improved if indicators were available which were relatively untainted by organizational processes. Then the activities involved in classifying and recording data could be treated as variables, not as part of the definition of crime.¹²

It was largely in response to these well-known problems that social sci-

⁶ *Chicago Tribune*, March 30, 1973.

⁷ Donald J. Black, "Production of Crime Rates," *American Sociological Review*, 35 (Aug., 1970), pp. 733-758.

⁸ Ronald H. Beattie, "Criminal Statistics in the United States," *Journal of Criminal Law, Criminology and Police Science*, 51 (May-June, 1960), pp. 49-65.

⁹ Charles R. Tittle, "Crime Rates and Legal Sanctions," *Social Problems*, 16 (Spring, 1969), pp. 411-412.

¹⁰ John J. Kitsuse and Aaron V. Ciccourel, "A Note on the Uses of Official Statistics," *Social Problems*, 11 (Fall, 1963), pp. 131-139.

¹¹ Black, "Production of Crime Rates," p. 734.

¹² Richard Quinney, "Structural Characteristics, Population Areas, and Crime Rates in the United States," *Journal of Criminal Law, Criminology and Police Science*, 57 (March, 1966), pp. 45-52.

entists began to explore alternative techniques for the measurement of crime. National sample-survey measurements were funded by the President's Commission on Law Enforcement in the mid-1960's, and they have become a widely used technique for gauging that "dark figure" of crime which does not come to our official attention.¹³

While survey-generated estimates of the incidence of crime are not filtered by the same set of social and organizational processes which shape official statistics, they also are plagued by measurement problems. Surveys uncover crime by asking people about their experiences. Because they are intended to uncover the dark figure, respondents are selected on some basis other than victimization, typically by household randomization. The interviewer thus initiates the notion of victimization, and probes with varying tenacity to elicit "self reports." The limited studies we have of the reliability and the validity of this process suggest that they are not very high. Among one sample of 400 persons who were *known* to have called the police to report victimization, 20 percent failed to volunteer that event in a follow-up interview.¹⁴ Experiments indicate that even minor variations in the wording and timing of the questions have substantial effects upon response patterns.¹⁵

Classification also presents difficulties. Less focused surveys shift responsibility for the definition of events to the respondent, and without the equivalent of police "unfounding" procedures the resulting reports are far from uniform. In one test, a lawyer screened interview schedules and judged that legally actionable offenses had not taken place in about 20 percent of the reported victimizations.¹⁶ All manner of events are reported to interviewers, and survey techniques do not easily convert these into data comparable to official statistics.

In addition, interview measurements are sensitive to all of the standard problems encountered in survey research. In Reiss' report to the Law Enforcement Commission, he noted that the response rate in one high-crime precinct was only 62 percent. High-income residents in the center-city are increasingly surrounded by a phalanx of security guards and building managers who define interviewers as undesirables, to be excluded.¹⁷ Like official statistics, survey measures of crime are affected by social and organizational processes which blur our image of victimization.

¹³ Albert D. Biderman and Albert J. Reiss, Jr., "On Exploring the Dark Figure of Crime," *Annals of the American Academy of Political and Social Science*, 374 (Nov., 1967), pp. 1-15; and Phillip Ennis, *Criminal Victimization in the United States: A Report of a National Survey* (Washington, D.C.: U. S. Government Printing Office, 1967).

¹⁴ Albert J. Reiss, Jr., "Measurement of the Nature and Amount of Crime," in President's Commission on Law Enforcement and Administration of Justice. *Studies in Crime and Law Enforcement in Major Metropolitan Areas: Field Surveys III, Vol. 1* (Washington: U. S. Government Printing Office, 1967), p. 150.

¹⁵ *Ibid.*, pp. 148-149.

¹⁶ *Ibid.*, pp. 151-152.

¹⁷ *Ibid.*, pp. 145-146.

VALIDITY AND THE CONVERGENCE OF MEASURES

The extent to which the *Uniform Crime Report's* "crimes known" figures systematically reflect the underlying distribution of victim-defined crime in local communities can be probed using multi-city survey data. Both survey and *U.C.R.* statistics are error-prone estimators of the true incidence of crime in a jurisdiction, and they give us two quite different views of that set of events. In measurement terms, they are maximally different indicators of the same variable. While potentially errorful, they share no "method variance." The overlap between these two views of events can thus be interpreted as a validity coefficient, a measure of the utility of indicators in social research.¹⁸ A validity coefficient is the best estimate of the extent to which a measure successfully reflects an underlying distribution of events. In this case, if the ranking of cities by crime on each measure is similar, the inference is that each usefully reflects the target phenomenon. If this is true, the easily available official statistics may be employed as indicators of the relative distribution of something more fundamental than the measurement artifacts which have been so widely described.

Comparison is made here between survey and official estimates of crime in ten major American cities. In other contexts this would be considered a small sample, and the restricted sample size greatly limits the sophistication of the analysis. But the requirement that the sampling frame produce enough respondents within each jurisdiction to enable us to characterize cities as units for analysis sets severe bounds upon the magnitude of any such study. The cities, which ranged in size, from Kansas City, Kansas (population 168,000) to Baltimore, Maryland (906,000), were surveyed in 1970 as part of the Urban Observatory Program of the Department of Housing and Urban Development.¹⁹ Within each city, sample households were drawn from an equally-weighted inventory of all households within the community's legal boundaries. Within each household, one adult respondent was selected using an objective procedure. No substitutions were permitted, and at least six call-backs were made before replacing hard-to-reach respondents. A total of 4,266 respondents were interviewed. A list of the cities and the number of respondents in each is presented in Table 1.

Each respondent was asked: "Here is a list of some crimes that happen to people. In the past year has anything like this happened to you or anyone living with you? Anything not on the list?"

The question and follow-up probes elicited 1,788 coded reports of victimization. These ranged from "peeping tom" incidents (12) to rape (2) and serious assault (277). If anyone in the respondent's family was victimized more than once in the same way, each incident was counted sepa-

¹⁸ Donald T. Campbell and Donald W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin*, 56 (March, 1959), pp. 81-105.

¹⁹ The Program is administered by the National League of Cities and the United States Conference of Mayors Secretariat. While they kindly made the data available to me for reanalysis, they bear no responsibility for its use and interpretation.

TABLE 1
Sample Cities and Crime Rates

City	Sample Size	Survey Rates per 10,000 Population		Official Rates per 10,000 Population	
		Auto	Robbery	Auto	Robbery
Boston, Mass.	507	331	169	238	50
Kansas City, Kan.	193	122	46	112	30
Kansas City, Mo.	383	113	40	123	56
Milwaukee, Wis.	443	100	50	70	9
Nashville, Tenn.	426	51	22	75	21
Albuquerque, N.M.	471	114	15	74	17
Atlanta, Ga.	469	146	58	88	32
Baltimore, Md.	500	138	69	114	109
Denver, Colo.	357	178	37	141	36
San Diego, Calif.	517	60	20	52	11

rately. Respondents reported 120 purses snatched and pockets picked, 106 mailboxes robbed, and six obscene telephone calls. Victimization by this measure is largely (although distorted by the initial list) respondent-defined, and occurrences reported in the survey vary greatly in seriousness and in the extent to which they reflect events which would be legally actionable. Part I crimes reported in the *Uniform Crime Report*, on the other hand, are a selected sub-set even of the data available to the F.B.I. They were chosen for public analysis because they are serious and among the more completely reported crimes.²⁰ Thus we cannot construct single, global measures of the incidence of crime in cities and compare them across measurement techniques. Summary crime rates based upon simple aggregations would be definitionally incomparable.

There were two (and only two) categories of responses to the survey which were relatively well-defined and roughly match categories of offenses known to the police summarized in the *Uniform Crime Report*: robbery and automobile theft. The analysis which follows deals only with events in these two offense categories. While this may limit the generality of our findings, there is some evidence that robbery and automobile theft rates may be representative of a larger domain. Robbery is an act of desperation which is often classified as both a property and a violent crime. It is regarded by some as an important "bellweather" indicator of patterns of both.²¹ In order to establish benchmark distributions for the variables reported in this study, 1970 census and crime statistics were analyzed for all

²⁰ David J. Pittman and William F. Handy, "Uniform Crime Reporting: Suggested Improvements," in Alvin W. Gouldner and S. M. Miller, eds., *Applied Sociology* (New York: The Free Press, 1965), pp. 180-181.

²¹ John E. Conklin, *Robbery and the Criminal Justice System* (Philadelphia: J. B. Lippincott, 1972), p. 3.

United States cities. Among cities in the size range we are considering here (between 150,000 and one million in 1970), robbery rates were highly correlated with rates for a variety of other Part I offenses: .62 with murder, .60 with rape, .58 with burglary. National survey studies indicate that auto theft is a very accurately reported crime, and its victims come from more varied class and racial backgrounds. Auto theft rates were not highly correlated with violent Part I offense rates, but they were representative of crimes of profit.

Reports of victimization in these two categories were used to generate survey measures of crime rates at the city level. For each city, the numerator of each measure is the sum of respondent-reported victimizations. As the survey item asked respondents to report victimizations of "you or anyone living with you," the denominator of each measure is the total number of persons living in the households sampled.²² The survey figures are corrected for jurisdiction: about two percent of the victimizations reported in the interviews took place outside of the cities under investigation. Boston, Baltimore, and Atlanta ranked as high crime cities on the two survey measures: Nashville and San Diego were low on both. The survey estimates of robbery and auto theft rates (victimizations per 10,000) for each of the ten cities are reported in Table 1.

Table 1 also presents comparable official statistics for the ten cities. The time frame of the survey item, "the past year," bridged two reporting periods, so the official figures reported here are averages for the years 1969 and 1970, as rates per 10,000.

The first question is, "How highly related are survey and official estimates of city crime rates?" While each is an error-prone indicator of the true incidence of crime, if they highly co-vary it suggests that the measurements reflect more than method variance, and that official statistics may be used as indicators of something more than citizen reporting and police record-keeping behavior. For our ten-city sample, the Pearson product-moment correlation between survey and official rates for auto theft is .94, and the correlation between survey and official rates for robbery is .39.

The relationship between survey and official estimates of automobile theft is so strong that they serve as virtually interchangeable indicators of inter-city variations in this type of crime. Again, they are errorful, and neither tells us what the *real* auto theft rate is. In measurement terms, this is the "true score" of the measured variable. Social scientists and policy makers who are interested in social indicators demand such "accurate" statistics. As Table 1 illustrates, high correlations between the results of these differing measurement techniques do not mean that we have successfully triangulated upon such a figure for auto theft: the two estimates of the auto theft rate differ considerably in *magnitude*. What is important for

²² Tests suggest that this household informant technique is less costly than individual-respondent items, and that it is accurate in white households. It does, however, tend to under-represent victimization in black households. See Reiss, "Measurement of the Nature and Amount of Crime," p. 147; and Ennis, *Criminal Victimization*, pp. 51-57.

correlational studies of inter-city variations in crime is that their distribution is quite similar. The high parametric correlation is affected by the extreme observed values for Boston, a city experiencing a very high automobile theft rate, whatever the measurement procedure. A conservative test of the "relative ordering" of cities which discounts most of the distributional properties of the data is the Spearman rank order correlation. For auto theft, the Spearman correlation between official and survey estimates of the crime rate is .76. Kansas City, Missouri, which falls four ranks "too high" on the official measure, contributes most of the error. In the case of auto theft we can be fairly confident that indicators generated from *Uniform Crime Report* data reflect popularly-defined and citizen-experienced crime, and that certain cities have more of it than others.

The moderate correlation between official and survey estimates of the robbery rate, on the other hand, does not mean that we must discount the utility of either indicator. The correlation, in fact, improves to .54 when the skewed values of two high scorers, Baltimore and Boston, are brought into line with a rank order Spearman measure of association. That the two are not highly related suggests that there is substantial amount of error introduced in the measurement process. Every social science measurement has an error component. Recent analysis by the United States Bureau of the Census suggests that the error in the 1970 estimate of the nation's population was 2.5 percent.²³ But typically we do not know what the error components of our measures are. Rather, we make (perhaps unwarranted) assumptions about them: that they are "random," and their error variance is uncorrelated with other measures in our study. We do this because, while knowledge about the relative size of the error component in a measure is interesting, knowledge about its distribution is crucial. If error in the dependent variable is randomly distributed across the independent variables in an investigation, then it will have at most a conservative effect which will not lead us dangerously astray in correlational studies, whatever its magnitude.

RATES AS VARIABLES

Analysis of the correlates of crime in this ten-city sample suggests that the assumption of random measurement error may often be warranted. Far from being artifactual, official statistics do not appear to lead us to make radically incorrect judgments. This section of the report examines the correlates of crime when it is measured in different ways, when the sources of error differ. Patterns in the data generally are invariant across measurement techniques.

A common enterprise among social scientists is the analysis of the correlates of crime (officially measured) at precinct, city, state, and national levels. These studies probe theories about such things as social disorganization and economic deprivation. They support a host of bivariate empirical generalizations about crime: the more deprived and disorganized a popu-

²³ *New York Times*, April 26, 1973.

lation, the more crime there is. Authors of these reports are often apologetic about the measures of their dependent variable, suggesting that they are at best inaccurate representations of true crime rates. Social scientists go ahead with these studies because what we find is congruent with empirical generalizations founded upon data gathered at the individual level: arrestees, prisoners, and self-reported perpetrators of various offenses are typically deprived and come from backgrounds that we would predict on the basis of social disorganization theory.

The difficulty inherent in our lack of confidence in our measures arises when we step beyond individual-level inferences and attempt to use official statistics to probe relationships which are not firmly documented by other kinds of research. We are simply not confident enough to decide if we are right or wrong. So, for example, Webb recently reported a test of Durkheim's theory about the effect of the division of labor (a characteristic of collectivities) upon deviance. His hypothesis:

There is a direct (and positive) relation between the extent of the division of labor in communities and the extent of deviance within these communities.²⁴

He employed crimes-known statistics aggregated at the city level as his measure of deviance, and he discovered very low correlations. There are a number of alternative interpretations of this finding, among them that (a) Durkheim was wrong, and (b) the indicators are unworthy. We usually have more faith in our theory than in our crime data, so Great Hypotheses are in practice seldom firmly rejected.

This study suggests that official statistics may be more worthy of confidence. The first question explored revealed that auto theft indicators were highly related across cities, and that they would serve as virtually interchangeable measures in correlational studies of their distribution. The lower congruence of robbery measures leads us to a second question probing the utility of official crime statistics, "Would existing error in measurement lead us to make incorrect inferences from the data?" Table 2²⁵ presents a number of variations in crime, and their relationship to both survey and official measures of crime rates.

Several indicators of independent variables are presented here: measures of racial heterogeneity, poverty, age distribution, ethnicity, employment, and wealth. These are common indicators of the causes of crime for which we have complementary individual-level support for city-level findings. Table 2 also includes several measures of the ecological or contextual characteristics of cities which are thought to be criminogenic: population, den-

²⁴ Stephen D. Webb, "Crime and the Division of Labor: Testing a Durkheimian Model," *American Journal of Sociology*, 78 (November, 1972), p. 645.

²⁵ The demographic and occupational data in Table 2 come from the United States Bureau of the Census, *1970 Census of Population: General Social and Economic Characteristics* (Washington, D.C.: U. S. Government Printing Office, 1972); and United States Bureau of the Census, *1967 Census of Manufactures: Area Statistics* (Washington, D.C.: U. S. Government Printing Office, 1971).

TABLE 2
Correlations with Survey and Official Estimates of Crime Rates

Indicators	Survey Auto	Official Auto	Survey Robbery	Official Robbery
Population	.13	.10	.33	.48
Percent Nonwhite	.07	.03	.23	.61
Percent Under 18	-.54	-.55	-.58	-.07
Percent Foreign Born and Stock	.64	.57	.68	-.04 ^a
Percent Unemployed	-.16	-.30	-.18	-.17
Percent Below Poverty	.27	.13	.30	.47
Percent Above \$15,000	-.25	-.34	-.40	-.43
Density	.77	.68	.84	.55
Manufacturing	.56	.64	.64	.63

^a This is the only relationship which varies in a statistically significant way across measurement technique, and the only reversal of sign.

sity, economic function. Of interest in Table 2 is the difference between the inferences we would make, or the hypotheses we would reject, if we employed survey or official indicators in our research. In this section, the goal is not to infer population estimates of correlations from this sample of ten. Rather, it is to compare descriptions generated from different data sources. Parametric product-moment correlations are used here to reflect what we know about our data, its distribution. It is precisely the great disparities in income, employment, race and ethnicity that we are interested in when we study the effects of these factors. Examination of the appropriate pairs of correlations reveals that the descriptions are similar. Of the 18 pairs of correlations presented in Table 2, 17 are of the same sign. Only inferences about the form of the relationship between ethnicity (percent foreign born and foreign stock) and crime would vary according to the measurement procedure employed. There the error would be conservative in studies using official statistics, for we would accept the null hypothesis on the basis of the low observed correlation. Very few social theories do more than predict the sign of hypothesized relationships (witness Durkheim), and official crimes-known statistics could usefully be employed to test them.

Differences in the convergence of robbery and auto theft measures are reflected in the magnitude of the observed correlations. Survey and official measures of auto theft rates for the ten-city sample were highly related, and the correlations between each of them and the selected variables are quite similar. On the average, correlations with official and survey measures differ only .08. The measures of robbery rates are less uniformly related to these variables. They differ an average of .25, but note that two indicators, percent under 18 years of age and ethnicity, account for the magnitude of this average. Although the two measures of robbery rates are only moderately correlated, the variance they share appears to be what is

systematically related to other variables. The similar pattern of the correlations with these variables suggests that neither measurement technique generates method-specific variance that is systematically related to them, increasing our confidence in the utility of crimes-known statistics. A formal test of the significance of the difference between correlations, when applied to these pairs, indicated that only the ethnicity-robbery comparison was sufficiently deviant to cause us to reject the conclusion that they are similar.²⁶ If we exercise some caution, the assumption may be warranted that the errors in our measures of crime can be treated as if they are "random."

The same conclusion results from an analysis of indicators of processes directly related to the measurement problem, the characteristics of police departments. While police organizational processes undoubtedly introduce error into official estimates of crime, they would have to do so in different ways and at substantially different rates in various cities for it to affect the relative distribution of the resulting figures. The correlation coefficient tells us how consistently cities lie in the same position relative to the mean on each of two variables. Of interest in Table 3²⁷ is whether police departments of various types fall into different positions vis-à-vis crime when the latter is measured in radically different ways. If they do not, we have more confidence that organizationally-induced errors do not alter the relative distribution of crime indicators across the cities.

Table 3 presents a number of commonly used measures tapping three major dimensions of police organization: the resources available to the sample departments, their allocation of these resources in "professionalizing" ways, and structural features of their relationship with the community. In addition, Table 3 analyzes the distribution of public opinion about crime and the police in the sample cities, based upon data aggregated from the 4,266 survey interviews. The correlations presented in Table 3 indicate that police-induced error may not change the relative distribution of cities when we use official statistics to measure crime rates. Of the 22 pairs of correlations presented in Table 3, 21 are of the same sign, and 20 are not statistically different. While the latter is not a difficult test to pass given the small sample size, the overall pattern again supports our general conclusion: we would not be led astray if we utilized official, crimes-known figures as indicators of the inter-city distribution of citizen-defined victimization.

This does not mean that we do not have serious measurement problems. More sophisticated models of social processes require measurements which

²⁶ Hubert M. Blalock, Jr., *Social Statistics* (1st ed.; New York: McGraw-Hill, 1960, p. 310.

²⁷ The police expenditure data in Table 3 come from the *Municipal Yearbook*, averaged for the years 1968–1970. Civilian employment data are from the same source, averaged for the years 1969–1970. Community relations unit data were drawn from the 1967 *Municipal Yearbook* and data processing information came from the same source, 1971 edition. Missing data in these sources were augmented by annual reports from the ten departments and personal communications. See International City Management Association, *Municipal Yearbook* (Washington, D.C.: ICMA, yearly). Public Opinion data in Table 3 were aggregated from the sample survey.

TABLE 3
**Characteristics of Police Departments as Correlates of Survey and
 Official Estimates of Crime Rates**

Indicators	Survey Auto	Official Auto	Survey Robbery	Official Robbery
Police Expenditures per capita	.68	.64	.78	.64
Data Processing Applications in Law Enforcement	-.20	-.28	-.41	-.49
Professionalism				
Percent Personnel Civilian	-.30	-.18	-.47	-.20
Expenditures per Officer	.23	.10	.32	.26
Capital Expenditures per Officer	-.47	-.54	-.42	.19 ^a
Community Relations				
Residency Requirement	-.26	-.06	-.35	-.23
Review Power	.67	.80	.63	.46
Citizen Complaint Board				
Age of Community Relations Unit	-.04	-.23	-.08	-.71 ^b
Public Opinion in Cities				
Percent Feel Safe at Night	-.58	-.70	-.67	-.79
Percent Rate Police Service Bad	.65	.58	.71	.74
Percent Favoring Increased Police Expenditures	.46	.59	.55	.72

^a Statistically significant difference and reversal of sign.

^b Statistically significant difference, but no reversal of sign.

accurately reflect the magnitude as well as the distribution of variables, and it is there where the use of official statistics may lead to difficulties. Any analysis which attempts to estimate numerical population parameters, the average level of some kind of crime or the impact of some policy upon the frequency of crime, will be affected by inexact indicators of the magnitude of these events. While correlations are not tied to the values of the indicators they relate, parameter estimates such as regression coefficients are. This is unfortunate, for it is regression coefficients which give us the laws of science and which are of most interest to policy-makers.²⁸

²⁸ Hubert M. Blalock, Jr., "Theory Building and Causal Inferences," in Hubert M. Blalock, Jr., and Ann B. Blalock, eds., *Methodology in Social Research* (New York: McGraw Hill, 1968), pp. 187-191.

TABLE 4
Impact Upon Crime Rates: Regression Coefficients

Variables	Survey Auto	Official Auto	Survey Robbery	Official Robbery
Percent Unemployed	-13.49	-17.33	-8.80	-5.49
Percent Below Poverty	8.83	2.74	5.52	5.71
Manufacturing Employ- ment per 1,000	151	117	97	64
Sworn Police per capita	530	340	346	172
Police Expenditures Per capita	.045	.030	.030	.017

Statistical statements of empirical laws often take the form:

$$Y = a + bX + e$$

Parameter estimates, the a's and b's in these statements, reflect the magnitude of the impact of independent variables upon dependent variables; they suggest to policy makers answers to questions like, "What will I get for what I spend?" Table 4 presents estimates of the impact of some policy-related variables upon both sets of crime indicators.

What is of interest in Table 4 is not the size or sign of specific regression coefficients, but the relationship between the magnitude of the estimates and the measurement technique upon which they are based. While some estimates are similar, in general it makes a substantial difference how crime is measured. Regression estimates utilizing survey measures of the dependent variable often range from two to four times those based upon official statistics in the case of auto theft, with the exception of unemployment, where the relationship is reversed. Similar differences appear in the estimates for robbery, with the impact of a unit increase in police manpower appearing to be twice as great when we use survey data to measure the dependent variable. Given the enormous cost involved in substantially changing the level of any of these variables, policy makers are going to demand better forecasting of the consequences of their actions. We can only reply that in order to do so we need better crime data: estimates are not independent of measurement technique when we attempt to make inferences about the magnitude of crime in cities.

CONCLUSIONS AND IMPLICATIONS

This report has drawn upon a sample survey of citizen opinions and experiences in ten major American cities for estimates of the incidence of criminal victimization. It is commonly suggested that official measures of crime are of limited utility in measuring such victimization, and largely reflect measurement artifacts produced on both the citizen and police sides of the official crime-measurement process. A direct comparison of survey estimates of crime and official reports of crimes known to the police sug-

gests a more optimistic picture of the status of crime statistics as social science data.

(1) Official statistics are at least moderately correlated with survey reports of victimization. The two procedures generate different kinds of error variance in their measurements, but the correlation between them suggests that they also partially reflect the true score, and are not totally measurement artifacts.

(2) The true-score component of each of the measures (their common variance: our best estimate of the "actual" crime rate) appears to be distributed in similar fashion across cities, and errors in each appear to be unrelated to the independent variables we commonly use to explain that distribution. Thus, measurement errors in official statistics do not seem to lead us to false conclusions, or to inferences which are measurement-specific.

(3) On the other hand, the magnitude of the various estimates of the incidence of crime varies considerably. Any analysis which describes or attempts to predict the level of crime in a community, or changes in its exact magnitude, will be errorful. The conclusions will be measurement specific. This magnitude problem does not bias correlation coefficients (although it does set upper bounds upon their size), but does bias means and regression coefficients. We need better official statistics.

These conclusions are tempered by two limitations: the range of crimes and cities analyzed. While it may be difficult to generalize from any subset to "crime" (a slippery concept) in general, robbery and auto theft rates may be instructive. They are different in character (one a crime of desperation and one a juvenile whim), they have quite different victims, they occur in different proportions across communities (their rates are not highly correlated), and they are reported at different rates. Given these contrasts, the extent to which they and their errors evidence similar patterns of relationships is remarkable. In no case do robbery and auto theft figures appear to be sensitive to independent variables in any systematic way. This suggests that the findings reported here may not be crime-specific, but may reflect more general measurement processes.

The small number of cases is an economic problem, and given the cost of generating city-level survey indicators it is not likely to be dramatically improved in any study. But the current research literature is so tentative that virtually any opportunity to enhance our understanding of the quality of the data should be pursued. It is only by more careful attention to problems of measurement that we will be able to test and reject theories with any confidence using official crime data. This report suggests that, within limits, such confidence may not be misplaced.