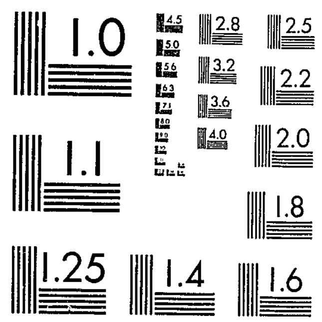

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

3-31-83

SOCIAL FORCES

AN INTERNATIONAL JOURNAL OF SOCIAL RESEARCH

VOLUME 60: NUMBER 3 MARCH 1982

81800-
81803

F ... ety

...a
...n

A ... a pioneering academic research organization and of the work of its distinguished founder, Howard W. Odum. The Institute for Research in Social Science, founded in 1924, quickly achieved a national reputation for exploring the controversial topics that impeded the movement of the South into the national mainstream.

Research in Service to Society describes the origin and early struggles of the Institute, the mechanics of its operation, and the researches of its staff members on such subjects as race relations, labor, farm tenancy, prison reform, and local government. Special attention is given to Odum's development of the concept of regionalism, as well as to the Institute's transformation in more recent decades in response to the use of computer technology in social science research. approx. 440 pp., $20.00

The University of North Carolina Press

Post Office Box 2288 Chapel Hill, NC 27514

MARCH 1982 VOLUME 60 NUMBER 3

Contents

9KB

ARTICLES

81802

Problems in Ratio Correlation: The Case of Deterrence Research*

CHARLES H. LOGAN, *University of Connecticut*

ABSTRACT

Potential problems in ratio correlation cannot be resolved outside a particular substantive context. Within the context of deterrence research, several approaches are examined: the "conceptual-meaning" resolution, the Pearsonian approximation formula and null comparison, simulation techniques, decomposition into component covariances, part correlation, and the use of residual scores. A simulation experiment shows that when the terms used in the measures of certainty of imprisonment and crime rate are randomly scrambled, the resulting ratios correlate in a manner comparable to what occurs with the data in their original form. These scrambled-data correlations, however, are due purely to artifactual effects of the common term. The most useful test for the existence of this common-term artifact appears to be the technique of part correlation. With empirical imprisonment data, the part correlations are lower than the zero-order correlations, supporting the possibility that the original correlations may have been at least partially artifactual.

Much sociological research consists of relating complex indexes to each other. These indexes are formed by combining separate but related indicators in some fashion, by adding, subtracting, multiplying, and dividing them in various ways. Such indexes are indispensable to many kinds of social scientific research, and it is therefore important to consider carefully the ways in which the construction of our indexes can affect the validity and interpretation of our research. This paper will consider in detail one such type of index construction: the formation of ratios, with specific attention to problems that may arise when attempting to correlate two ratios in which the numerator of one is the same as the denominator of the other. In one rapidly expanding body of research, the deterrent or preventive effects of legal sanctions on crime, the strongest findings of most of the major

*Expanded version of a paper presented to the American Society of Criminology, 1977. This research was supported in part by NIMH Grant #MH 24574-01 and the University of Connecticut Research Foundation. Computer time was provided by the University of Connecticut Computer Center.

studies involve correlations between two ratios having exactly this kind of common-term problem. This research will be used for illustration throughout the discussion.

Although recognition of the problems involved in correlating ratios having common terms dates back to the turn of the century (Pearson), it has only very recently been given concentrated attention by sociologists. Excellent general discussions of problems in the correlation of ratios or difference scores having common terms have been provided, with a range of sociological examples (Fuguitt and Lieberson; Long; Schuessler a,b). The discussion has recently reached practically a stage of debate (Freeman and Kronenfeld; Kasarda and Nolan; MacMillan and Daft). No single set of rules or formulas can be derived from existing statistical discussions that will be adequate for all data based on ratios with common terms. For this reason, the discussion of problems in ratio correlation will be applied to one set of data in one substantive context (deterrence).

The Problem of Spurious Ratio Correlations in Deterrence Research

In critiques of deterrence research thus far (Greenberg; Nagin), relatively little attention has been paid to one very basic difficulty common to much of this research. The early core of deterrence research consisted of a group of studies that discovered a remarkably consistent negative relationship between crime rate and some measure of certainty of criminal sanction (punishment) for crime. These studies all used cross-sectional data aggregated by state, county, or city units. The sanctions they dealt with were *execution* (Ehrlich, c), *imprisonment* (Antunes and Hunt; Bailey et al.; Bean and Cushing; Chiricos and Waldo; Ehrlich, a, b; Erickson and Gibbs; Gibbs; Gray and Martin; Kobrin et al.; Logan, b; Tittle; Vandaele), and *arrest* (Block; Chapman; Kobrin et al.; Logan, c; Phillips and Votey; Tittle and Rowe). Thus, units with higher rates of punishment (arrest, imprisonment, or execution) tended to show lower rates of crime, and vice versa.

The problem common to these studies consists of an alleged definitional dependency, or contamination, between the measures of the independent and dependent variables. The independent variable, certainty of sanction, consists of the ratio of sanctions to crimes, denoted here as A/C. ("A" is used because in most studies the sanction is admission to prison or arrest clearances.) The dependent variable is per capita crime rate, denoted as C/P. The common term, C, has been either identical or very highly correlated between the two ratios, in the data analyzed by the studies described above.

The question at issue here is whether an increase in the common term, C, might tend both to deflate the value of A/C and inflate the value of C/P. If so, this would produce a negative correlation, as an artifactual effect, in the same direction as the relation predicted by the deterrence hypothesis. Thus, the first problem in interpreting the negative correlations found in deterrence research is to try to determine whether they are real or artifactual. (Other problems, such as the direction of causal effect, also exist in the deterrence research but are not discussed here.)

One data set from the literature described above (Logan, b) will be used for illustration. These particular data were chosen for their availability and because they are similar in character—and in some cases overlapping or identical—to data analyzed by the several imprisonment studies cited above. However, the analysis here will carry implications for all deterrence studies that measure certainty of sanction by means of a ratio of sanctions administered to crimes known, whether the sanction be arrest, conviction, imprisonment, or execution.

Responses to the Problem of Spuriousness in Ratio Correlation

The problem of correlating ratios having common terms has been addressed specifically in the deterrence literature (Chiricos and Waldo; Logan, a, b, c; Tittle) as well as in more general methodological literature (Bollen and Ward; Freeman and Kronenfeld; Fuguitt and Lieberson; Kasarada and Nolan; Long; MacMillan and Daft; Pearson; Schuessler, a, b). Several resolutions that have been suggested will be critically examined as to difficulties in their application and interpretation:

1. The *conceptual-meaning* resolution.
2. The Pearsonian null comparison.
3. Simulation techniques.
4. Decomposition into component covariances.
5. Part correlation and the use of residual scores.

1. THE CONCEPTUAL-MEANING RESOLUTION

One early and continuing disclaimer has been that if our primary conceptual and theoretical interest is in the ratios per se, rather than in the components that make up the ratios, then the issue of spuriousness in the correlation of two ratios does not logically arise, even if they do have a common term (Fuguitt and Lieberson; Kasarda and Nolan; Kuh and Meyer; McNemar; Rangarajan and Chatterjee; Schuessler, a; Yule). Spuriousness, according to this argument, becomes a problem only when the common term is used to standardize, or otherwise adjust, both of the two major variables of interest (e.g., city taxes and city expenditures, both being divided by population size). If our hypothesis is stated in terms of the correlation between two ratios, we cannot legitimately test it with correlations between the unadjusted component terms, nor vice versa, since correlations between ratios will generally differ from the correlations between the component terms. With regard to the deterrence data, it has been argued (Logan, a, b) that in the ratio A/C, neither component is conceptually sufficient by itself; the deterrence hypothesis refers to their relation to each other and not to the absolute values of either component. An analogy could be made to the well-known positive association between speed (miles per hour) and gas consumption (gallons per mile), where we are also not interested in the relation among the components but in the relation between the ratios. In this example, it is clear that the ratio variables are not mere epiphenomena, created by the artifact of dividing one measurement by another, but are conceptually meaningful variables in themselves. Further, we expect these ratios to have a causal correlation opposite in sign to what would be expected as a result of indexical artifact.

Several recent papers have made strong cases for the argument that there is nothing inherently spurious or biased in the empirical correlation between ratios containing common terms and that where ratios properly express the concepts or variables of interest they ought not be abandoned in favor of some alternative measures (Kasarda and Nolan; Long; MacMillan and Daft a, b). Long for example, demonstrates that a correlation between two ratios can take any sign, depending on the slopes and intercepts of the relations among the components. As will be shown in a later section on decomposition, this is merely an analytic truth. Correlations between ratios are affected by the nature of the components (their variances, covariances, slopes and intercepts) *and vice versa*. This, however, does not tell us whether in any given instance we are to take the sign of the ratio correlation as given, which then constrains the components, or to take as given the constraints on the components, which then affect the sign of the ratio correlation. Thus Long et al. have demonstrated that there is no *necessary* bias in the empirical correlation of ratio variables, but just because bias does not *automatically* occur does not mean that it *cannot* occur.

In sum, the problem with the conceptual or theoretical-meaning resolution, along with the demonstration that common-term ratios are not inherently biased, is not that they are wrong, but that they are insufficient. They do not provided us with any means of distinguishing causal correlations from artifactual ones.

2. THE PEARSONIAN NULL COMPARISON

In perhaps the earliest treatment of this problem, Karl Pearson developed a formula giving an approximation of the correlation between two ratios in terms of the correlations and rel-variances (or coefficients of variation) of the component terms. The formula for the general case is:

$$r_{(Y/Z)(X/W)} = \frac{r_{YX}V_YV_X - r_{YW}V_YV_W - r_{XZ}V_XV_Z + r_{ZW}V_ZV_W}{\sqrt{V^2_Y + V^2_Z - 2r_{YZ}V_YV_Z}\ \sqrt{V^2_X + V^2_W - 2r_{XW}V_XV_W}} \tag{1}$$

where V_X = the rel-variance (or coefficient of variation) of X (i.e., the standard deviation divided by the mean). Formulas for the correlation of ratios having common terms can be easily derived from this basic formula. For the case of A/C and C/P we have, with appropriate changes in notation:

$$r_{(A/C)(C/P)} = \frac{r_{AC}V_AV_C - r_{AP}V_AV_P - V^2_C + r_{CP}V_CV_P}{\sqrt{V^2_A + V^2_C - 2r_{AC}V_AV_C}\ \sqrt{V^2_C + V^2_P - 2r_{CP}V_CV_P}} \tag{2}$$

The most commonly suggested approach (Chayes; Fuguitt and Lieberson; Kuh and Meyer; Pearson; Rangarajan and Chatterjee; Yule; Yule and Kendall) to the problem of spuriousness in ratio correlation involves a manipulation of formula (2) to produce a null value representing the correlation between A/C and C/P that would be expected to occur under the condition that the components (A, C, and P) are uncorrelated. If $r_{AC} = r_{CP} = r_{AP} = 0$, formula (2) becomes:

$$\frac{-V^2_C}{\sqrt{V^2_A + V^2_C}\ \sqrt{V^2_C + V^2_P}} \tag{3}$$

a value that will obviously be negative. Pearson referred to this value as the "spurious correlation," while Chayes refers to it as the "null correlation." Comparing the values computed by (2) and (3) provides a method of determining whether the empirical correlation approximated by (2) is larger than the correlation that would be obtained from these data if the ratio components were uncorrelated, as estimated by (3). This null comparison is supposed to be analogous to a test of the ordinary null hypothesis that $r = 0$. Thus, an empirical correlation larger than the value from (3) would be needed to reject the null hypothesis of a non-artifactual correlation between the ratios.

As an illustration, if we applied these formulas to our deterrence data (Logan, b), for the case of total felonies, the approximation formula (2) gives $r = -.79$ while the null formula (3) gives $r = -.68$. The (approximate) ratio correlation is larger than the null value.

To use this comparison to test whether the ratio correlation is real or artifactual, however, would be misleading, for two reasons.

First, the ratio correlation approximated by formula (2) is seriously in error when applied to the data in the example above. Using decimal values for A/C and C/P, the standard Pearsonian correlation coefficient (exact) is $-.44$, compared to the approximation formula value of $-.79$. Because Pearson's formula was derived from a binomial expansion in which terms higher than the second order were dropped, it is an approximation that becomes subject to serious error when any of the values of V rise above *0.15* (Chayes, 15). In the deterrence data on imprisonment used here for illustration, V ranges from *0.94* to *1.86*. For the "total felonies" data used just above, V_A, V_C, and V_P are *0.94*, *1.45* and *1.06*, respectively.

The second, and more important, problem with the Pearsonian procedure is that, strictly speaking, the null comparison does not directly test the hypothesis of noncorrelation between two ratios; rather, it tests the hypothesis of noncorrelation among the *components*. Finding a correlation larger than the null value, and thus rejecting the null hypothesis, should therefore only mean rejection of the assumption that the components are uncorrelated in the population. We cannot reject the possibility that the ratios may still be correlated only as a function of the common term, because we do not know how strong that artifactual correlation would be under the different condition of correlated components. Moreover, there is a paradox in this procedure: in order to apply formulas (2) and (3) we must already have data on the components; hence, the test is unnecessary, because we will already know whether the components are correlated.

In sum, it is not clear to me that Pearson's null comparison can properly be used as a test of the null hypothesis that a given empirical ratio correlation is not greater than some calculated value that would be expected to occur with those data if the correlation were due entirely to the operation of a common term in the two ratios.

3. SIMULATION TECHNIQUES

Since Pearson's formula provides a poor approximation, Chayes suggests substituting a simulation technique as an alternative means of determining a null value for the correlation of two ratios formed from three uncorrelated components. For a large number of simulated cases, Chayes generates quasi-random values for each of three components in such a way as to reproduce the means and standard deviations that these components have in some empirical sample data. These three components are then combined for each case into two ratios having a common term, and then those constructed ratio values are correlated. Chayes proposes that this correlation provides a proper substitute for the null correlation figured by formula (3). When the coefficients of variation for all components are small, the null correlations computed by formula (3) and by Chayes' simulation technique should be about the same. Otherwise, Chayes' technique is to be preferred. However, since Chayes' simulation technique is directly analogous to the Pearsonian null comparison, the criticisms already made of that comparison also will apply to the simulation technique when it is being used as a substitute for Pearson's null comparison.[1]

One very recent discussion of ratio correlation, using simulated

data, includes an application to the deterrence problem as one of several illustrations. MacMillan and Daft produce simulated data for crimes (C), imprisonments (I), and population (P) to test for bias in the correlation of I/C and C/P. They criticize earlier simulations (Chiricos and Waldo) for generating C, I, and P to be independent of each other, and argue that "the appropriate simulation experiment to test for bias is to have C, P, and I increase at the same rate across jurisdictions." This is the point at which their logic departs from that of all previous simulation research on ratios (Chayes; Chiricos and Waldo; Freeman and Kronenfeld; Logan, a; Tittle), in which the whole purpose is to allow the components to vary independently of one another.

For 100 cases, MacMillan and Daft set $C = a + .10P + u$ and $I = a + .31C + u$, where a = a constant intercept and u = a small random error term. P is set randomly between 10,000 and 100,000. Thus, I is a strongly predetermined linear function of C, which in turn is a linear function of P. By varying the intercepts, they predetermine the relation of the ratios to their components, and thereby to each other. Whenever the intercept for the production of C from P is zero *or* when the intercept for the production of I from C is zero there will be no correlation between I/C and C/P. In the former case, for example, C will always be about 10 percent of P (adjusted by error term), regardless of the size of P. Since C/P will be constrained to remain virtually constant at about *.10*, it will not correlate with I/C (or anything else). MacMillan and Daft seem to believe that only if the impossible occurred and I/C correlated with C/P under these circumstances would we have any evidence of a definitional dependency, or common term, effect.

MacMillan and Daft have overconstrained their data. They have predetermined, in the above case of a zero intercept, that I/C will not vary much from *.10*, regardless of the size of P or C. It may be sensible to place limits on the generation of the separate components, but MacMillan and Daft have placed severe limits on the relations among the components. This goes against one of the main purposes of producing ratios from simulated data, which is to measure the effect of a common term when the components are known to be independent.

In summary, MacMillan and Daft constrain not only the components, but their ratios. They then discover with scatterplots and correlations that the ratios behave in the ways in which they have been constrained to behave. This is used to prove that they could not have behaved otherwise. Hence (they say), the use of ratios is valid and unbiased with respect to the effect of a common term.

There is another way of interpreting a simulation technique, in terms of measurement errors, that seems at least intuitively compelling. In producing quasi-random component values, it would seem desirable not only to reproduce the means and standard deviations of the empirical data, as

Chayes does, but the ranges as well, particularly with data that cannot take on negative values or values above or below a certain range because of some practical or theoretical limits.[2] With all these restrictions on the random generation, it is as if we took the actual empirical data and scrambled the values among the cases, thereby creating massive measurement error.[3] For example, our new "case" representing Alabama might be randomly assigned California's number of prison admissions (*A*), Ohio's number of crimes (*C*), and Nevada's population size (*P*). These components would be random with respect to each other, and the resulting ratios (*A*/*C*; *C*/*P*) would be meaningless.

Thus, if it could be shown that an empirical correlation found between conceptually meaningful ratios formed from real-data components is about the same as that which is found between conceptually meaningless ratios formed by first scrambling the component data, this would certainly raise troublesome questions about the nature of the empirical correlation, even if it could not be said to definitely prove that the empirical correlation is spurious or artifactual.

A Scrambled-data Simulation Technique

Using computer programs for random number generation and a Fortran sorting procedure, the real values for admissions to prison (*A*), crimes (*C*), and population size (*P*) for 48 states were scrambled in a random fashion and then recombined into ratios *A*/*C* and *C*/*P*. These scrambled-data ratios were then correlated and their scatterplots examined. As was the case with the original (unscrambled) data, *A*/*C* and *C*/*P* correlated negatively, at moderate strength, for all felony categories. However, examination of the scatterplots showed that for every felony category a few extreme outliers gave the data a strong curve, asymptotic on both axes. This curve was present in the original data, but became more pronounced with the scrambling and recombining procedure. To adjust for the skewed distributions of *A*/*C* and *C*/*P*, and the curvilinearity of their correlation, log transformations were made on each ratio, both in the original data and in the scrambled data.[4]

The program written to scramble the values of *A*, *C*, and *P* randomly, recombine them into *A*/*C* and *C*/*P*, and correlate these scrambled ratios, was looped 1,000 times. This produced a sampling distribution of the correlations that would obtain between *A*/*C* and *C*/*P* in a hypothetical population where *A*, *C*, and *P* were distributed as they are in the real world but without any correlation to each other. Since in the scrambled data *A*, *C*, and *P* were not related[5] and the resultant ratios *A*/*C* and *C*/*P* had no substantive meaning or causal relation to each other, the sampling distribution of scrambled-data correlations represents correlations due entirely to the artifactual effect of having a term common to both ratios. This sampling distribution can thus be used to test whether the empirical correlations

found in the original data are significantly different from the average correlations that occur artifactually in the corresponding sets of scrambled data. Table 1 presents the results of this test.

Only for sex offenses is the original (real-data) correlation significantly greater (in the predicted negative direction) than what occurs under the conditions of the scrambled-data simulation. For most felonies, the scrambled-data correlations are as strong or stronger than those in the original data. This does not prove that the real-data correlations are artifactual, but it does at least shake our faith in the results of published deterrence research to think that they could be replicated using data so randomly scrambled as to be meaningless. This is more remarkable than a roomful of monkeys pounding eternally on typewriters and eventually reproducing a work of Shakespeare. This is like 1,000 monkeys shredding and recomposing a data set and each one tending, on the average, to reproduce the findings of the original researcher. Does this make researchers out of monkeys? Or monkeys out of researchers?

In any case, this simulation experiment is sufficient to raise serious questions about the reality of at least some of the original ratio correlations that have been published in deterrence research. But although simulation tests may demonstrate a very real possibility of spuriousness in the correlation of deterrence ratios, they do not prove its existence. Simulation tests follow the logic of a syllogism in which the major premise is this: If *A*, *C*, and *P* are constrained in certain realistic ways but are random with respect to each other, then *A*/*C* and *C*/*P* will correlate to a certain degree as a function of their common term, plus the constraints on *A*, *C*, and *P*. Thus if

Table 1. ORIGINAL VS. SCRAMBLED DATA CORRELATIONS BETWEEN CERTAINTY OF IMPRISONMENT (A/C) AND CRIME RATE (C/P) (LOG TRANSFORMED)

	Scrambled-Data Correlations			
Offense	Mean	Standard Deviation	Original-Data Correlation	p-value*
Total felonies	-.55	.08	-.47	.37
Homicide	-.57	.08	-.19	<.001
Sex offenses	-.60	.07	-.76	.02
Robbery	-.62	.07	-.67	.44
Assault	-.69	.05	-.71	.76
Burglary	-.54	.08	-.45	.32
Larceny	-.45	.10	-.26	.04
Auto theft	-.38	.11	-.31	.52

*Probability of a value (at least) as extreme as the original-data correlation, under the sampling distribution of 1,000 scrambled-data correlations.

a significantly different correlation occurs in our empirical data, we can conclude that something else, beyond the common term and the constraints, is affecting the relationship between the ratios. But, as with any syllogism, affirming the consequent does not lead to any valid conclusion. If the empirical correlation is close to what occurs artifactually in simulated data, that suggests the possibility but not the valid conclusion that the empirical results are artifactual.

4. DECOMPOSITION INTO COMPONENT COVARIANCES

Schuessler calls attention to the point that the relation between two ratios can be expressed as a function of the relation among the components with exact rather than approximate results if the variables are first converted by log transformations. This is because $\log (A/C) = \log A - \log C$, and, unlike the correlation of ratios, the correlation of difference terms may be expressed with exact, rather than approximate, results as a function of the moments of the component terms (Schuessler, b, 380). Schuessler notes (a, 217) that with log transformations, the covariance of A/C and C/P can be expressed as

$$\sigma_{12} = (\sigma_{ac} + \sigma_{cp}) - (\sigma^2_c + \sigma_{ap})$$

where σ_{12} = the covariance of $\log (A/C)$ and $\log (C/P)$,

$a = \log A - \overline{\log A},$
$c = \log C - \overline{\log C},$
$p = \log P - \overline{\log P},$
$\sigma_{ac}, \sigma_{cp}, \sigma_{ap}$ = the respective covariances of a, c, and p, and
σ^2_c = the variance of c.

Thus, the covariance of log transformed ratios can be expressed as a direct function of sums and differences of component covariances and the common term variance. When this is done, "it is possible to gauge the statistical weights of the constituent covariances in the covariance of ratios (as logs)" (a, 216).

Like Pearson's formula, Schuessler's technique expresses the correlation of two ratios in terms of the moments of the constituent components, but Schuessler's decomposition procedure has two distinct advantages. First, with the use of log transformations the results are exact, and thus not subject to the occasionally serious distortions of Pearson's approximation. Second, the decomposition procedure does not require the making of unrealistic assumptions in order to determine the contribution of the common term to the ratio correlation.

For the purpose of illustration, Schuessler (a, 217–9) took data on imprisonment and offense rates (for all *Index* felonies combined) for 1966 and analyzed them by this decomposition technique. Plugging his empirically obtained values into formula (4) yielded:

$$-.03 = (.24 + .22) - (.28 + .21).$$

A comparable outcome is obtained with our deterrence data. For total felonies:

$$-.10 = (1.13 + 1.14) - (1.37 + 1.00).$$

Schuessler remarks of his finding that it "reflects the weight of the common term variance (.*28*) relative to the approximately equal covariances. An implication is that a positive relation between the admission rate and the crime rate is unlikely in the absence of a weak relation between the numbers of admissions and the population. Such a relation could occur, but it is not very probable" (a, 219).

It is important to note that Schuessler's conclusion that the negative covariance of the ratios reflects the relative weight of the common term variance and the component covariances could just as easily be stated the other way around. It depends on what is taken as given. (Should we say that $2 + 2 = 4$, or that $4 - 2 = 2$?) We could say that *given a negative relation* between A/C and C/P, we are not likely to find a large variance in c, or a strong covariance between a and p, compared to the other component covariances (when data are transformed to logs). The reversibility or symmetricality of the decomposition procedure is noted by Schuessler in his introduction (a, 203), where he indicates that it may be "instructive to analyze ratio variables in terms of their components, *and vice versa*" (emphasis added), and again in his concluding discussion, where he emphasizes that "the decomposition of moments of ratios into moments of components does not carry the implication that components are the causes of ratios, or vice versa" (226).

However, this point deserves even more emphasis than Schuessler gives it, because anything less than a rigorous reading of the literature on ratio correlations leaves the strong impression that expressing a correlation between ratios in terms of the moments of the components is the same as showing that the former is *due to* the latter, which is not necessarily the case. There is no inherent reason why the (co)variances of components should be said to account for ratio correlations, rather than vice versa.[6]

Analytically, of course, apart from any substantive interpretation, the correlation and the component (co)variances constrain each other. Which way it will be stated in any given substantive example depends on the theoretical context and meaning of the component and ratio variables.

The decomposition exercise is thus analytic, rather than explanatory. Technically, it is tautological, although it may occasionally be a revealing and useful tautology. For this reason, Schuessler cautions that decomposition should not be done for its own sake, but only to test a hypothesis or to clarify an otherwise puzzling relationship.

With regard to the question of a possible artifactuality in the correlation of the deterrence variables A/C and C/P, it is not enough to be able to express the covariance of these ratios (after log transformations) in terms of moments of components, or vice versa. We still want to know what the correlation of A/C and C/P would be if it were not for the statistical effect of the common term—i.e., if the effect of that common term could somehow be removed or adjusted for. This adjustment can be accomplished through partialling, or residualizing, procedures.

5. PART CORRELATION AND USE OF RESIDUAL SCORES

One solution to the common-term problem that has been suggested in the deterrence literature (cf. Logan, b) is the use of part correlation. In part correlation, one variable is related to a second variable from which the effect of a third variable has been removed. Thus, the part correlation $r_{1(2.3)}$ correlates one variable with the residuals of a second, as opposed to the partial correlation $r_{12.3}$, which correlates the residuals of each of the first two variables after they have each been regressed on a third. The general formula for part correlation is:

$$r_{1(2.3)} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r^2_{23}}}, \tag{5}$$

which expresses the correlation between 1 and the residuals of 2 regressed on 3 (Dubois, 60–62; McNemar, 186). Its significance can be tested by the formula:

$$F = \frac{r^2_{1(2.3)}}{(1 - R^2_{1.23}) / (N-3)}, \tag{6}$$

where $R^2_{1.23}$ is the multiple correlation of 1 with 2 and 3 (McNemar, 322). The choice between part and partial correlation may be arbitrary, or it may in some cases be guided by conceptual or theoretical considerations in the specification of the model.[7]

In terms of the present problem, the part correlation $r_{C/P(A/C \cdot 1/C)}$ expresses the correlation of crime rate with certainty of punishment after the effects of the common term, C, have been removed from the certainty measure. $1/C$ is used here instead of C in order to meet the linearity assumption. A/C is a curvilinear function of C, but a linear function of $1/C$ (Fleiss and Tanur, 44).[8]

Fuguitt and Lieberson (140) express reservations about the part correlation technique and conclude that "until more work is done on the rationale for part or partial correlations to remove the effect of a common term we prefer the straightforward Pearson procedure." Citing Fleiss and Tanur, they point out that it can be shown for the special case of data generated such that A, C, and P are unrelated, the value of $r_{A/C\ C/P\ \cdot\ C}$ and of $r_{A/C\ C/P\ \cdot\ 1/C}$ will be 0. Hence, for those data the part correlation $r_{C/P(A/C \cdot 1/C)}$ will also = 0, thus providing a rationale for the use of the part correlation technique *"but under the assumption that [A, C, and P] are independent. Hence it is not clear that Logan has circumvented the objection that an independence assumption is unreasonable"* (emphasis added).

However, this confuses the conditions under which a technique is proven to work, with the conditions under which the technique may be applied. Fleiss and Tanur created random data in order to demonstrate that a partial correlation would be zero when applied to data where it was known that it should be zero. Unlike the Pearson procedure, however, the application of the part correlation to real data does not *require* that we assume, even provisionally for purposes of test, that the ratio components are uncorrelated. Its advantage is that it can be applied to data where the components are known to be correlated.

The rationale for the use of part correlation in the analysis of ratio variables is the same as the rationale for any use of part or partial correlation: (1) Let A, C, and P vary as they do in reality; (2) hypothesize that some part or all of the correlation between A/C and C/P is the spurious result of the simultaneous relation of C to both A/C and C/P; (3) adjust the values of A/C by regressing on $1/C$ and taking the residual scores as our new values; (4) compute a part correlation as expressing the covariation between the two ratios that cannot be attributed to the operation of the variable being statistically controlled. No difference is made by the fact that in ordinary partialling analyses the controlled third variable is thought to be *causally* antecedent to or intervening between the other two variables, whereas in the present problem, the common term is thought to be definitionally or analytically related to the ratio variables of interest, as well as perhaps causally prior to one of them (A/C).

Thus, it would seem that part correlation ought to provide a sensible answer to the question of how A/C and C/P would correlate to the absence of any effect of the common term. In Table 2, columns 1 and 2 compare the zero-order and part correlations as calculated from the original imprisonment data. When adjustments are made for the effects of the common term, the correlations reduce considerably but do not vanish completely.[9]

Columns 3 and 4 of Table 2 illustrate the application of part correlation to the scrambled data, where the values of A/C, C/P, and $1/C$ are known to be meaningless and unrelated except for the effect of the common term. As expected, the part correlations for those data reduce to an average value of zero, though they may depart from zero by chance in any one trial.

Table 2. ZERO-ORDER VS. PART CORRELATIONS BETWEEN CERTAINTY OF IMPRISONMENT (A/C) AND CRIME RATE (C/P), USING ORIGINAL AND SCRAMBLED DATA VALUES (LOG TRANSFORMED)

	Original Data		1000 Scrambled-Data Part Correlations	
Offense	$r_{\frac{A}{C}\frac{C}{P}}$	$r_{\frac{C}{P}}$ (A/C.1/C)	Mean	Standard Deviation
Total felonies	-.47	-.30	.00	.09
Homicide	-.19	-.11	.00	.09
Sex offenses	-.76	-.22	.00	.10
Robbery	-.67	-.20	.00	.08
Assault	-.71	-.13	.00	.08
Burglary	-.45	-.12	.00	.10
Larceny	-.26	-.14	.00	.10
Auto theft	-.31	-.22	.00	.10

In comparing column 2 with columns 3 and 4, it can be seen that several of the original-data part correlations are low enough that their reliability or reality might be questioned. The values for homicide, burglary, assault, and larceny are that low, while the values for total felonies, sex offenses, auto theft, and robbery remain high enough to have confidence in. The standard used here for "confidence" is two standard deviations (column 4) around the mean of zero.

In sum, then, the application of part correlation techniques to these data suggests that for all eight offense categories, the common term may be acting as an artificial inflator of the correlations between sanction rates and crime rates. For about half the felony categories, there may even be no non-chance correlation once the common term is removed.[10]

Part correlation provides a method of testing for the existence and degree of artifactual[11] common-term effects in a simple two-variable analysis. However, it is still possible that the operation of other variables may be obscuring the relation between the two major variables of interest. Thus, a non-zero part correlation might be shown to be spurious by controlling for other variables. Alternatively, a zero part correlation might be shown to be the net result of counteracting effects of some third variable or set of variables that mask a causal relation between the two major variables. The desirability of entering additional variables into the analysis indicates the need for a simple method of adjusting the data to remove common term effects prior to the application of conventional techniques for multivariate analysis. Residualization[12] seems to provide this solution.

It has been argued (DuBois; Fuguitt and Lieberson) that a ratio may be seen as a kind of residualized variable, under certain conditions. Stated

conversely, a residualized score may sometimes provide a substitute for a ratio measure. Fuguitt and Lieberson (138) go so far as to suggest that residual scores will generally be preferable to ratios except under those certain conditions where the two are equivalent. Thus, it might be thought that we could just respecify the measure of certainty of punishment, not in terms of a ratio (*A/C*), but in terms of a residual (*A.C*).

Whether a ratio or a residual is preferable as a measure of a concept, however, depends in good part on the nature of the concept. It is never a purely methodological problem. In the case of certainty of punishment, a residual score would be less faithful to the meaning of the concept being measured. The residuals of *A* regressed on *C* are the values of *A* relative to what one would predict from the value of *C*. The ratio *A/C*, in contrast, refers to the value of *A* relative to the total potential for *A*. It constitutes a proportion that can range from zero to one. This makes the ratio measure much closer to the concept of *certainty* or *probability* of punishment, and thus preferable to the simple residual of *A* on *C*.[13]

It is possible, however, to combine the conceptual advantage of a ratio measure with the statistical benefits of residualization. The solution is simply to remove the effects of the common term from the ratio itself, rather than from the numerator only. Thus, we could use as our independent variable, "residual arrest rate," defined as the residuals of *A/C* regressed on *1/C*. This residual variable, freed of possible measurement contamination with the dependent variable, could then be treated just like any other variable in a multivariate analysis of the determinants of crime rate.[14]

Summary and Conclusions

A recent and rapidly expanding line of deterrence research, which has received much attention with potentially significant policy implications, has involved the cross-sectional correlation of crime rates with measures of certainty of punishment. Where the index used to measure certainty of sanction has been the ratio of prison admissions or arrest clearances to crimes known to police (*A/C*), there arises the possibility of definitional contamination with the dependent variable, crime rate (*C/P*), due to the presence of a common term in the two measures. Arguments to the effect that *A/C* and *C/P* are meaningful as ratios, and that there is no inherent bias in their correlation are valid but insufficient. They cannot rule out the possibility of an artifactual correlation resulting from measurement error in conjunction with the common term in the two ratios.

In a simulation experiment, the results obtained with some previously analyzed deterrence data (Logan, b) were replicated after randomly scrambling the components *A*, *C*, and *P* in such a way that the resulting ratios were meaningless and unrelated, but for the effect of the common term. These results do not prove that the original empirical data are meaningless, nor that previously published empirical correlations are artifactual. However, they do at least raise questions about the reality of the negative correlations between A/C and C/P found in much deterrence research. Therefore, several other techniques for dealing with problems in ratio correlation were examined, with an eye toward their applicability to deterrence data.

Analytic techniques have been used to express the correlation between ratios either approximately, in terms of the moments of the components (Pearson), or exactly, with log transformations, in terms of the (co)variances of the components (Schuessler, a,b). These techniques, however, must not be interpreted as showing that the characteristics of the components *produce* the correlation between the ratios, any more than vice versa. Thus, they do not provide a means of demonstrating that a correlation between two ratios is a spurious result of the operation of a common component. This is not to deny the general utility of these analytic techniques,[15] but only to assert that they do not provide a satisfactory test for common-term spuriousness.

As a test for this source of spuriousness, the technique of part correlation is suggested. Part correlation allows one to remove the effects of a common term from one of the two ratios, thereby correlating the residual variation in that ratio with the full variation in the other ratio. An extension of part correlation to the multivariate case would be to residualize one ratio on the common term, creating a new variable (the residual score). This score would then be used in all further analyses as a measure of the first variable that should have no definitional contamination with the second variable. Additional variables could then be entered into the analysis using standard multivariate techniques.

The capability of part correlation to discount the artifactual effect of a common term was verified by computing part correlations on the randomly scrambled data for A/C and C/P, since in those data it was known that the two ratios correlated solely by virtue of the common term. Across 1,000 trials, the average part correlation between these ratios was *0.00*, as expected. When part correlations were run on the original, unscrambled data, the correlations were greatly reduced from their zero-order levels. For about half the felony categories, the part correlations were within a range where they could have occurred by chance.

Two general conclusions—one substantive and one methodological—may be offered. The substantive conclusion is that confident assertions that deterrence works (e.g., Tullock) are premature. The questions raised in this analysis ought to at least encourage some more systematic examination of the issue of possible artifactuality in the findings of much recent and ongoing deterrence research. The general methodological conclusion is that the most useful test for the existence of a common-term artifact in a

bivariate ratio correlation is the technique of part correlation. Where additional variables are included, the best approach would seem to be to residualize one of the two ratio variables by regressing it on the common term, then to proceed from there.

Notes

1. As an exercise, Chayes' simulation technique was compared to the empirical deterrence correlation for total felonies. For 1,000 simulated cases, values of A, C, and P were randomly generated in such a way as to reproduce the means and standard deviations of these components in the empirical data. The ratios formed by these quasi-random components correlated at *.003*, compared to the empirical correlation of *−.44*. By the same logic that rejected the Pearsonian comparison, however, this ought not to be regarded as a test of the artifactuality of the empirical correlation.
2. Previous attempts in the deterrence literature to use simulation techniques to investigate the possibility of a common-term artifact have either failed to set any limits at all on the simulated data (Tittle), or constrained the simulated components only by the ranges of the real data (Chiricos and Waldo; Logan, a).
3. That measurement error is a key source of bias in ratio correlation is more rigorously documented by Long. Both underestimation and random error in the measurement of C would create bias in the correlation of A/C and C/P, as would overestimation in the measurement of A. That official statistics on crime (C) underestimate is well established. That admissions to prison (A) could be overestimated has been completely overlooked in deterrence research. National Prisoner Statistics on admissions to prison include not only new commitments (which are relevant to certainty of punishment) but also transfers and technical parole revocations (which are not). The type of measurement error produced here, however, is not the normal type of random error, since it applies not to the measurement of the component terms, but to the accuracy of their recombination into ratios. Errors that may exist in the original component measurements still exist, of course.
4. In both cases, scatterplots showed that the transformations made the relations linear.
5. The truth of this was verified empirically with the scrambled data.
6. Thus, even Schuessler is simplifying (by leaving out "and vice versa") when he concludes that "decompositions are a reminder that a restriction on the correlation of ratios inheres in one or more restrictions on the component variables from which the ratios were formed" (b, 394).
7. Where the two ratios involve common deflations or standardizations of different components, then partial correlation seems appropriate, and simpler. Where the conceptual focus is on A relative to B, rather than on B relative to A, then A ought not to be partialled out. In practice, this may generally mean that one should partial common terms out of denominators and not numerators, but perhaps not necessarily. Rather than attempting to establish some abstract methodological rule, it would be better to ask in each case whether it makes conceptual sense to partial the effect of the common term out of the ratio. If it does so for both ratios, use partial correlation; if it makes more sense for one ratio than for the other, use part correlation.
8. The effect of C (or rather, $1/C$, because of the curvilinearity problem) is removed from A/C rather than from C/P because this makes more sense substantively. We are interested in the effects of that part of the variation in certainty of sanction beyond that which is a function of number of crimes.
9. The part correlations may even be understated, since the log transformation procedure does not seem to have the same effect of improving the fit of the part correlations as it had with the zero-order correlations. In fact, log transformations on A/C, C/P, and $1/C$ reduce the values of the part correlations below what is obtained with untransformed data. Using the untransformed data, the values in column 2 of Table 2 would be, in order: *−.34, −.12, −.39, −.45, −.30, −.42, −.33, −.22*. Those values are considerably stronger for most offenses, but still below the corresponding zero-order correlations for all offenses except assault. Using untransformed data, the zero-order correlations, in order, are: *−.44, −.20, −.55, −.56, −.24, −.45, −.38, −.29*.

10. This is in contrast to earlier conclusions based on part correlations (Logan, b), where the effects of log transformations on part correlations were not considered, nor were comparisons made with the application of part correlation to scrambled data.
11. One reviewer suggested that these effects be referred to as structural rather than artifactual. This was based on the reasoning that C would never be purely random with respect to A and P when using real data. Rather, C for any state would probably be smaller than P for the smallest state but bigger than A for most states. This structural constraint would inflate the negative correlation between the two ratios. This is plausible, but it does not seem important to me at this point to develop the distinction between artifactual and structural effects. The technique of part correlation would serve as a check on their existence under either name.
12. Residualization refers to the procedure of regressing one variable on another and creating a new variable (the residual) equal in value to the difference between predicted and actual values of the first variable. Thus, "residual Y" is the deviation in Y above or below what would be predicted from the associated value of X. Residual $Y = Y - \hat{Y}$, where $\hat{Y} = a + bX$ (the regression of Y on X).
13. While it is argued above that there is a basic *conceptual* difference between a ratio and a residual, there is also an empirical difference. The two are equivalent only under limiting conditions (Bollen and Ward; Fuguitt and Lieberson; Kuh and Meyer). The following pairs of coefficients contrast the correlation of crime rate (C/P) with certainty of punishment, measured first as a ratio (A/C) then as a residual ($A.C$): $-.44$ vs $-.09$; $-.20$ vs $.06$; $-.55$ vs $-.21$; $-.56$ vs $-.13$; $-.24$ vs $.14$; $-.45$ vs $-.07$; $-.38$ vs $-.13$; $-.29$ vs $-.15$.
14. For the bivariate case, calculating the part correlations by formula (5) produces exactly the same results as first calculating the residual values for ($A/C \cdot 1/C$) and then correlating these residual scores with C/P. Therefore, the use of "residual certainty" ($A/C \cdot 1/C$) as the measure of the independent variable ought not to cause any problems in any multivariate analyses.
15. Schuessler's discussions, in particular, demonstrate that it can often be quite illuminating to decompose a ratio correlation into the covariances among the components and the variance of the common term. This would be particularly helpful where a ratio correlation that did not make any substantive sense could be expressed in terms of covariances among the components that did make sense.

References

American Friends' Service Committee. 1971. *Struggle for Justice*. New York: Hill & Wang.

Antunes, G., and A. L. Hunt. 1973. "The Impact of Certainty and Severity of Punishment on Levels of Crime in America States: An Extended Analysis." *Journal of Criminal Law, Criminology and Police Science* 64:486–93.

Bailey, W. C., D. J. Martin, and L. N. Gray. 1971. "Crime and Deterrence: A Correlation Analysis." Mimeograph.

Bean, F., and R. Cushing. 1971. "Criminal Homicide, Punishment and Deterrence: Methodological and Substantive Reconsiderations." *Social Science Quarterly* 52:277–89.

Block, Michael K. 1972. *An Economic Analysis of Theft With Special Emphasis on Household Decisions Under Uncertainty*. Unpublished Ph.D. dissertation, Stanford University.

Bollen, K., and S. Ward. 1980. "Ratio Variables in Aggregate Data Analysis; Their Uses, Problems, and Alternatives." In Edgar F. Borgatta and David J. Jackson (eds.), *Aggregate Data; Analysis and Interpretation*. Beverly Hills: Sage.

Chapman, Jeffrey, I. 1973. "The Impact of Police on Crime and Crime on Police: A Synthesis of the Economic and Ecological Approaches." UCLA, Institute of Government and Public Affairs. Mimeograph.

Chayes, Felix. 1971. *Ratio Correlation: A Manual for Students of Petrology and Geochemistry*. Chicago: University of Chicago Press.

Chiricos, T. G., and G. P. Waldo. 1970. "Punishment and Crime: An Examination of Some Empirical Evidence." *Social Problems* 18:200–17.

Dubois, Philip H. 1957. *Multivariate Correlational Analysis*. New York: Harper & Bros.

Ehrlich, I. a:1972. "The Deterrent Effect of Criminal Law Enforcement." *Journal of Legal Studies* 1:259–76.

———. b:1972. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy* 81:521–65.

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

Erickson, M., and J. Gibbs. 1973. "The Deterrence Question: Some Alternative Methods of Analysis." *Social Science Quarterly* 54:534–51.

Fleiss, J. L., and J. M. Tanur. 1971. "A Note on the Partial Correlation Coefficient." *The American Statistition* 25:42–45.

Freeman, J. H., and J. E. Kronenfeld. 1973. "Problems of Definitional Dependency: The Case of Administrative Intensity." *Social Forces* 52:108–21.

Fuguitt, G. V., and S. Lieberson. 1974. "Correlation of Ratios or Difference Scores Having Common Terms." In Herbert L. Costner (ed.), *Sociological Methodology 1973–1974*. San Francisco: Jossey-Bass.

Gibbs, J. P. 1968. "Crime, Punishment and Deterrence." *Southwestern Social Science Quarterly* 48:515–30.

Gray, L. and D. J. Martin. 1969. "Punishment and Deterrence: Another Analysis of Gibbs' Data." *Social Science Quarterly* 50:389–95.

Greenberg, D. F. 1977. "Deterrence Research and Social Policy." In Stuart Nagel (ed.), *Modeling the Criminal Justice System*. Beverly Hills: Sage.

Kasarda, J. D., and P. D. Nolan. 1979. "Ratio Measurement and Theoretical Inference in Social Research." *Social Forces* 58:212–27.

Kobrin, Solomon, Stephen G. Lubeck, E. Wayne Hansen, and Robert Yeaman. 1973. *The Deterrent Effectiveness of Criminal Justice Sanction Strategies: Summary Report*. Washington: Law Enforcement Assistance Administration, National Institute of Law Enforcement and Criminal Justice.

Kuh, E., and J. R. Meyer. 1955. "Correlation and Regression Estimates When Data are Ratios." *Econometrica* 23:400–16.

Logan, Charles H. a:1971. "On Punishment and Crime (Chiricos and Waldo 1970): Some Methodological Commentary." *Social Problems* 19:280–84.

———. b:1972. "General Deterrent Effects of Imprisonment." *Social Forces* 51:64–73.

———. c:1975. "Arrest Rates and Deterrence." *Social Science Quarterly* 56:376–89.

Long, S. B. 1979. "The Continuing Debate Over the Use of Ratio Variables: Fact and Fiction." In Karl F. Schuessler (ed.), *Sociological Methodology, 1980*. San Francisco: Jossey-Bass.

MacMillan, A., and R. L. Daft. a:1979. "Administrative Intensity and Ratio Variables: The Case Against Definitional Dependency." *Social Forces* 58:228–48.

———. b:1980. "Relationships Among Ratio Variables with Common Components: Fact or Artifact." *Social Forces* 58:1109–28.

McNemar, Quinn. 1969. *Psychological Statistics*. 4th ed. New York: Wiley.

Morris, Norval. 1974. *The Future of Imprisonment*. Chicago: University of Chicago Press.

Nagin, D. 1978. "General Deterrence: A Review of the Empirical Evidence." In Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin (eds.), *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*. Washington: National Academy of Sciences.

Pearson, Karl. 1897. "Mathematical Contributions to the Theory of Evolution: On a

Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs." *Proceedings of the Royal Society of London* 60:489–98.
Phillips, L., and H. Votey, Jr. 1972. "An Economic Analysis of the Deterrent Effect of Law Enforcement on Criminal Activities." *Journal of Criminal Law, Criminology, and Police Science* 63:335–42.
Rangarajan, C., and S. Chatterjee. 1969. "A Note on Comparison Between Correlation Coefficients of Original and Transformed Variables." *The American Statistician* 23:28–29.
Schuessler, K. a:1973. "Ratio Variables and Path Models. In Arthur S. Goldberger and Otis Dudley Duncan (eds.), *Structural Equation Models in the Social Sciences*. New York: Seminar Press.
———. b:1974. "Analysis of Ratio Variables: Opportunities and Pitfalls." *American Journal of Sociology* 80:379–96.
Tittle, C. R. 1969. "Crime Rates and Legal Sanctions." *Social Problems* 16:408–28.
Tittle, C. R., and A. R. Rowe. 1974. "Certainty of Arrest and Crime Rates: A Further Test of the Deterrence Hypothesis." *Social Forces* 52:455–62.
Tullock, G. 1974. "Does Punishment Deter Crime?" *The Public Interest* 36:103–11.
Vandaele, Walter. 1976. "Participation in Illegitimate Activities: I. Ehrlich Revisited." Harvard University, unpublished manuscript.
Wilson, James. 1975. *Thinking About Crime*. New York: Basic Books.
Yule, G. U. 1910. "On the Interpretation of Correlations Between Indices or Ratios." *Journal of the Royal Statistical Society* 73:644–47.
Yule, G. Udny, and M. G. Kendall. 1950. *An Introductic : To The Theory Of Statistics*. 14th ed. New York: Hafner.

END