SIMULTANEOUS SCALING OF OFFENSE SERIOUSNESS AND SENTENCE SEVERITY THROUGH CANONICAL CORRELATION ANALYSIS

JAMES C. McDAVID*
BRIAN STIPAK

This article describes and illustrates a new, easily applied method of scaling the severity of different types of criminal sentences and the seriousness of different types of crimes. In contrast to past approaches, this method is based on actual judicial performance, not on opinions or subjective scaling procedures. Using this approach, the effects on sentencing of variables other than type of crime can be examined.

The severity of different types of criminal sentences and the seriousness of different types of crimes are two key variables for research. Past studies have relied on the subjective assessments of the researchers, judges, college students, or some other group in determining scale values for sentences and crimes. In contrast to these studies, the empirical approach presented in this paper uses data from the actual performance of the judicial system in order to assess the severity of sentences and the seriousness of crimes. Possible biases in judicial performance resulting from different treatment accorded various socio-demographic groups, differences among judges, and other factors can be taken into account and examined using this approach. Because this approach is objective, in the sense that it relies on actual results of the judicial system rather than on the opinions of a

* The authors gratefully acknowledge the assistance they received from others. Carl Hensler offered useful suggestions for analysis, and Herbert Kritzer provided helpful comments on a draft of the manuscript. Glen Kreider developed the program for splitting the data into two random subsamples. Ralph Perez and Bob Stevenson contributed to making the data base useful for analysis. Finally, Jeffrey Bastuscheck of the Central Regional Advisory Committee, Pennsylvania Commission on Crime and Delinquency, provided the data tape, as well as advice on its use.

particular group of people, the authors believe it offers a valid method of scaling sentences and crimes.[1]

I. ALTERNATIVE APPROACHES TO SCALING THE SEVERITY OF SENTENCES AND THE SERIOUSNESS OF CRIMES

The simplest approach to scaling sentences or crimes is to subjectively assign numerical scale values. Tiffany *et al.* (1975: 370-371), for example, assigned estimated interval-level scale values of 0 - 50 to 16 sentence categories ranging from suspended sentence to over ten years imprisonment.[2] A slightly different alternative is the graphic rating scale method (Guilford, 1954: 265), which requires each subject to mark on a line representing a severity or seriousness scale the relative position of each sentence or crime.

More sophisticated psychometric scaling techniques can also be applied to subjective responses obtained from judges, lawyers, the researchers themselves, or some other group. For example, Buchner (1979) used a sample of criminal court judges to assess relative severity for a number of paired comparisons of different types of sentences, and then used Thurstone scaling methods to estimate interval-level severity scores. Buchner then constructed a multiple regression model, which regresses the sentence severity score on different attributes of the sentences. The model can then be used to calculate severity scores for other sentences that were not included in the original analysis.

Psychometric scaling methods have also been applied to the scaling of crimes, especially since the seminal work of Sellin and Wolfgang (1964). Using subjective data obtained from undergraduate college students, Sellin and Wolfgang developed a widely used scale of crime seriousness. They concluded that the subjective seriousness of a crime was determined by several objective characteristics of the crime, such as the amount of money stolen and the extent of the victim's physical injuries. Subsequent research has replicated this scaling analysis with a variety of different groups of

---

1 This approach is appropriate when researchers desire severity and seriousness measures that are implicit in a judicial system's performance—for example, in order to statistically control for crime seriousness when estimating the effects of other factors on sentence severity. However, these scale values do not necessarily have normative significance and may not, for example, be appropriate for establishing sentencing guidelines.

2 This scale is a slight modification of a scale first developed by the Administrative Office of the U.S. Courts (1965: 29).

subjects (e.g. Normandeau, 1966; Ackman *et al.*, 1967; McConnell and Martin, 1969; Figlio, 1975), including subjects from different cultures, and has obtained generally similar results.

Regardless of the sophistication of the analysis, scaling methods that use subjective data from judges, lawyers, the researchers themselves, or any other group have several disadvantages. First, subjective data reflect the opinions of a specific group of people, which may differ from the opinions of other groups; for example, McCleary *et al.* (1981) found that judges and other criminal justice professionals disagreed with the general public about the relative seriousness of some types of crimes. Second, such data are not normally available, and require a special data collection effort. Finally, directly asking judges or other subjects to assign scale values or to make paired comparisons exemplifies the "obtrusive" techniques of social science data collection that have been criticized as artificial and likely to generate invalid data (e.g., Webb *et al.*, 1966). The subjective responses provided by subjects are not necessarily related to how the judicial system operates, or even to the subjects' own behavior in normal situations, such as the decisions judges make in the actual adjudication of cases.

A completely different approach is to scale sentences and crimes according to how the judicial system assigns sentences to crimes. In the following sections it will be shown how the statistical technique of canonical correlation can be used to implement such an approach. The approach avoids the aforementioned problems. First, the results are not subjective but rather are based on the performance of the judicial system itself. Second, the method requires no data beyond that kept in judicial records; special data collection efforts are avoided, and comparisons can be easily made of results computed for different court systems.

II. APPLICATION OF CANONICAL CORRELATION TO SCALING SENTENCES AND CRIMES

Our basic approach is to scale sentences according to how the judicial system assigns them to crimes of varying seriousness, and to scale crimes according to how the judicial system assigns to them sentences of varying severity. Each of these tasks would be relatively easy, given a pre-existing interval scale for the seriousness of crimes or a pre-existing interval scale for the severity of sentences, respectively.[3] Since the current purpose is to scale both sentences and crimes according to actual judicial performance, it is not appropriate to rely on subjectively derived values for either scale. Fortunately, the statistical technique of canonical correlation provides a neat solution to this problem by allowing researchers to simultaneously estimate scale values for the seriousness of different crimes and the severity of different sentences.[4]

Canonical correlation is not one of the common statistical techniques used in social research, but its purpose can be easily understood by anyone familiar with the more common multivariate methods of factor analysis and regression analysis. Canonical correlation extracts a linear combination, called a "canonical variate," from each of two sets of variables. Thus, a canonical correlation analysis is similar to a factor analysis on each of the two sets of variables. However, whereas factor analysis extracts factors that maximally explain the common variance among the variables, canonical correlation extracts pairs of canonical variates such that the correlation, called the canonical correlation, between the variates from each set of factors is maximized. Just as factor analysis can extract more than one factor, canonical correlation can extract more than one pair of canonical variates. The first pair is that pair of linear combinations having the largest possible canonical correlation; the second pair is that pair of linear combinations, each of which is constrained to be uncorrelated (orthogonal) with the corresponding first canonical variate, having the largest possible canonical correlation, and so on. Canonical correlation analysis can also be viewed as similar to multiple regression analysis, except that there is not only more than one independent variable but also more than one dependent variable.[5]

The use of canonical correlation as a scaling technique was introduced by Klatzky and Hodge (1971), who applied it to

---

3 The simplest case would merely require calculating the mean seriousness score for each sentence category or the mean severity score for each crime category. However, under some conditions it might be necessary to transform the pre-existing scores and to incorporate other variables into the statistical analysis (see Hensler and Stipak, 1979: 639-640).

4 A conceptually similar, but less efficient or rigorous approach for scaling the seriousness of crimes is to compare the average number of years' imprisonment or the relative frequency of probation for different types of crimes.

5 For more technical discussions of canonical correlation see Cooley and Lohnes (1971), Morrison (1967), and Tatsuoka (1971). See McKeon (1965) for an overview of the relationship of canonical correlation to other techniques.

scaling the occupational prestige of fathers' and sons' occupations. Hensler and Stipak (1979: 642-643) demonstrated how the technique could be used to estimate interval-scale values for survey item response categories. When canonical correlation is used as a scaling technique, dummy (binary, 1/0) variables are used to represent the categories of the variable(s) to be scaled. As in dummy variable regression analysis, n categories are represented with n-1 dummy variables, and the omitted category is considered to have a coefficient of zero. The unstandardized dummy variable coefficients (weights) calculated for the first set of canonical variates are the estimated scale values.

When applying canonical correlation to scaling the severity of discrete sentence categories or crime categories, the two sets of dummy variables for the sentence categories and the crime categories become the two sets of input variables. Which sentence category and which crime category are omitted is not important.[6] However, it is important to remember that the unstandardized, not the standardized, coefficients become the scale values.[7] Since most computer programs produce standardized coefficients, it is necessary to divide these standardized coefficients by the standard deviations of the corresponding dummy variables to obtain the scale values. The resulting scale values can be linearly transformed to yield a convenient range, perhaps from zero to one or from zero to ten.[8]

The calculated scale values can be considered interval-level scale values, or approximately interval-level, provided that the true underlying relationship between sentence severity and seriousness of crimes is linear or approximately linear, respectively. Linearity implies that equal increases on the seriousness of crime dimension result, on average, in equal

[6] Researchers may find it convenient to omit the categories assumed *a priori* to have the lowest scale value. However, this is a minor consideration, since any set of calculated scale values will be linearly related, regardless of the category omitted, and since the estimated values can be linearly transformed as desired.

[7] Since the variance of a dummy variable depends on the proportion of cases falling in that category, using the standardized coefficients would produce scale values that depend on the frequency distribution across categories.

[8] In most analyses data will be available on the actual length of the prison sentences received. In that case, rather than only using dummy variables to represent the different types of sentences (e.g. suspended, probation), the researcher can include an interval-level length of sentence variable, in addition to the dummy variables for non-prison sentence categories. After the unstandardized canonical coefficient is obtained for the length of sentence variable, a severity scale score can then be calculated for each case receiving a prison sentence.

increases on the severity of sentence dimension. If only monotonicity is assumed, then the scale values can be considered only ordinal. That is, if the analyst is only willing to assume that more serious crimes tend to receive more severe sentences, the calculated scale values have only a rank-order interpretation. The authors view this interpretation as over-cautious, since it does not exclude any possible monotonic relationship between seriousness of crime and severity of sentence. Even if linearity does not strictly apply, many nonlinear, monotonic transformations would result in only minor distortions of the relative scale values. In addition, one purpose for estimating scale values is to facilitate using interval-level statistics like regression analysis.[9] Since the variance a variable shares with a nonlinear but monotonic transformation of itself is extremely high for a variety of monotonic transformations (Labovitz, 1970: 519), minor nonlinear transformations of the true interval scale values will usually have little effect on further statistical analyses. Moreover, since the calculated scale values maximize the linear relationship between the severity and seriousness scales, transformation errors that could interfere with the ability to statistically control for crime seriousness when estimating the effects of other factors on sentence severity are minimized.

Analysts should be alert to indications that the assumptions underlying the analysis are wrong. Since this approach assumes that the main reason types of crimes are related to types of sentences is the correspondence between seriousness of crimes and severity of sentences, the strength of association between the second pair of canonical variates should be much weaker than the first pair. Therefore, the analyst should consider it a warning if the first canonical correlation is not substantially larger than the second canonical correlation.[10] In practice, we expect that this will not be a

---

[9] Some past research using subjectively assigned scale values has curiously contradicted itself in the choice of scale values and in the choice of statistical methods. Baab and Furgeson (1967), for example, rotely assigned rank-order numbers to 12 sentence categories of increasing severity—despite their belief (Baab and Furgeson, 1967: 483) that some differences between adjacent categories were far larger than between other adjacent categories—and then used that scale as an interval-level variable in multiple regression analysis. Similarly, Uhlman (1979: 111-112) assigned rank-order numbers as scale values to 93 types of sentences in order to obtain a severity scale for interval-level statistical analysis. Rather than rotely assign rank-order numbers, research that treats a severity measure as an interval-level variable should whenever possible assign scale values that more appropriately represent the relative intercategory distances.

[10] As a rough rule of thumb, the authors suggest that analysts consider it a cautionary warning if the second pair of canonical variates share more than a quarter as much variance as the first pair, and a serious warning if they share

problem for analyses that include a broad range of crimes of varying seriousness and sentences of varying severity. The broader the range, the more reasonable is the assumption that the main source of correspondence between crimes and sentences results from the seriousness-severity relationship. However, within a narrow range of crimes and sentences it would not surprise us if some other factors were more important in determining the correspondence between crimes and sentences. Therefore, researchers using this technique should include a fairly wide range of crimes and sentences within the analysis.

A second warning signal occurs when the results do not appear reasonable, especially when the ordering of crime and sentence categories is grossly contrary to *a priori* expectations. This warning may mean that an omitted variable affecting sentence severity needs to be brought into the analysis; we discuss this in a later section. An unreasonable ordering can also result from an inadequate sample size, resulting in large sampling variances and unstable coefficient estimates. Of course, an "unreasonable" ordering could simply indicate the incorrectness of the *a priori* subjective expectations.

III. THE DATA

The Pennsylvania Commission on Crime and Delinquency (PCCD) initiated a program in 1969 to develop an offender-based record of each arrested person's "track" through the Common Pleas Court system. The state, which was divided into planning regions at that time, managed the collection of these data for the counties in each region. Local county court personnel were responsible for recording the information on coding forms, and it was then keypunched and stored on computer tapes, presumably for purposes of management and analysis.

The data for this analysis are from the PCCD records for 1977 for the sixteen counties in the south central region. The data set includes information on offender characteristics, offense characteristics, case characteristics (for example,

---

more than half as much. Note that this occurs when the ratio of the canonical correlations is less than 2 and 1.4 respectively. In that case, some types of differences between crimes and between sentences, other than the seriousness of crimes and the severity of sentences, are important in determining the correspondence between crimes and sentences. Therefore, in order to use the calculated scale values, the researcher must feel confident that the first pair of canonical variates do in fact represent severity of sentence and seriousness of crime.

public versus private defender), and court characteristics. The quality of the data leaves something to be desired; in fact, the Commission is currently sponsoring a statewide project to check on the reliability and validity of the information and to upgrade coding procedures. There is a total of 2,321 cases, but because of missing data the number of cases used in one of the analyses is less.[11]

In this analysis, type of crime is represented using seven categories, based on the Pennsylvania criminal code, from summary offenses to felony one offenses.[12] In addition to dummy variables for fines/costs, suspended sentences, and probation, prison sentences of varying lengths are represented using dummy variables for sentences of less than 6 months, 6 to 12 months, 12 to 24 months, and greater than 24 months.[13]

When ungrouped data are available for the actual length of the prison sentence for each case, the researcher can include an interval-level length of sentence variable, rather than dummy variables corresponding to different length sentences. Using an interval length of sentence variable avoids measurement error created through categorization, but does require choosing an appropriate functional form for representing the prison length variable.[14] The dummy variable approach is used in this analysis in order to avoid the functional form problem,[15] and to check whether the calculated scale values appear reasonable, i.e., whether severity increases with sentence length.

---

[11] Except for Table 3, all analyses are based only on those cases having complete data on all variables being analyzed. An alternative approach, often termed "pairwise" deletion in contrast to "listwise" deletion, is to exclude a case from the calculation of a correlation coefficient only if it has a missing value on one of the two variables. This approach was used for the analysis presented in Table 3, since less than half of the cases have complete data on all of the variables used in that analysis. Although pairwise deletion results in less sample attrition, it should be avoided unless necessary, since it can lead to misleading results and computational problems.

[12] Although finer crime categories would be desirable and should be used when possible by future researchers, finer categorization for this analysis was not possible because of the number of cases and the information available in the data set.

[13] The lengths used for the prison sentences are the minimum times to be served by convicted defendants. Thus, it was possible to create a set of sentence categories with a clean break between probation and prison.

[14] Representing prison length with a linear term would imply that the marginal increase in sentence severity for each additional time increment remains constant. A more reasonable assumption probably is that sentence length has a decreasing marginal effect on severity, and therefore a nonlinear transformation of sentence length is appropriate.

[15] Note that examining the coefficient estimates for the prison dummies can aid in selecting an appropriate transformation for an interval prison length variable.

Table 1 presents, for each sentence and crime category, the frequency, the unstandardized canonical coefficient based on the first pair of canonical variates,[16] and a scale value transformed to a zero to ten range. The results are consistent with prior expectations of the rank order of the sentence and crime categories. The lowest estimated scale value for sentences is for fines, and the highest is for the longest prison sentences. The lowest estimated scale value for crimes is for summary offenses, and the highest is for felony one offenses. The reasonableness of these results supports the use of canonical correlation for this type of scaling problem.

Table 1. Canonical Correlation Results for Predicting Severity of Sentence from Seriousness of Crime

Variable	Category	Frequency	Unstandardized Canonical Coefficient	Scale Value
Type of Sentence	Fines, Costs	450	*	0.00
	Suspended	115	.939	2.13
	Probation	770	1.076	2.44
	<6 mo. Prison	711	1.671	3.79
	6-12 mo. Prison	158	3.041	6.91
	12-24 mo. Prison	83	3.996	9.07
	>24 mo. Prison	34	4.404	10.00
Type of Crime	Summary	93	*	0.00
	Misdemeanor 3	179	.252	0.66
	Misdemeanor 2	960	1.469	3.82
	Misdemeanor 1	251	1.995	5.08
	Felony 3	335	2.208	5.74
	Felony 2	225	2.923	7.60
	Felony 1	278	3.845	10.00

Canonical correlation = .49
Total number of cases = 2321

*No dummy variable was included, since this is the reference category. Coefficient is considered to be zero for scaling purposes.

In order to check the stability of these scale values,[17] the total sample was randomly divided into two subsamples and the canonical correlation analysis was done for each separately. As Table 2 shows, some minor differences resulted in the relative scale values. Although the rank order for Random Sample One remains the same as in Table 1, for Random Sample Two the order of suspended sentences and probation, and of the last two prison categories are reversed. Thus, for

[16] The second pair of canonical variates was much less strongly related, since its canonical correlation was only .21. This is consistent with the assumptions underlying the scaling technique.

[17] Unfortunately, convenient statistical tests and methods for constructing confidence intervals are not available for the canonical coefficients, although tests can be made of the statistical significance of the canonical correlations.

Table 2. Canonical Correlation Results for Randomly Split Half-Samples

		Random Sample One			Random Sample Two		
Variable	Category	Frequency	Unstandardized Canonical Coefficient	Scale Value	Frequency	Unstandardized Canonical Coefficient	Scale Value
Type of Sentence	Fines, Costs	219	*	0.00	231	*	0.00
	Suspended	57	1.483	3.50	58	1.746	4.22
	Probation	386	1.566	3.70	384	1.561	3.77
	<6 mo. Prison	353	2.100	4.96	358	2.025	4.89
	6-12 mo. Prison	73	3.218	7.59	85	2.951	7.13
	12-24 mo. Prison	45	3.703	8.74	38	4.139	10.00
	>24 mo. Prison	16	4.237	10.00	18	3.960	9.57
Type of Crime	Summary	52	*	0.00	41	*	0.00
	Misdemeanor 3	91	.522	1.30	88	.434	1.04
	Misdemeanor 2	474	2.185	5.43	486	2.173	5.20
	Misdemeanor 1	118	2.692	6.69	133	2.354	5.64
	Felony 3	174	2.856	7.10	161	2.844	6.81
	Felony 2	102	3.378	8.40	123	3.056	7.32
	Felony 1	138	4.022	10.00	140	4.175	10.00

Random Sample One		Random Sample Two	
Canonical correlation =	.52	Canonical correlation =	.49
Total number of cases =	1149	Total number of cases =	1172

*No dummy variable was included, since this is the reference category. Coefficient is considered to be zero for scaling purposes.

moderate sample sizes, the scale values appear fairly stable, but larger size samples with more cases in each category would be desirable to yet further decrease sampling error, yielding more stable scale values.

Controlling for Other Variables When Estimating Scale Values

Sometimes variables in addition to type of sentence and type of crime must be included in the canonical correlation scaling analysis. To understand why, recall what happens in multiple regression analysis if a relevant independent variable (an independent variable that has an effect on the dependent variable) is not included. Such omission, termed a "specification error," biases the coefficient estimates for other independent variables that are correlated with the omitted variable.[18] Similarly, in this application of canonical correlation analysis, failing to include relevant predictors of sentence severity in addition to the type of crime may distort the calculated scale values.[19] By including such variables in the canonical correlation analysis, possible distortions are minimized and at the same time the effect of those variables on sentence severity can be examined.

Although researchers can include in the analysis any additional variable that may have an effect on sentencing which the researchers want to investigate, for purposes of estimating scale values only those variables need to be included that not only may have an effect on sentencing, but also are statistically associated with the type of offense.[20] Omitting predictors of sentence severity that are statistically independent of the crime categories cannot distort the calculated scale values.[21] As an example of a variable that does need to be included

[18] For a discussion of the effects of this type of specification error see Kmenta (1971: 391-395), Wonnacott and Wonnacott (1970: 312-314), or other econometrics texts.

[19] In this analogy to multiple regression analysis, the biased canonical coefficients for the crime dummy variables are analogous to the biased regression coefficients. However, any nonlinear biases (see following footnote) will bias the coefficients for the sentence dummy variables also, since the coefficients for each canonical variate are calculated to maximize the correlation with the other canonical variate.

[20] More exactly, it is not necessary to include variables that cause only linear biases in the coefficient estimates—i.e., variables that have linear effects on sentence severity and are linearly related to crime seriousness—since our objective is to estimate scale values for a linear function of the underlying sentence severity dimension. Linear biases will not distort the calculated scale values, although they can affect the efficiency of estimation.

[21] This is analogous to the principle that omission of an independent variable from a regression equation will not affect the coefficient estimates for the other independent variables if the regressors are orthogonal (Wonnacott and Wonnacott, 1970: 259; Kmenta, 1971: 381).

in the analysis, assume that black offenders are disproportionately represented in the most serious crime categories, and that the court system being studied tends to discriminate against black offenders by giving them more severe sentences, *ceteris paribus*. In that case, it would be necessary to include a dummy variable distinguishing between black and white offenders in the canonical correlation analysis, in addition to the crime category dummies. Otherwise, the scaling results would be distorted by the discrimination against black offenders, resulting in over-serious estimated scale values for the most serious crime categories.[22]

Fortunately, it is easy to include any variables in the analysis which the researcher thinks might cause such distortion. Categorical variables, such as the race or sex of the defendant, can be represented by dummy variables. Besides characteristics of the defendant, characteristics of the case and of the court can also be included. Almost any theoretically relevant characteristics can be represented by including appropriate variables in the analysis. For example, if a particular judge is unusually severe in sentencing, the researcher can include a dummy variable for that judge; if, in addition, the researcher thinks that judge may be even more severe in sentencing defendants of a particular demographic group, that effect can be represented by a dummy variable interaction term.[23] Interval-level variables can be included directly as interval variables in the analysis. If the effect of an interval variable on sentence severity is expected to be nonlinear, the researcher can first subject the variable to an appropriate transformation.[24]

Table 3 shows the results obtained when the scaling analysis presented in Table 1 was redone, including additional predictors of sentence severity other than type of crime. The sex of the defendant is represented by a dummy variable which

---

22 Discrimination under these assumptions causes the most serious crime categories to correspond on average to more severe sentences than they would without discrimination, biasing their estimated scale values upward. Conversely, discrimination implies that the most severe sentence categories correspond on average to less serious crimes than they would without discrimination, biasing their estimated scale values downward.

23 The researcher would compute the dummy variable interaction term by multiplying the dummy variable that has a value one for that judge by the dummy variable that has a value one for defendants of that demographic group.

24 A logarithmic transformation is often an appropriate transformation for a variable expected to have a decreasing marginal effect—in particular, when roughly equal changes in the dependent variable (sentence severity) are expected to result from equal proportional changes in the independent variable.

has the value one for males, and zero for females. Similarly, other predictors represented by single dummy variables include the defendant's race (white versus nonwhite), whether the defendant was released on bail, whether there was a presentence investigation, the type of counsel (public versus private defender), the type of court administrator (traditional versus professional), and the type of conviction (guilty plea

Table 3. Canonical Correlation Results for Predicting Severity of Sentence from Seriousness of Crime, Controlling for Other Variables

Variable	Category	Frequency	Unstandardized Canonical Coefficient	Scale Value
Type of Sentence	Fines, Costs	450	*	0.00
	Suspended	115	.489	1.14
	Probation	770	.595	1.39
	<6 mo. Prison	711	1.403	3.28
	6-12 mo. Prison	158	2.652	6.20
	12-24 mo. Prison	83	4.117	9.63
	>24 mo. Prison	34	4.276	10.00
Type of Crime	Summary	93	*	0.00
	Misdemeanor 3	179	.026	0.13
	Misdemeanor 2	960	.711	3.52
	Misdemeanor 1	251	.928	4.59
	Felony 3	335	.967	4.79
	Felony 2	225	1.054	5.22
	Felony 1	278	2.019	10.00
Sex of Defendant	Male	1983	.348	—
Race of Defendant	White	1838	−.065	—
Type of Conviction	Guilty Plea	2068	−.154	—
Released on Bail	Yes	1717	−2.150	—
Presentence Investigation	Yes	750	.435	—
Type of Counsel	Private	693	−.315	—
Type of Court Admin.	Professional	765	−.346	—
Number of Offenses	—	—	.180	—
Time to Adjudicate	—	—	.000	—
Yearly Court Caseload	—	—	.001	—
Age of Defendant	—	—	.006	—

Canonical correlation = .60
Total number of cases = **

*No dummy variable was included, since this is the reference category. Coefficient is considered to be zero for scaling purposes.

**Pairwise deletion of missing data was used (see note 10). The minimum number of valid cases for calculating a correlation was 1831.

versus guilty by trial). Interval-level variables are included for the number of offenses with which the defendant was charged, the total time (in days) to adjudicate the case, the yearly caseload of the court, and the age of the defendant.

The scaling results shown in Table 3 are very similar to those obtained (Table 1) when no additional variables were included in the analysis. The only noteworthy differences are that in Table 3 suspended sentences and probation are scaled somewhat less severely. The additional variables included in Table 3 did not introduce any large distortions in the original set of calculated scale values. However, in some cases it could be possible, as discussed above, for other variables that have strong effects on sentencing and are strongly related to the type of crime, to cause large distortions of scale values calculated without considering those variables.

The additional variables (Table 3) were included primarily to illustrate how additional variables can be controlled when estimating scale values. However, these results can also be used to examine the effects of additional variables on sentencing severity.[25] The sign for each canonical coefficient indicates whether that variable predicts to more severe or more lenient sentences, *ceteris paribus*. For example, these results suggest that, for the same type of crime, males, nonwhites, and older defendants tend to receive more severe sentences. The coefficients also indicate the relative magnitude of the estimated effects. For example, the greater severity predicted for male versus female defendants (.348) is about the same as the estimated difference for public versus private counsel (.315) or for an additional charge of two offenses (.360).

As has just been illustrated, researchers interested in examining the effects of variables other than type of crime on sentence severity can do so by including those variables in the canonical correlation analysis. Variables can be included in this way for purposes of examining their effect on sentencing, regardless of whether their inclusion is necessary for purposes of estimating the scale values. However, the researcher can choose an alternative, two-step strategy of analysis. First, the

[25] As was already discussed, for purposes of calculating scale values, variables affecting sentence severity need not be included if they are statistically independent of the crime categories. However, for purposes of estimating the effect of other variables on sentence severity, researchers must attempt to include in the analysis all variables that 1) have an effect on sentencing, and 2) are correlated with any of the variables whose effects on sentence severity are being estimated. Omission of variables that affect sentencing will bias the estimated effects for correlated variables that are included in the analysis.

researcher can do the canonical correlation scaling analysis, yielding scale values for sentence severity and crime seriousness. Second, the researcher can do a multiple regression analysis that regresses sentence severity on crime seriousness and the other independent variables. This analysis strategy offers the advantage of separating the scaling analysis from the analysis of the impact of other variables, which the researcher may think of conceptually as distinct steps, as well as the advantage of allowing the use of the more familiar technique of multiple regression. However, researchers should use this strategy cautiously, being careful to include necessary variables in the canonical correlation analysis, and paying special attention to any variables that are strongly related to type of crime.[26]

IV. CONCLUSION

The scaling approach presented in this article is easily applied and offers the convenience of requiring no data beyond those usually available in offender-based court records. It also offers the validity of severity and seriousness measures scaled according to how actual judicial performance assigns sentences to crimes. Thus, it is an appropriate empirical method for research on a court system when the researchers desire severity and seriousness scales that are implicit in the system's performance, rather than scales based on consensual validity among "experts" or based on some *a priori* standards.

Scaling results from studies using this approach should be published and compared in the research literature. The scaling results themselves provide substantively interesting information about how a judicial system operates. In fact, perhaps the most interesting result would occur if the scaling analysis broke down,[27] indicating that sentencing was not based primarily on assigning appropriately severe sentences according to the seriousness of the crime, and alerting the researchers to other factors having greater importance. However, the authors do not expect this to occur when researchers apply this analytical approach to data from U.S. court systems, provided that there are a large number of cases

---

[26] Also, researchers cannot use standard statistical tests and confidence intervals when interpreting the results from the multiple regression analysis, because the prior scaling analysis makes the observations on the dependent variable not independent and makes error in the dependent variable correlate with error in the crime seriousness variable.

[27] That is, in some applications the first pair of canonical variates may not correspond to seriousness and severity.

representing a wide range of seriousness. Comparisons of results should be made to identify differences across states and court systems, to help understand the errors that inevitably occur in statistical social research, and to promote model building and theoretical development.

REFERENCES

ACKMAN, D.D., A. NORMANDEAU, and S. TURNER (1967) "The Measurement of Delinquency in Canada," 58 *Journal of Criminal Law, Criminology and Police Science* 330.

ADMINISTRATIVE OFFICE OF THE U.S. COURTS (1965) *Federal Offenders in the United States District Courts 1964*. Washington: Administrative Office of the U.S. Courts.

BAAB, George William and William Royal FURGESON, Jr. (1967) "Texas Sentencing Practices," 45 *Texas Law Review* 471.

BUCHNER, Deborah (1979) "Scale of Sentence Severity," 70 *Journal of Criminal Law and Criminology* 182.

COOLEY, William W. and Paul R. LOHNES (1971) *Multivariate Data Analysis*. New York: Wiley.

FIGLIO, Robert M. (1975) "The Seriousness of Offenses: An Evaluation by Offenders and Nonoffenders," 66 *Journal of Criminal Law and Criminology* 189.

GUILFORD, J.P. (1954) *Psychometric Methods*. New York: McGraw-Hill.

HENSLER, Carl and Brian STIPAK (1979) "Estimating Interval Scale Values for Survey Item Response Categories," 23 *American Journal of Political Science* 627.

KLATZKY, Sheila R. and Robert W. HODGE (1971) "A Canonical Correlation Analysis of Occupational Mobility," 66 *Journal of the American Statistical Association* 16.

KMENTA, Jan (1971) *Elements of Econometrics*. New York: Macmillan.

LABOVITZ, Sanford (1970) "The Assignment of Numbers to Rank Order Categories," 35 *American Sociological Review* 515.

McCLEARY, Richard, Michael J. O'NEIL, Thomas F. EPPERLEIN, Constance JONES, and Ronald H. GRAY (1981) "Effects of Legal Education and Work Experience on Perceptions of Crime Seriousness," 28 *Social Problems* 276.

McCONNELL, J.P. and J.D. MARTIN (1969) "Judicial Attitudes and Public Morals," 55 *Journal of the American Bar Association* 1129.

McKEON, James J. (1965) "Canonical Analysis: Some Relations Between Canonical Correlation, Factor Analysis, Discriminant Function Analysis, and Scaling Theory." *Psychometric Monographs*, No. 13. Chicago: University of Chicago Press.

MORRISON, D.F. (1967) *Multivariate Statistical Methods*. New York: McGraw-Hill.

NORMANDEAU, A. (1966) "The Measurement of Delinquency in Montreal," 57 *Journal of Criminal Law, Criminology and Police Science* 172.

SELLIN, Johan T. and Marvin W. WOLFGANG (1964) *The Measurement of Delinquency*. New York: Wiley.

TATSUOKA, M.M. (1971) *Multivariate Analysis*. New York: Wiley.

TIFFANY, Lawrence P., Yakov AVICHAI, and Geoffrey W. PETERS (1975) "A Statistical Analysis of Sentencing in Federal Courts: Defendants Convicted After Trial, 1967-1968," 4 *Journal of Legal Studies* 369.

UHLMAN, Thomas M. (1979) *Racial Justice*. Lexington, Mass.: Lexington Books.

WEBB, Eugene J., Donald T. CAMPBELL, Richard D. SCHWARTZ, and Lee SECHREST (1966) *Unobtrusive Measures*. Chicago: Rand McNally.

WONNACOTT, Ronald J. and Thomas H. WONNACOTT (1970) *Econometrics*. New York: Wiley.