Pergamon

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Criminal Justice 32 (2004) 547–555

JOURNAL OF CRIMINAL JUSTICE

Establishing the statistical relationship between population size and UCR crime rate: Its impact and implications

James J. Nolan III*

West Virginia University, Division of Sociology & Anthropology, 316 Knapp Hall, Morgantown, WV 26506, United States

Abstract

The fact that the *volume* of crime is related to the size of a jurisdiction's population has been well established. The relationship between *crime rate* and population size, however, is less clear. Crime rate presents crime on a per capita basis, and is intended to adjust for population size so that comparisons can be made. In this article, the author first establishes the statistical relationship between crime rate and population size. Once established, he conducted an analysis of crime rates in jurisdictions of various sizes and in a variety of population-based strata using data obtained from the Uniform Crime Reports (UCR). Based on his findings, the author discusses implications for research and analysis, database management, and making jurisdictional comparisons of crime rates.

Introduction

The crime statistics collected through the Uniform Crime Reports (UCR) are law enforcement agency-level counts of specific crimes known as "index crimes."[1] Although it was well-established over the years that official crime data had their limitations (Black, 1970; Kitsuse & Cicourel, 1963; Merton, 1957; Skogan, 1975), UCR was considered by many to be a popular source of crime statistics in the United States (Regoli & Hewitt, 2000; Schmalleger, 1999; Siegel & Senna, 2000) and, as some argued, a valid indicator of the index crimes (Gove, Hughes, & Geerken, 1985).

The UCR data are most recognizable in two main formats: crime *volume* and crime *rate*. Crime volume is simply a count of the number of crimes that occurred in a specific jurisdiction in a given year, while crime rate is a relativized number that presents crime on a per capita basis. Crime rate is generally expressed as the number of crimes per 100,000 residents in the population.

Each year when the UCR Program releases its publication, *Crime in the United States*, the news media, among many others, hurry to use the reported crime figures to compare jurisdictions and perhaps even compile rankings. Many argued against comparing or ranking jurisdictions in this way (Brantingham & Brantingham, 1994; Federal Bureau of Investigation, 2001; McCleary, Nienstedt, & Erven, 1982). In fact, in the very front of *Crime in the United States*, the FBI included a section entitled Crime Factors in which the agency warned against comparing crime rates and ranking cities (Federal Bureau of Investigation, 2001, p. iv). The Federal Bureau of Investigation (FBI) argued that many factors, such as economic conditions; degree of urbanization; climate; effective strength of law enforcement agencies, among many other things, affected the crime rate. So, unless other factors were considered, jurisdictional crime rates should not be ranked or compared.

* Tel.: +1 304 293 5801x3210.
E-mail address: jim.nolan@mail.wvu.edu.

0047-2352/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jcrimjus.2004.08.002

At first glance, this argument seemed logical; after all criminologists uncovered many explanations for geographic differences in crime rates. On closer examination, however, one could easily find fault with this logic.

It was true that there were many factors that affected the level of reported crime in a jurisdiction. Viewed in this way, crime was a *dependent* variable, i.e., it was dependent on all the other factors that might give rise to crime, such as the factors listed by the FBI. The crime rate, however, was a presentation of crime based on a unit size. The degree to which many independent variables combined to impact the reported crime rate of a jurisdiction was *reflected* in the UCR crime rate.

Consider this argument from a slightly different angle. Instead of crime, suppose the issue to be compared was "quality of life." One might look at crime—now an independent variable—as one of the many factors that affected the quality of life of a city or town. If crime were the only factor considered when comparing the quality of life in different cities, then the argument presented by the FBI became more reasonable. The same thing was true when considering other social indicators such as unemployment. There were many factors that affected the unemployment rate of a locale; however, to argue that the unemployment rates in major cities could never be compared unless the factors associated with unemployment were also considered made little sense. It was important to note that it was *not* the author's position that comparing jurisdictional crime rates was a good idea. Instead, he intended to show that the incomparability of jurisdictional crime rates lay in another place, i.e., the statistical relationship between crime rate and population size.

In defining the crime rate, it was envisioned that the impact of population on the volume of crime would be mostly eliminated through a unit size of the population. If so, the crime rate could provide a crude but standardized basis of jurisdictional comparisons. One could argue that the growth of population generally brings forth such factors as urbanization, higher population density, changes in the area's economic conditions and life styles, etc., therefore the crime rate might still be related to the population. If this is the case, a non-zero correlation would be observed between the crime rate and population size.

It was well established that the *volume* of crime within a jurisdiction was highly correlated with the population size. For decades, research on the impact of population on crime focused primarily on population characteristics such as changes in size and density or its demographic and socioeconomic makeup (Blau & Blau, 1982; Gibbs & Erickson, 1976; Green, Strolovitch, & Wong, 1998; Schuessler, 1962; Tarling, 1986; Watts, 1931 as examples). What prompted the present inquiry was that it was not clear in the literature whether a jurisdiction's *crime rate* was also related in such a consistently positive way with the jurisdiction's population size. Some studies found small, positive correlations between population size and crime rate (O'Brien, 1983; Ogburn, 1935; Skogan, 1974), while Reiman (2001) argued that no such correlation between these two variables existed. Reiman wrote, "A comparison of crime rates for Standard Metropolitan Areas. . .reveals a striking *lack* of correlation between crime rate and populations size" (p. 24).

In theory, if the volume of crime were proportional to the jurisdiction's population size, then the crime rate, being a constant across jurisdictions, would have no correlation with the population. If the occurrence of crime, however, were to increase say, in a nonlinear fashion with the jurisdiction's population, then the jurisdiction's crime rate would change with the population. Therefore, the relationship between the *crime rate* and the *population* depended upon the way the *crime volume* related to the *population*.

There were several reasons why understanding the nature and extent of the relationship between crime rate and population size might be important. As mentioned previously, the appropriateness of making jurisdictional comparisons of crime rates might hinge on this relationship. Also, it was a common practice for criminologists and other criminal justice professionals to use secondary data like UCR for research and analysis. Understanding the strength and direction of the relationship between these variables might be important when preparing crime data for analytical studies, particularly when it was necessary to impute for missing or erroneous data. Finally, government agencies like the FBI currently used statistical methods for processing crime data (i.e., imputation and outlier detection) that assumed crime rate and population size were *not* correlated. For example, the FBI used a ratio estimation procedure for imputing crime data. According to this method, the average crime rate of a population stratum was applied to the jurisdictions with missing data. The limitations of this (and similar) methods in light of a positive or negative relationship between crime rate and population size will be considered.

The primary purpose of this study, then, was to first establish the statistical relationship between crime rate and population size. Although one might assume that a relationship between crime rate and population size existed, the nature and extent of this relationship had yet to be defined. Once this relationship was established, an empirical analysis of UCR

crime data in various size populations and within a variety of population-based strata would be presented. Based on this examination, the author will discuss the implications of his findings.

The statistical relationship: establishing a measure of covariance

The key to understanding the statistical relationship between crime rate and population size was found in the two alternative methods for calculating the crime rate. The first method, denoted here as $\hat{R}$, was calculated by, first, adding the total number of crimes reported in an area consisting of n jurisdictions, then dividing by the combined population of the area. For example, in the state of California in the year 2000, there were 237 cities with populations over 25,000 (Federal Bureau of Investigation, 2001). The total number of UCR index crimes reported in all of these cities that year was 980,426 and the combined populations of all these cities was 24,984,233. According to this first method, $\hat{R}$, the area crime rate was 3,924.[2] The second method, denoted here as $\bar{R}$, was calculated by summing each agency's crime rate, then dividing by the total number of jurisdictions. Continuing with the above example, the sum of all crime rates for these 237 California agencies in the year 2000 was 817,723; therefore, the crime rate calculated according to this second method, $\bar{R}$, was 3,450. The difference between $\hat{R}$ and $\bar{R}$ multiplied by $\bar{P}$ (the averaged agency population) was the covariance between crime rate and population size.[3] Again, it was important to note that $\hat{R}$, the crime rate for the combined total of n agencies, was an average crime rate (averaged through the population-weights, P_i/P), while $\bar{R}$ was an arithmetic mean of the agencies' crime rates. The latter was obtained by assigning the uniform weight ($1/n$) to each agency.

The correlation between population size and crime rate, then, lay in (or was expressed through) these two types of crime rates. The difference between these two crime rates determined the strength and direction of this relationship. $\hat{R}$ would be greater than $\bar{R}$ in situations where the more highly populated agencies tended to have the higher crime rates. On the other hand, $\hat{R}$ would be less than $\bar{R}$ in situations where the less populated agencies tended to have higher crime rates. When the crime volume was proportional to the jurisdictional populations, $\hat{R}$ and $\bar{R}$ would be approximately equal.

Now, consider an ordinary linear regression of crime rate on population size, where the slope of the line was positive (i.e., $\hat{R}$ minus $\bar{R}$ was a positive number), an increment in the population would be accompanied by an increase in an agency's crime rate. Under this condition, and where the jurisdictions to be compared had different size populations, crime rates were not comparable unless they were first adjusted by the slope.[4] The difficulty created by this situation was that the slope of the line was not a universally applicable constant; it depended on the group of jurisdictions under study (i.e., it was "group dependent"). This group dependence, and its implications, will be examined in more detail in the following sections. Refer to Appendix A for a more technical discussion of the statistical relationship between crime rate and population size.

Methods

In the year 2000, there were 1,294 cities in the United States with populations over 25,000 that reported crime statistics to the UCR Program (Federal Bureau of Investigation, 2001). These cities were located in all corners of the U.S. and everywhere in between. To add some perspective on this group of cities, the list included Newton, Connecticut, Homewood, Alabama, Belmont, California, and Smyrna, Tennessee, all with populations just over 25,000. It also included cities like Pittsburgh, Pennsylvania and St. Louis, Missouri with populations in the 300,000 range. In addition, it also included the largest cities in the U.S., such as New York (population 8,008,278), Los Angeles (population 3,694,820), Chicago (population 2,896,014), Houston (population 1,953,631), and Philadelphia (population 1,517,550). Using the 2000 UCR crime data from these cities, the author examined the relationship between crime rate and population size. These 1,294 cities were first examined together as one group and then disaggregated into one of four population strata defined by the FBI: Group 1, population 250,000 and above (N=68); Group 2, population 100,000 to 249,999 (N=166); Group 3, population 50,000 to 99,999 (N=371); and Group 4, population 25,000 to 49,999 (N=689).

The author calculated area crime rates ($\hat{R}$) and average crime rates ($\bar{R}$) as a first step in examining the direction of the relationship between crime rate and population size. He, then, subjected the data to a linear regression to confirm the direction and determine the strength of the relationship between the two variables. The findings are examined and discussed in the following sections.

Findings

As described earlier, the correlation between crime rate and population size was expressed through

Table 1
Comparing the calculations of crime rate in various size U.S. cities[a]

Population group	Number of jurisdictions	Area crime rate $\hat{R}$	Average of crime rates in area $\bar{R}$	$\hat{R}-\bar{R}$
Cities 25,000 +	1,294	5,419	4,519	900
Group 1-250,000 +	68	6,283	6,932	−649
Group 2-100,000 to 249,999	166	5,663	5,486	177
50,000 to 99,999	371	4,634	4,596	38
25,000 to 49,999	689	4,163	4,155	7

[a] $\hat{R}$ was calculated by summing the total number of crimes in a large area consisting of n agencies, then dividing by the total population of the entire area. $\bar{R}$ was the average of the crime rates for the n agencies in the area. $\hat{R}-\bar{R}$ was the difference between the two ways of calculating crime rate.

the relationship between two types of crime rates, the area crime rate ($\hat{R}$) and an average crime rate ($\bar{R}$). The difference between the area crime rate and the average crime rate in all 1,294 cities in this study was 900 (see Table 1). This was a positive number which meant the correlation between crime rate and population in these cities was positive. Compare this with the findings in the other city groupings. When the top 68 cities were considered together (i.e., Group 1 cities) the direction of the relationship was reversed. In this stratum, when the average crime rate was subtracted from the area crime rate the result was −649, indicating that the higher crime rates could be located in the less populated jurisdictions in the stratum. In the other three population strata, the difference between these two crime rates was positive, giving the preliminary indication that the relationship was positive.

In Table 1, the direction of the relationship between crime rate and population size was established. The next step was to see if this relationship was significant. In order to assess the strength of the relationship between crime rate and population size, the author conducted a linear regression of crime rate, violent crime rate, and property crime rate on population, first, in all cities together and, then, in each of the population strata. The results of this analysis are outlined in Table 2.[5]

When all cities were examined together (N=1,294), the relationship between crime rate and population size was positive and statistically significant. This finding was true for all three types of crime rates. This meant that higher crime rates in these 1,294 jurisdictions were found in the places that had the largest populations; therefore, as the population size rose in these cities, one would expect to find higher crime rates. This observation, however, did not hold when examining only the Group 1 cities (i.e., the largest sixty-eight cities in the U.S.). Consistent with the findings presented in Table 1, the relationship between crime rate and population appeared to be negative, but as one could see it was *not* statistically significant (p=.09). The nature of this relationship became clearer when the overall crime rate was disaggregated into violent and property crime rates. Violent crime rates in Group 1 cities were clearly not related to the size of the population (p=.643), but the property crime rates were related to the population size in a negative direction(p=.029). This meant that in the Group 1 cities, the highest property crime rates

Table 2
Results of linear regressions of crime rates on population sizes in U.S. cities with populations over 25,000

Population group	Number of jurisdictions	Crime rate	Violent crime rate	Property crime rate
Cities 25,000+	1,294	r = .109 β = .000947 $p.$ = .000	r = .200 β = .000300 $p.$ = .000	r = .084 β = .000646 $p.$ = .002
Group 1-250,000+	68	r = −.207 β = −.000438 $p.$ = .090	r = .057 β = .000028 $p.$ = .643	r = −.265 β = −.000467 $p.$ = .029
Group 2-100,000 to 249,999	166	r = .277 β = .016020 $p.$ = .000	r = .244 β = .002391 $p.$ = .002	r = .266 β = .013630 $p.$ = .001
50,000 to 99,999	371	r = .079 β = .013700 $p.$ = .128	r = .156 β = .004799 $p.$ = .003	r = .058 β = .008906 $p.$ = .263
25,000 to 49,999	689	r = .014 β = .004800 $p.$ = .717	r = .047 β = .002539 $p.$ = .219	r = .007 β = ..002261 $p.$ = .848

Source: Federal Bureau of Investigation (2002).

were observed in jurisdictions with the smaller populations, and that as population size decreased, one should expect to see an increase in the property crime rate.

In the Group 2 cities (N=166) the relationship between crime rate and population size was positive and significant. This finding held true for all three crime rates. Therefore, in this group of agencies, larger population sizes predicted higher crime rates. In the Group 3 jurisdictions (N=371), only the violent crime rate was related to population size, and the relationship was positive (p=.003). In the Group 4 jurisdictions (N=689) the crime rate did not appear to be related to population size at all.

The findings presented in Table 2 indicated that the relationship between the crime rate and the population size was statistically significant in Group 2 cities, but not in the larger Group 1 cities. In order to consider the practical implications of this finding, the author examined twenty cities randomly selected from the Group 1 and Group 2 strata (ten from each). These cities are presented in Table 3. One of the first observations to note was the difference between the average crime rates in these two population groupings. As one might expect, the average crime rate in the Group 1 cities (6,932) was significantly higher than the average crime rate in the Group 2 cities (5,486) (t=4.33, p=.000). Based on the linear regressions described earlier, the author calculated expected crime rates for each of the twenty cities in this sample and, then, calculated a ratio of the expected crime rate to the average crime rate (see last column of Table 3). This analysis highlighted an important difference between these two groupings. In the sample of Group 1 cities, i.e., where the relationship between crime rate and population size was not significant, the expected crime rates differed from the average crime rate by no more than 3 percent (e.g., Mobile, Alabama=1.03). In the sample of Group 2 cities, i.e., where the relationship between crime rate and population size was significant, however, the difference between the expected crime rates and the average crime rates was much larger, from 11 percent lower than the average (in Carrollton, Texas) to 21 percent higher than the average (in Rochester, New York).

Discussion and conclusions

The purpose of this research was to establish the statistical relationship between crime rate and population size and discuss its implications. The author demonstrated that crime rate and population were clearly related, and that the extent to which they were related was dependent on the group of jurisdictions under consideration. This relationship was expressed mathematically as a measure of covariance found through the different ways of calculating crime rate within a prescribed set of agencies. The difference

Table 3
Comparing expected crime rates with average crime rates in a random sample of Group 1 and Group 2 cities

Group 1 agencies	Population	Crime rate	Average crime rate	Expected crime rate	Ratio exp/average
Mobile, AL	254,779	6,913	6,932	7,149	1.03
Anaheim, CA	328,014	3,021	6,932	7,117	1.03
Santa Ana, CA	337,977	3,093	6,932	7,113	1.03
Sacramento, CA	407,018	6,717	6,932	7,083	1.02
Miami, FL	362,470	10,968	6,932	7,102	1.03
Jacksonville, FL	735,617	6,943	6,932	6,939	1.00
New Orleans, LA	484,674	6,979	6,932	7,049	1.02
Detroit, MI	951,270	10,067	6,932	6,844	0.99
Charlotte, NC	625,810	7,904	6,932	6,987	1.01
Corpus, Christi, TX	277,454	7,212	6,932	7,139	1.03
Group 2 agencies	**Population**	**Crime rate**	**Average crime rate**	**Expected crime rate**	**Ratio exp/average**
Tempe, AZ	158,625	9,587	5,486	5,657	1.03
Fremont, CA	203,413	2,610	5,486	6,375	1.16
Glendale, CA	194,973	2,518	5,486	6,239	1.14
Orlando, FL	185,951	12,030	5,486	6,095	1.11
Naperville, IL	128,358	1,788	5,486	5,172	0.94
Worcester, MA	172,648	5,143	5,486	5,882	1.07
Rochester, NY	219,773	7,849	5,486	6,637	1.21
Eugene, OR	137,893	7,181	5,486	5,325	0.97
Carrollton, TX	109,576	3,385	5,486	4,871	0.89
Beaumont, TX	113,866	7,239	5,486	4,940	0.90

between the area's crime rate ($\hat{R}$) and the average crime rate per agency ($\bar{R}$) can result in a non-zero correlation between crime rate and population size. Positive values in the difference between these two types of crime rates ($\hat{R}-\bar{R}$) indicates that the relationship between crime rate and population size was positive; negative values indicated that the relationship was negative. A small difference between these two types of crime rates would indicate a zero correlation between crime rate and population size.

The analysis of the 2000 UCR crime data for the 1,294 cities with populations over 25,000 revealed a significant positive relationship between crime rate and population size, indicating that the higher populated cities reported the higher crime rates. When these cities were disaggregated into population strata and the crime rate was disaggregated into violent and property crimes, however, the relationship between crime rate and population size became a little more complex. In some groupings of cities, the relationship between crime rate and population size could be positive (e.g., the Group 2 agencies), in others it could be negative (e.g., the Group 1 agencies), and in other groupings there was no correlation between crime rate and population (e.g., the Group 4 agencies).

The first implication of these findings related to the practice of comparing (and ranking) jurisdictions according to their crime rates. As stated at the outset, the arguments generally advanced by the FBI and others to discourage this practice made more sense when crime rate was an independent variable, i.e., one of many things to consider when trying to assess something like "quality of life" in a city. Based on the findings in this study, the practice of comparing jurisdictional crime rates would be strongly discouraged until one could establish the nature and extent of the relationship between crime rate and population. When a significant positive or negative relationship existed between these variables, comparisons of crime rates should *not* be made unless adjustments were made for the size of the population. What made this assessment difficult, however, was that the relationship between crime rate and population size was not constant, i.e., it depended on the group of jurisdictions involved in the study.

Another implication of these findings related to the handling of missing or incomplete crime data. This was especially relevant when dealing with UCR, because it was a *voluntary* program. Each year, somewhere between 4 and 10 percent of U.S. law enforcement agencies either *did not* participate or had their data rejected by the FBI during standard quality control checks. These missing data could cause problems for researchers who must develop justifiable methods for filling in the missing numbers for these agencies. The practice of imputing crime data to the missing agencies was well known within the field of criminology, and often used by researchers and analysts who worked with secondary data sets (Reidel, 2000). Likewise, agencies charged with managing large crime data sets like the FBI and the National Center for Juvenile Justice (NCJJ), must also address issues relating to imputation because their aim is to obtain global totals. At the present time a popular cross-sectional method for imputing crime data–known as ratio estimation– was being used by the FBI (Federal Bureau of Investigation, 2000).[6] The problem with this method, based on what the researcher showed about the relationship of crime rate to population size, was that it did not give appropriate adjustment weights to the larger and the smaller agencies within the stratum. One way to apply such weights would be to base an estimate of a missing jurisdiction's crime rate on a linear regression of crime rate on population size.

A third implication of this study applied to researchers and crime analysts who often want to detect jurisdictions whose crime rates extended outside the normal range. A common practice was to subtract the jurisdiction's observed crime rate from the *average* stratum crime rate and divide by the standard deviation. This resultant score (a Z score) quickly assessed how a particular agency fit in relation to the other agencies in the stratum. A Z score of 0 meant that the agency's crime rate was the same as the stratum's average crime rate. A score of 1 indicated that the agency was one standard deviation above the stratum average, while a score of 1.96 or above indicated that the jurisdiction's crime rate was outside of the ninety-fifth percentile of agencies in the stratum and could be considered an outlier. Based on the findings in this study, it could easily be seen that subtracting the agency's crime rate from the average stratum rate could be problematic when there was a significant positive or negative relationship between crime rate and population size. A better practice in these situations would be to base this analysis on the expected crime rate given a certain population size.

In conclusion, although the impact of population size on crime rate may appear relatively small, it can be statistically significant. Therefore, understanding the exact nature of the relationship between these two variables may be helpful to criminologists and criminal justice professionals who use UCR crime data for research and analysis, who manage large crime databases, and who sometimes must make *responsible* comparisons of jurisdictional crime rates.

Acknowledgements

The author would like to thank Dr. Yoshio Akiyama for his thoughtful insights and assistance on the more substantive and technical issues in this article. He would also like to thank an anonymous reviewer for suggestions that helped make the article more focussed and readable.

Appendix A

The statistical relationship between crime rate and population size was found in the two alternative methods for calculating a crimes rate in a given geographic area. $\hat{R}$ represented the area crime rate, i.e., $\frac{C}{P}$ where C = the total number of crimes reported in the specified area consisting of n agencies, and P = the total population of the area. $\bar{R}$, on the other hand, was the average crime rate per agency, i.e., $\frac{R}{n}$, were R = the sum of jurisdictional crime rates and n = the number of jurisdictions under study.

Let p represent the variable P_i (the population of a particular jurisdiction) and r represent the variable R_i (the crime rate in the same jurisdiction). The covariance γ between the two variables, p and r, is then expressed by

$$\gamma = cov(p,r) = \frac{1}{n}\sum_{l}^{n}\left(P_i - \bar{P}\right)\left(R_i - \bar{R}\right)$$

$$= \bar{C} - \bar{P}\bar{R} = \bar{P}\left(\frac{C}{P} - \bar{R}\right) = \bar{P}\left(\hat{R} - \bar{R}\right). \quad (1)$$

Again, it was important to note that $\hat{R}$, the crime rate for the combined total of n agencies, was an average crime rate (averaged through the population-weights P_i/P, while $\bar{R}$ was an arithmetic mean of the agencies' crime rates. The latter was obtained by assigning the uniform weight $(1/n)$ to each agency. Therefore, both represented averaged crime rates, although averaged through different weights, i.e.,

$$\hat{R} = \sum_{i=l}^{n} w_i R_i, \ where\ w_i = \frac{P_i}{P}.$$

$$\bar{R} = \sum_{i=l}^{n} \omega_i R_i, \ where\ \omega_i = \frac{1}{n}. \quad (2)$$

According to Eq. (1), the difference $(\hat{R}-\bar{R})$ multiplied by $\bar{P}$ (the averaged agency population) is the covariance between p and r. Therefore, the correlation between the population p and the crime rate r lay in (or expressed through) the difference of the two types of crime rates $\hat{R}$, and $\bar{R}$. The difference between these two crime rates determined the strength and sign (+ or −) of the relationship between crime rate and population size. $\hat{R}$ would be greater than $\bar{R}$ in situations where the more highly populated jurisdictions tended to have the higher crime rates. This was the result of the larger agencies receiving more weight than the smaller agencies as in Eq. 2, i.e., $(P_i/P)R_i$). On the other hand, $\hat{R}$ would be less than $\bar{R}$ in situations where the less populated agencies showed a tendency to have the higher crime rates. And when the crime volume was proportional to jurisdictional populations $\hat{R}$, and $\bar{R}$ would be approximately equal. The significance of this situation was that population p and crime rate r would be positively correlated when $(\hat{R}-\bar{R})$ resulted in a positive number.

The researcher now adds a few more notations below.

σ_p^2 = the variance of the population variable p,
σ_r^2 = the variance of the crime rate variable r, and
$\rho = (\gamma)/(\sigma_p\sigma_r)$ = the correlation coefficient between p and r.

Consider an ordinary linear regression of r on p:

$$r = \alpha + \beta p, \quad (3)$$

where the constants α and β are given by the following formulas:

$$\beta = \frac{\gamma}{\sigma_p^2} = \rho\left(\frac{\sigma_r}{\sigma_p}\right) = \left(\hat{R} - \bar{R}\right)\left(\frac{\bar{P}}{\sigma_p^2}\right), \ where\ \bar{P} = \frac{P}{n}. \quad (4)$$

$$\alpha = \bar{R} - \beta\bar{P}. \quad (5)$$

It was noted that the first factor $(\hat{R}-\bar{R})$ in β in Eq. (3) expressed a measure of association between p and r, while the second factor $\bar{P}/\sigma_p^2$ was an adjustment of scale. Therefore, in β, the first factor $(\hat{R}-\bar{R})$ carried the information with respect to the association between p and r.

The following comments were made with respect to the association between population (p) and crime rate (r):

(1) According to the regression line in Eq. (3), an increment Δp in population was accompanied by an increase in the agency's crime rate, on the average, by $\Delta r = \beta(\Delta p)$. It should be noted, however, that the coefficient β was not a universally applicable constant in that it depended on the group of agencies under study (i.e., it was "group-dependent"). This group-dependence of β raised difficulty when an agency's crime rate was compared to another agency. If the difference between the two agencies to be compared was Δp, then their

crime rates were not comparable unless adjusted by the factor $\Delta r = \beta(\Delta p)$. As mentioned above, the difficulty here was that β was not a universally applicable constant.

(2) Although one might, for example, resort to compare agencies for a given population size P_o by the statistic $Z_i = (R_i - E(r\_P_o))/(\sigma(r\text{u } P_o))$, this was also group-dependent and the value of Z_i depended on the group of agencies chosen to compute the regression line.

(3) If variables were relativized as $\eta = r/\bar{R}$ and $\xi = r/\bar{P}$, the covariance between ξ and η was given by the formula:

$$\breve{\gamma} = cov(\xi, \eta) = \frac{\gamma}{\bar{P}\bar{R}} = \frac{(\hat{R} - \bar{R})}{\bar{R}}. \qquad (6)$$

Therefore, the covariance $\breve{\gamma}$ was the rate of change from $\bar{R}$ to $\hat{R}$ and hence was a relativized quantity.

(4) If ς represented the variable "number of crimes" whose observations were denoted by C_i, then $r = \varsigma/p$. Therefore, the regression line, Eq. 3, suggested a quadratic relationship such as $\varsigma = \alpha p + \beta p^2$.

(5) A positive association was suggested by positive values in $\hat{R} - \bar{R}$, β, ρ, or $\breve{\gamma}$. It was through testing significance to establish that such association was significant. It was known that the test of independence (H_0: $\rho = 0$) was a test of the regression coefficient (H_0: $\beta = 0$) (Stuart & Ord, 1994).

Notes

1. The UCR Program used the following eight "index crimes" to gauge the annual fluctuation of crime in the United States: murder and non negligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny-theft, motor vehicle theft, and arson.

2. Crime rates were expressed per 100,000 in the population.

3. This fact was first established by Dr. Yoshio Akiyama, the chief statistician at the UCR Program, and relayed to this author by way of a personal communication.

4. Whether the slope was significant, however, was a matter for statistical testing.

5. Crime rate included murder and non-negligent manslaughter, forcible rape, aggravated assault, robbery, burglary, larceny/theft, and motor vehicle theft. Violent and property crime rates were subsets of the crime rate. Violent crime rates included murder and non-negligent manslaughter, forcible rape, aggravated assault, and robbery. Property crime rate included burglary, larceny/theft, and motor vehicle theft.

6. Following this method, FBI officials simply applied the average crime rate within a population stratum to a missing agency. Crime estimates using the ratio estimation method were calculated in the following way: $C_{estimate} = \hat{R}P_i/100{,}000$, where $\hat{R}$ was the estimated crime rate of the stratum and P_i was the population of the missing agency for which the crime estimates were being generated.

References

Black, D. (1970). Production of crime rates. *American Sociological Review*, *35*(4), 733–748.

Blau, J., & Blau, P. (1982). The cost of inequality: Metropolitan structure and violent crime. *American Sociological Review*, *47*(1), 114–129.

Brantingham, P. L., & Brantingham, P. J. (1994). Crime analysis using location quotients. In D. Zalm & P. Cromwell (Eds.), *Proceedings of the International Seminar on Environmental Criminology and Crime Analysis* (pp. 83–94).

Federal Bureau of Investigation. (2000). *Refined quality control, imputation, and estimation algorithms*. Unpublished manuscript.

Federal Bureau of Investigation. (2001). *Crime in the United States, 2000*. Washington, DC: U.S. Government Printing Office.

Federal Bureau of Investigation. (2002). *Crime in the United States, 2001*. Washington, DC: U.S. Government Printing Office.

Gibbs, J. P., & Erickson, M. L. (1976). Crime rates of American cities in an ecological context. *American Journal of Sociology*, *82*(3), 605–620.

Gove, W., Hughes, M., & Geerken, M. (1985). Are Uniform Crime Reports a valid indicator of index crimes? An affirmative answer with minor qualifications. *Criminology*, *23*, 451–501.

Green, D., Strolovitch, D., & Wong, J. (1998). Defended neighborhoods, integration, and racially motivated crime. *American Journal of Sociology*, *104*(2), 372–403.

Kitsuse, J., & Cicourel, A. (1963). A note on the use of official statistics. *Social Problems*, *11*, 131–139.

McCleary, R., Nienstedt, B., & Erven, J. (1982). Uniform Crime Reports as organizational outcomes: Three time series experiments. *Social Problems*, *29*, 361–372.

Merton, R. (1957). *Social theory and social structure, revised*. New York: The Free Press.

O'Brien, R. (1983). Metropolitan structure and violent crime: Which measure of crime. *American Sociological Review*, *48*, 434–437.

Ogburn, W. (1935). Factors in the variation of crime among cities. *Journal of the American Statistical Association*, *30*(189), 12–34.

Regoli, R., & Hewitt, J. (2000). *Delinquency in society.* Boston: McGraw Hill.

Reidel, M. (2000). *Research strategies for secondary data: A perspective for criminology and criminal justice*. Thousand Oaks, CA: Sage.

Reiman, J. (2001). *The rich get richer and the poor get prison: Ideology, class, and criminal justice*. Boston: Allyn and Bacon.

Schmalleger, F. (1999). *Criminal justice today.* Upper Saddle River, NJ: Prentice Hall.

Schuessler, K. (1962). Components of variation in city crime rates. *Social Problems*, *9*, 314–323.

Siegel, L., & Senna, J. (2000). *Juvenile delinquency: Theory, practice, and law* (7th ed.). Belmont, CA: Wadsworth/Thompson Learning.

Skogan, W. G. (1974). The validity of official crime statistics: An empirical investigation. *Social Science Quarterly*, *54*, 25–28.

Skogan, W. G. (1975). Measurement problems in official and survey crime rates. *Journal of Criminal Justice*, *3*, 17–32.

Stuart, A., & Ord, J. K. (1994). *Kendall's advanced theory of statistics*. New York: Halsted Press.

Tarling, R. (1986). Statistical applications in criminology. *Statistician*, *35*(3), 369–388.

Watts, R. (1931). The influence of population density on crime. *Journal of the American Statistical Association*, *26*(173), 11–20.