6-2008

The Conjunctive Analysis of Case Configurations: an Exploratory Method for Discrete Multivariate Analyses of Crime Data

Terance D. Miethe
University of Nevada, Las Vegas, miethe@unlv.nevada.edu

Timothy C. Hart
University of Nevada, Las Vegas

Wendy C. Regoeczi
Cleveland State University, w.regoeczi@csuohio.edu

Original Citation
Miethe, T. D., Hart, T. C., & Regoeczi, W. C. (2008). The Conjunctive Analysis of Case Configurations: An Exploratory Method for Discrete Multivariate Analyses of Crime Data. Journal Of Quantitative Criminology, 24(2), 227-241. doi:10.1007/s10940-008-9044-8

ORIGINAL PAPER

The Conjunctive Analysis of Case Configurations: An Exploratory Method for Discrete Multivariate Analyses of Crime Data

Terance D. Miethe · Timothy C. Hart · Wendy C. Regoeczi

Abstract Derived from comparative approaches in both qualitative and quantitative research, the current study describes a simple exploratory technique for the multivariate analysis of categorical data. This technique is referred to as the conjunctive analysis of case configurations. After describing the logic and underlying assumptions of this conjunctive method, it is applied and illustrated in the study of the federal sentencing of drug offenders. The relative value of this conjunctive approach for purposes of exploratory data analysis and its overall utility as a method for confirmatory research are also discussed.

Keywords Conjunctive analysis · Case configurations · Discrete multivariate analysis · Exploratory methods

Introduction

Most quantitative research begins with a preliminary, exploratory analysis of the data. These initial inquiries focus on basic summary measures of univariate and bivariate distributions, often using visual representations like Tukey's (1977) specific methods of exploratory data analysis (EDA). The major value of these preliminary explorations is that they help identify particular problems (e.g., skewed distributions, outliers, non-linearity) that may affect descriptive summaries of the observed results and subsequent analyses of the data.

When applied to multivariate analysis, exploratory methods are often expressed in the language of "diagnostic tools". Diagnostic tests for multicolinearity, for example, are essential before reaching informed conclusions about the net effects of any particular

T. D. Miethe (✉) · T. C. Hart
Department of Criminal Justice, University of Nevada, Las Vegas, 4505 Maryland Parkway, Las Vegas, NV 89154-5009, USA
e-mail: miethe@unlv.nevada.edu

W. C. Regoeczi
Department of Sociology, Cleveland State University, 2121 Euclid Avenue, RT1733, Cleveland, OH 44115-2214, USA

variable. Preliminary assessments for autocorrelation serve a similar role in multivariate analysis across spatial units and time-series modeling. The investigation of the marginal distributions and minimum expected cell frequencies are the primary diagnostic methods used in multivariate contingency table analysis.

Drawing upon existing methods for discrete multivariate analysis, the current study describes an alternative technique for exploring causal relationships among categorical variables. We refer to this technique as the conjunctive analysis of case configurations. After describing the logic of this approach and its underlying assumptions, it is illustrated in the study of the federal sentencing of drug offenders. The relative value of this conjunctive approach for EDA and its overall utility as a primary method for confirmatory research within criminology are also discussed.

Comparative Methods for Categorical Data

There are various methods of cross-case comparative analysis of categorical data. Bivariate contingency table analysis, for example, is the most basic method for exploring the joint distribution of two categorical variables. Techniques of discrete multivariate analysis (e.g., elaboration models, log-linear analysis, logit models, configural frequency analysis) extend this approach to multiple categorical variables and the analysis of the main and interaction effects among them. Many of these analytic procedures begin with a saturated model of all possible effects among variables and then proceeds through a series of tests of nested model to derive a more parsimonious representation of these relationships (see Bishop et al. 1975; Goodman 1972; von Eye 2002).

The method of qualitative comparative analysis (QCA) represents an alternative orientation to categorical data analysis. As developed by Ragin (1987), QCA is designed to bridge the gap between case-oriented qualitative research and variable-oriented quantitative studies. Similar to most qualitative research, QCA views cases as complex configurations of elements and assumes causal complexity (i.e., there are multiple causes of the same outcome and that any particular variable may or may not be causally related to an outcome depending upon context and the nature of the other elements of the case). However, QCA also shares with conventional quantitative research the pursuit of generality and broader patterns across cases and contexts (e.g., the possibility that a variable's impact is consistent across context, as assumed in a main-effects statistical model). The variable-oriented emphasis on assessing the relative importance of different variables is achieved primarily in QCA through the application of algorithms for minimizing casual complexity among all possible combinations of case attributes (see Ragin 2000; Drass and Ragin 1992).

Previous criminological applications of QCA have focused on its role as an alternative method to traditional quantitative approaches for cross-case comparative research. For example, Ragin (1987) uses QCA to explore the nature of empirical typologies of juvenile courts. Miethe and colleagues (Miethe and Drass 1999; Miethe and Regoeczi 2004) employ QCA to identify the common and unique features of different types of US homicides and changes in their situational contexts over time. QCA has also been the primary method in a cross-national study of the socio-political variability in death penalty laws (Miethe et al. 2005). Outside of criminology, QCA has been applied in a wide variety of studies of the multiple causes of various social policies (see Amenta and Halfmann 2000; Amenta et al. 1992; Ragin 1987, 2000). However, these previous applications of QCA have not addressed directly its potential role for EDA and its value in augmenting more conventional techniques for discrete multivariate analysis.

The Conjunctive Analysis of Case Configurations

Although derived from different disciplinary traditions of social research, QCA shares many of its structural features with other methods for multivariate analysis of categorical data (e.g., log-linear models, configural frequency analysis, latent class analysis, multiple classification analysis [MCA] for discrete data, cluster analysis of multidimensional attribute space]). For example, these methods often view cases as configurations or combinations of attributes and aggregate individual observations to develop “counts” for each distinct case configuration. These particular aggregations of observations that share similar attributes are called the joint cell frequencies in contingency tables and they are represented by the distinct rows of a truth table in QCA. Many of these approaches are also similar because they begin with a full matrix of all possible combinations of case attributes (i.e., a saturated model of all possible interactions) and then use various procedures to minimize this complexity and find more parsimonious representations of the underlying structure of these case configurations.

Given these similarities in their structural forms, QCA and other discrete multivariate methods may be viewed as specific instances of a more general type of conjunctive analysis of case configurations. Our selection of this generic name derives from its emphasis on cases as complex configurations of elements and its preliminary assumptions about the causal importance of the joint distribution of multiple attributes that determine different outcomes.

Similar to the conceptual logic of EDA, our approach to conjunctive analysis involves visual representations of case configurations that convey important information about their nature, diversity, and distribution for subsequent analysis. The simple ways to produce a data matrix table of case configurations and the specific terminology used to describe characteristics of them are examined below.

The Data Matrix Table of Case Configurations

A conjunctive analysis of case configurations begins with an aggregated compilation of all possible combinations of attributes considered simultaneously. The number of possible case configurations depends on the number of independent variables and categories within them. For a conjunctive analysis involving 5 dichotomous independent variables, there are 32 qualitatively distinct case configurations ($2^5 = 32$). If one of these independent variables involves 3 categories, the number of complete case configurations would increase to 48 ($2^4 \times 3^1 = 48$).[1] Once the possible case configurations are identified, conjunctive analysis proceeds by aggregating each observation into their respective case configuration and exploring the relative distribution of particular categories of the outcome variable across these configurations.

To illustrate the basic structure of conjunctive analysis, let’s assume some independent variables [X_1, X_2, X_3, X_4 ..., X_j] are hypothesized to influence the relative likelihood of a particular category of an outcome variable [Y_k]. Each variable is binary coded in terms of

[1] Technically, there are no limits on the number of case configurations to be included in conjunctive analysis. Miethe and Regoeczi (2004), for example, studied the nature of homicide situations by examining a maximum number of 32,768 possible case configurations involving the conjunctive interrelationships among 15 dummy variables. Most of their major analyses, however, focused on a substantially smaller number of dominant case configurations (n = 25) that represented at least 1,000 homicides per decade. Practical problems of greater interpretative complexity and small cell sizes often limit most applications of conjunctive methods to the analysis of far less than 100 distinct case configurations.

Table 1 Data matrix of case configurations

Config #	X1	X2	X3	X4	Xj	N_Cases	Y
1	0	0	0	0	…	nc1	y1/nc1
2	0	0	0	1	…	nc2	y1/nc2
3	0	0	1	0	…	nc3	y1/nc3
4	0	0	1	1	…	nc4	y1/nc4
5	0	1	0	0	…	nc5	y1/nc5
6	0	1	0	1	…	nc6	y1/nc6
7	0	1	1	0	…	nc7	y1/nc7
8	0	1	1	1	…	nc8	y1/nc8
9	1	0	0	0	…	nc9	y1/nc9
10	1	0	0	1	…	nc10	y1/nc10
11	1	0	1	0	…	nc11	y1/nc11
12	1	0	1	1	…	nc12	y1/nc12
13	1	1	0	0	…	nc13	y1/nc13
14	1	1	0	1	…	nc14	y1/nc14
15	1	1	1	0	…	nc15	y1/nc15
16	1	1	1	1	…	nc16	y1/nc16
.	.	.	.	.	…	.	.
.	.	.	.	.	…	.	.
.	.		.	.	…	.	.
ci					Xij	nci	Y1/nci

the presence [1] or absence [0] of specific attributes. When displayed in a table of i rows and j columns, each row represents a particular case configuration. The row entries also include the number of observations in the case configuration (nc_i's) and the proportional distribution of a particular category of Y within this configuration (e.g., $Y_{1/nci}$). This general structure of the data matrix for conjunctive analysis is shown in Table 1.

Many readers will notice the similarity between this table of case configurations and those used in multivariate contingency table analysis.[2] In fact, these data matrices are virtual identical. The only basic difference is that conjunctive tables display the relative proportions of cases in only the focal category of the dependent variable (i.e., Y[1 = present]), whereas all categories of the dependent variable are shown in most contingency table analysis. However, we prefer the data matrix illustrated in Table 1 for conjunctive analysis because it provides a more parsimonious and concise representation of the nature and distribution of case configurations for both EDA and confirmatory research. The ease of rearranging the order of the variables (e.g., reordered from ABCD to BCDA) for theoretical reasons or to better highlight specific comparisons across categories is another practical benefit of the matrix display in Table 1.

The Conjunctive Approach to EDA

As developed by Tukey (1977), EDA is an approach that uses a variety of graphical techniques and numerical summaries for the fuller dissection, investigation, and

[2] Appendix 1 illustrates the computer syntax and procedures for constructing a conjunctive table of case configurations within several common software packages (e.g., SPSS, STATA, SAS).

interpretation of a data set. Through the applications of EDA, researchers are often able to uncover underlying structures in the data, extract trends among variables, detect outliers and other anomalies, and develop more parsimonious models.

Conjunctive analysis through QCA and multivariate contingency table analysis is most often applied in confirmatory research to summarize trends and test hypotheses. However, many of the basic ideas and concepts associated with our version of conjunctive analysis seem especially fruitful for purposes of EDA. In particular, how conjunctive analysis addresses preliminary analytic questions about dominant configurations, case diversity, outliers and influential observations, and specification of functional forms is described below.

A basic starting point in any exploratory analysis is the examination of central tendencies and variability in the data. For conjunctive analysis, these questions are readily answered through a simple visual inspection of the matrix of case configurations. In particular, patterns of case concentration or what is called "situational clustering" (LaFree and Birkbeck 1991) are observed by exploring the relative frequencies of observations within particular case configurations (i.e., the column marked "N_Cases" in Table 1).

High levels of situational clustering are easily recognized in the conjunctive matrix by a large number of observations within only a few case configurations and minimal frequencies in others. This pattern of extreme clustering may be substantively important, but it also suggests high multicollinearity among particular categories of variables—a statistical problem that often yields unstable estimates of the net effects of specific variables. Similarly, visual inspection of the data matrix will provide immediate evidence of low-frequency configurations that may adversely affect subsequent analysis because of their possible role as outliers and otherwise influential cases. When low-frequency case configurations are present, minimum frequency rules (e.g., delete case configurations with N's < 10) are often used with conjunctive methods to reduce their influence on substantive conclusions (see Ragin 1987; Miethe and Regoeczi 2004).

Substantive questions about the causal importance of variables and particular combinations of them are addressed in conjunctive analysis by examining the column of relative proportions (i.e., the last column in Table 1). Through simple methods of paired comparisons and rearrangements of the data matrix, conjunctive analysis offers a preliminary way of evaluating the relative importance of particular variables and the nature of the functional form of relationships among them. The specific ways that conjunctive analysis addresses these substantive questions is illustrated shortly in our study of the federal sentencing of drug offenders.

Research Questions

Methods of conjunctive analysis investigate the nature of the interrelationships among categorical variables. Two questions, however, have not been adequately addressed when these techniques are applied in criminological research. First, as a method of EDA, does the visual representation of case configurations in a conjunctive analysis provide a useful diagnostic function for subsequent variable-oriented quantitative analyses? Second, as a primary method of discrete multivariate analysis, does the conjunctive analysis of case configurations yield similar results to those provided by more traditional multivariate analysis of categorical data?

To provide some answers to these questions, the current study applies the method of conjunctive analysis to the study of the federal sentencing of drug offenders. Only a brief literature review is provided in this substantive area because the primary focus of this study is to examine the relative value of the method of conjunctive analysis for both exploratory and confirmatory research.

The Risks of Imprisonment for Federal Drug Offenders

Over the last four decades, a voluminous empirical literature has emerged on the relative importance of legal and extra-legal factors in sentencing decisions (see, for example, Blumstein et al. 1983; Hagan 1974; Johnson 2005; Ulmer 1997). The study of racial differences in sentencing practices has been a focal concern in this research. These studies have often shown that the impact of race and other factors is highly contextual, depending upon the nature of other extra-legal and legal factors (see Chiricos and Crawford 1995; Miethe and Moore 1986; Myers and Talarico 1987; Peterson and Hagan 1984; Steffensmeier et al. 1998; Zatz 1987). By focusing on context-specific effects and multiple causal factors, the study of criminal sentencing is an ideal substantive area for the conjunctive analysis of case configurations. Case configurations in this research domain involve the conjunctive distribution of both legal and extralegal factors that are expected to influence sentencing decisions.

For this illustration of conjunctive analysis, we examine the federal sentencing of 1,358 drug offenders from 1997 through 1998.[3] The specific independent variables include the type of offense (1 = drug trafficking; 0 = possession and other drug offenses) and the offender's prior record (1 = prior arrest record; 0 = no prior arrests), gender (1 = male; 0 = female) and race (1 = Black; 0 = White). The dependent variable is whether the offender received a prison sentence (1 = yes; 0 = no).[4] Under these federal sentencing guidelines, 92% of the drug offenders in this particular sample received a prison sentence.

The Exploratory Analysis of Drug Cases

As an exploratory method for studying sentencing decisions, conjunctive analysis begins with an examination of the patterns of clustering and variability among the case configurations of legal and extralegal attributes. To more easily observe these patterns of clustering and variability, case configurations in the conjunctive matrix are initially rank-ordered by the relative size of their cell frequencies. Table 2 displays this rank-ordering of case configurations for our sample of federal drug offenders by their relative frequencies (i.e., N_Cases).

While all 16 possible case configurations are empirically observed in this analysis, Table 2 reveals that their relative cell sizes vary substantially. In particular, there are two dominant case configurations among these federal drug offenders (N = 622 for Config #1 and N = 337 for Config #2). These two case configurations of "male drug traffickers with

[3] These data were collected by the Federal Sentencing Guidelines Commission and are available for secondary analysis through ICPSR at the University of Michigan.

[4] As the initial step in any quantitative inquiry, a brief inspection of the univariate frequency distributions shows that many of these variables are highly skewed. The modal categories for each variable in this sample include drug trafficking (94% of the cases), having a prior record (83%), race (Black = 59%), gender (Male = 86%) and type of sentence (Prison = 92%).

Table 2 Case configurations among drug offenders ranked by their cell frequencies

Config #	Drug traffic	Prior record	Male	Black	N_Cases	Prison sent
1	1	1	1	1	622	.99
2	1	1	1	0	337	.95
3	1	0	1	0	78	.83
4	1	0	1	1	77	.99
5	1	1	0	0	56	.84
6	1	1	0	1	52	.94
7	1	0	0	0	29	.79
8	0	1	1	0	26	.54
9	1	0	0	1	23	.74
10	0	1	1	1	22	.77
11	0	1	0	0	14	.36
12	0	0	0	0	9	.22
13	0	0	1	0	5	.00
14	0	1	0	1	4	.25
15	0	0	0	1	2	1.00
16	0	0	1	1	2	.50

a prior record" account for 71% of all observations in this sample. In contrast, all low-frequency cells (i.e., N's < 10) involve non-drug traffickers (i.e., cases of Drug Trafficking = 0), and the offender does not have a prior arrest record in most of these rarely occurring configurations.

These simple observations about the concentration of case configurations and variability in their relative frequencies have direct implications for our substantive analyses and conclusions for them. In fact, two critical points emerge from this exploratory analysis. First, the visual recognition of the uneven distribution of cell frequencies across case configurations is important because of the adverse impact of small cell frequencies and high multicollinearity on estimating net effects within multivariate analyses. The application of minimum cell frequency rules (e.g., delete all configuration with N < 10) would prohibit us from estimating a completely saturated model of all possible interactions, but such a decision to eliminate low-frequency configurations may be prudent in this example for generating more stable estimates of net effects and their standard errors. Second, the high concentration of drug traffickers with prior records in this sample places direct limits on our substantive inferences about other types of drug offenders. It is the relatively low cell frequencies (N's < 30) within these other types of configurations that hamper our inferences about them. Both of these critical observations may have escaped detection without the type of multivariate exploratory analysis provided by the visual inspection of conjunctive matrix of case configurations.

The Nature of Causal Complexity in Imprisonment Risks

Substantive questions about the causal factors in sentencing decisions are examined in conjunctive analysis by the systematic study of the variability in incarceration risks across case configurations. This variability may be assessed relative to the overall incarceration

Table 3 Case configurations among drug offenders ranked by their relative risks of imprisonment

Config #	Drug traffic	Prior record	Male	Black	N_Cases	Prison sent
1	1	0	1	1	77	.99
2	1	1	1	1	622	.99
3	1	1	1	0	337	.95
4	1	1	0	1	52	.94
5	1	1	0	0	56	.84
6	1	0	1	0	78	.83
7	1	0	0	0	29	.79
8	0	1	1	1	22	.77
9	1	0	0	1	23	.74
10	0	1	1	0	26	.54
11	0	1	0	0	14	.36

risk (92%) or by making specific paired-comparisons across sets of case configurations. As shown below, both comparative approaches are easily applied to the conjunctive matrix to determine the relative importance of different variables and the best empirical specification of their functional form.

One basic way to use conjunctive analysis to assess the main- and interaction-effects of particular variables involves the examination of the particular characteristics of drug cases that are associated with the lowest and highest risks of imprisonment. This approach involves two basic steps: (1) rank the case configurations according to the relative risks of imprisonment within them and (2) compare the relative prevalence of particular categories of each variable among the highest and lowest ranked groups of case configurations. Table 3 displays this ranking of relative risks of imprisonment among case configurations with a minimum cell frequency of 10 observations within them.

The ranking of case configuration's relative risks of imprisonment in Table 3 shows the wide variability in these risks across contexts. It also indicates the nature of the case profiles above the overall mean risks (e.g., those with prison risks above 92%), those profiles substantially below the mean (e.g., 77% and lower), and the configurations between these two groups. These three groups are identified by the brackets in Table 3. The mere fact that there is wide variability in imprisonment risks across case configurations (i.e., from a low of 36% to a high of 99%) confirms that these variables have some influence on sentencing decisions. However, a closer examination of the data matrix is required to determine which variables are most important and the nature of their joint impact on this sentencing outcome.

Comparing the nature of the case configurations above and below the mean imprisonment risks provides one basis for substantive conclusions about the relative importance of particular variables. For example, these comparisons indicate that the type of offense (trafficking vs. other drug crimes) provides the most discriminatory power because drug traffickers are included in all 4 of the highest risk profiles but they are found in only 1 of the 4 configurations that represent the lowest risks of imprisonment. The categories for the other independent variables (i.e., prior arrests, gender, and race) are more evenly dispersed between the high and low risk profiles. Having a prior record, for example, is found in 3 of the 4 highest risk configurations and in 3 of 4 of the lowest risk profiles. These results indicate that the impact of offender characteristics on sentencing decisions are highly

Table 4 Structure of conjunctive matrix for main-effects and interaction-effects of race on prison risks

Config #	Drug traffic	Prior record	Male	Black	N_Cases	Prison sent
1	1	1	1	1	622	.99
2	1	1	1	0	337	.95
3	1	1	0	1	52	.94
4	1	1	0	0	56	.84
5	1	0	1	1	77	.99
6	1	0	1	0	78	.83
7	1	0	0	1	23	.74
8	1	0	0	0	29	.79
9	0	1	1	1	22	.77
10	0	1	1	0	26	.54
11	0	1	0	0	14	.36

contextual, depending upon the particular combination of other variables included in the case configuration.

An alternative method for assessing the nature of causal complexity involves variable-based comparisons across each set of case configurations that share the same profile except the variable in question. The discovery of large differences of equal magnitude in imprisonment risks between levels of a category variable across each set of configurations would indicate a significant main-effect for that variable. However, if the magnitude of differences between these categories varies widely across case configurations, this pattern would reflect some type of context-specific interaction effect. The specific order of that interaction (i.e., 1st-, 2nd-, 3rd- or 4th-order interactions) is determined by the particular pattern of differences across configurations. The application of this paired-comparison method to explore the nature and magnitude of racial differences in imprisonment risks can be illustrated by the conjunctive matrix in Table 4.

If race has a strong main-effect, Blacks and Whites should have substantially different imprisonment risks *and* the direction and magnitude of these differences should be virtually identical across contexts (i.e., pairs of case configurations that differ only in terms of the offender's race). As arranged in Table 4, these paired comparisons involve the successively numbered configurations that are highlighted with brackets.

Contrary to a main-effects specification, a simple visual inspection of the paired-comparisons in Table 4 suggests that racial differences are primarily context-specific. For criminal cases involving male drug traffickers with prior arrests (i.e., Config #1 and #2), Blacks are only slightly more likely than Whites to receive a prison sentence (i.e., 99% vs. 95%). When the case involves males with priors who are convicted of non-trafficking offenses (Config #9 and #10), however, racial differences are large and clearly detrimental to Black defendants (77% vs. 54%). In some other contexts (see Config #7 and #8), Black defendants have a slightly lower risks of imprisonment than Whites (74% vs. 79%).

Specific patterns of statistical interaction can also be recognized by making case comparisons within the conjunctive matrix of Table 4. For example, a 3-way interaction between the offender's race, gender, and prior record is revealed by the following comparisons: (1) For drug traffickers *with* a prior record, gender differences in imprisonment risks are most pronounced among White than Black defendants (i.e., compare the differences between Config #2 and #4 with Config #1 and #3) and (2) for drug traffickers *without* a prior record, gender differences are far larger among Black than White defendants (i.e., compare the differences between Config #5 and #7 with Config #6 and #8). This pattern is indicative of a 3-way interaction because the nature of race and gender differences in the risks of imprisonment depends on whether or not the defendant has a prior record.

To assess the comparability of substantive conclusions reached by conjunctive analysis and more conventional multivariate statistical methods, we conducted a logistic regression analysis on the likelihood of receiving a prison sentence. Two models were estimated: (1) a main-effects model and (2) a "modified" saturated model of lower and higher order interactions. This latter model is not a completely saturated model because some interaction effects are not estimable (e.g., the 4-way interaction among all independent variables, the 2-way interaction between offense type and prior arrest history) due to the exclusion of low-frequency cells containing empirically rare configurations.[5] The results of this logistic regression analysis are presented in Table 5 and summarized below.

As shown in Table 5, each independent variable has a significant main-effect on the risks of imprisonment among these federal drug offenders. The relative net odds of imprisonment are about 15 times higher for drug traffickers than other types of drug offenders. While significant main effects are found for each offender characteristic, this functional form is clearly contrary to the dominant pattern of context-specific effects observed through the conjunctive analysis of case configurations.

When a modified saturated model is estimated, several significant interaction effects are found. For example, the imprisonment risks for black males are significantly higher than their counterparts and there is a significant 3-way interaction between race, gender, and prior record. The numerical value of the odds ratio (.05) for the 3-way interaction represents the differential risks of imprisonment for Black males with prior records compared to other groups after adjusting for the main effects and other lower-level interactions among the included variables. This same 3-way interaction and its specific pattern were also revealed by the visual inspection of the conjunctive matrix of case configurations in Table 4.

Discussion and Implications

Derived from comparative methods within both qualitative and quantitative research, the present study has described and applied an alternative approach for the discrete multivariate analysis of crime data. We called this approach the conjunctive analysis of case configurations to emphasize its assumptions about multiple conjunctive causes and its view of cases as representing distinct combinations of attributes.

As a general approach for multivariate analysis of categorical data, there remain several questions about the relative utility of this conjunctive method for exploratory and

[5] When a saturated model of all possible main and interaction effects was estimated on the full sample ($n = 1{,}358$), many of the estimated interaction effects were highly unstable, resulting in unusually large standard errors (e.g., $se = 17{,}974$ for the gender x trafficking interaction) and extreme odds ratios (e.g., odds ratio of 2.5 million-to-1 for this same interaction). These dubious estimates are due directly to the adverse impact of the non-random distribution of case attributes within the low-frequency cells (e.g., 18/22 of these cases involve non-drug traffickers without prior records and the majority of them also involve offenders who are white and/or female). No statistically significant interaction effects are found in this saturated model and the type of drug crime (i.e., trafficking vs. other offenses) is the only variable with a significant main-effect on imprisonment risks. This absence of any interaction effects in the saturated model is in sharp contrast to the observed patterns of interaction visually revealed in the conjunctive matrix of Table 4 and confirmed by estimating the "modified" saturated model in Table 5.

Table 5 Logistic regression models of incarceration risks

Variables	Main-effects model Odds ratios	Interaction model[a] Odds ratios
Black	3.05*	.63
Male	3.22*	.71
Prior record	2.90*	1.36
Drug trafficker	15.49*	9.40*
Black × Male		20.56*
Black × Prior record		4.23
Black × Drug traffic		1.16
Male × Prior record		2.95
Male × Drug traffic		1.83
Black × Male × Prior record		.05*
Model Chi-square	135.3*	144.3*
df	6	10
N	1,336	1,336

[a] Interaction Model is a "modified" saturated model that includes all estimable interactions
* $p < .05$

confirmatory research, its extensions to other research, and the limitations of this method. Each of these questions is addressed below.

Utility of Conjunctive Method for Exploratory and Confirmatory Studies

When conventional methods for discrete multivariate analysis are used in most research applications, they are applied primarily for purposes of confirmatory research (i.e., testing hypotheses, assessing the overall fit of a model). Exploratory inquiries within this framework often involve little more than a brief inspection of the univariate distributions or bivariate relationships. Unfortunately, if multiple and complex interrelationships exist among the categorical variables, this exploratory approach will be insufficient for identifying strong interdependencies and exceptional cases that will affect any subsequent multivariate analysis.

As illustrated in our study of the federal sentencing of drug offenders, the method of conjunctive analysis of case configurations offers a more efficient and comprehensive approach for conducting EDA. In particular, the relative advantages of this conjunctive method for EDA include (1) the succinct manner in which the conjunctive matrix identifies patterns of case clustering, diversity, and low-cell frequencies that are problematic for most multivariate statistical analyses and (2) the ability to easily generate these conjunctive data matrices from various types of statistical packages (see Appendix 1 for software applications).

For purposes of confirmatory research (e.g., testing hypotheses about net effects, evaluating functional forms, providing summary measures of goodness of fit), the conjunctive method has both advantages and disadvantages compared to other discrete multivariate procedures. The primary limitation of our version of conjunctive analysis for confirmatory research is the absence of general summary measures for quantifying the strength of interrelationships across case configurations and the overall fit of the model. However, given that formal statistical tests are used in other types of conjunctive analysis (e.g., tests of equal probability across cells, quasi-independence, or the relative fit of nested models within loglinear and CFA), it is easy to see how comparable tests could be applied in the current approach to conjunctive analysis (see Bishop et al. 1975; von Eye 2002; von Eye et al. 2006)

The primary advantage of conjunctive analysis for confirmatory research involves the visual acuity that the conjunctive data matrix provides for directly evaluating hypothesized effects. For example, if a main-effect model is suggested by a particular theory, visual inspection of the conjunctive matrix will easily reveal the validity of this specification by using the method of paired-comparisons illustrated in Table 4. If one's theory suggests complex causal relationships that are entirely contextual, this specification would be visually confirmed by a conjunctive data matrix with no main effects and largely idiosyncratic effects across groups of case configurations. However, neither main nor interaction effects would be visually confirmed when the configuration matrix exhibits a pattern of equal probability of a particular outcome variable across all case configurations. After visual recognition of the patterns underlying a conjunctive matrix, conventional methods of hypothesis testing within multivariate statistical analyses could then be performed to formally assess the statistical significance of these findings. It is this type of integration of visual and statistical methods that we recommend as a sound research strategy for the multivariate analysis of categorical data.

Another advantage of the conjunctive method for substantive analysis is its "top-down" approach to data analysis. In particular, some econometricians have recently exalted the virtues of this modeling approach over their "bottom-up" counterparts (see Charemza and Deadman 1997). Top-down modeling starts with a general model and seeks a more parsimonious representation of the data by applying theory and proceeding in a structured and ordered manner (see Charemza and Deadman 1997: 78). Bottom-up modeling, in contrast, takes the specific-to-general route for testing models and is often more haphazard in its search of competing models. Given that conjunctive analysis begins with the general assumption of causal complexity and then seeks a more simplified model, this top-down approach may offer a more defensible statistical framework for confirmatory research on criminal sentencing and other areas of criminology.

Extensions and Limitations

Conjunctive analysis in this current study was applied to a rather simple model of causal complexity involving the simultaneous distribution of four binary independent variables. However, this approach can be easily extended to other research situations involving a larger number of variables and multiple categories within them. For example, Miethe and Regoeczi (2004) apply a conjunctive approach (QCA) to describe the clustering and variability in homicide situations that are formed by the conjunction of 7 dichotomous and 4 trichotomous independent variables. Ragin (2000) has also extended conjunctive methods to continuous variables through the use of fuzzy sets. As is true of all conjunctive methods, substantive theory is crucial in these fuzzy set applications to identify the most important variables for the analysis and to help define meaningful classes of group membership within them.[6]

[6] The research group for comparative methods for the advancement of systematic cross-case analysis and small-n studies (COMPASS) provides numerous bibliographic sources and software links for conducting various types of comparative configurational analyses (e.g., QCA, fs/QCA (fuzzy set), and mvQCA (multi value). Software for conducting QCA that has been developed by Charles Ragin and associates can be downloaded from reference links in the COMPASS website (http://www.compasss.org).

When applied to more complex causal structures, conjunctive analysis requires large sample sizes and the development of rules for minimum cell frequencies. Rules for minimum cell frequencies are important in conjunctive analysis so that idiosyncratic patterns from low-frequency cells do not adversely effect the interpretation of more dominant patterns of case concentration within a study. However, before simply deleting these low-frequency configurations, it is substantively important to document their relative prevalence within a conjunctive matrix because they are indicative of the overall level of case diversity. While the requirement of large samples may be viewed as a limitation of conjunctive analysis, the same problem with small cell frequencies also plagues other methods of discrete multivariate analysis.

For most variable-oriented researchers, the conjunctive method described in this paper will be criticized on several grounds. The most likely criticisms are that the conjunctive method is more descriptive than predictive and it is less theoretically-informed than more quantitative approaches (see Ragin 1987). However, these potential criticisms seem unwarranted for several reasons. First, we contend that the descriptive value of the conjunctive method is indispensable for augmenting variable-oriented research because the conjunctive matrix provides a succinct and clear picture of (1) the proper functional form among variables and (2) the magnitude of case clustering, diversity, and low-cell frequencies that affect statistical estimates in discrete multivariate analyses. Second, conjunctive analysis uses theory to identify the major variables influencing some outcome variable and specifies a model of multiple causality and joint effects that is also derived from substantive theories (see, for application, Ragin 2000; Amenta et al. 1992). Under these conditions, we consider conjunctive analysis to be as theoretically informed as other methods and its focus on description of data patterns as a major strength of this method.

Conclusions

The conjunctive analysis of case configurations is a simple method of discrete multivariate analysis that can be easily applied to both exploratory and confirmatory research. This method offers a middle ground between (1) the focus on specificity and multiple causality that underlies most qualitative research and (2) the variable-oriented search for general patterns across contexts in most quantitative research (see Ragin1987, 2000). By assuming maximum causal complexity and then using basic methods to visually identify patterns within these case configurations, the conjunctive approach described here provides a simple way of addressing both of these concerns.

Even for researchers who prefer more variable-oriented quantitative methods, conjunctive analysis can augment and inform their substantive analysis by identifying possible problems with multicollinearity and low-cell frequencies. Although conjunctive methods have been used for confirmatory research in criminology, the simple analysis of a conjunctive matrix of case configurations provides a succinct and visually appealing way to both explore and confirm the nature of case concentration, diversity, and complex causal patterns among multiple categorical variables.

Appendix 1: Software Syntax for Conjunctive Analysis

For each of the following examples, ABCD = categorical independent variables and Y = categorical dependent variable.

SPSS Syntax for Generating Conjunctive Data Matrix:

```
AGGREGATE
/OUTFILE = 'cdmatrix_file'
/BREAK = A B C D
/Y_mean = MEAN(Y)
/N_Cases = N.
```

STATA Syntax for Generating Conjunctive Data Matrix:

```
egen N_Cases = count(Y), by (A B C D)
collapse (count) N_Cases (mean) Y_MEAN = Y, by (A B C D)
list A B C D Y_MEAN N_Cases
```

SAS Syntax for Generating Conjunctive Data Matrix:

```
proc means data = yourdata nway;
class a b c d;
var y;
output out = cdmatrix(drop=_type_ _freq_) mean = n= / autoname;
run;
proc print data = cdmatrix;
run;
```

References

Amenta E, Carruthers BG, Zylan Y (1992) A hero for the aged? The Townsend movement, the political mediation model, and U.S. old-age policy, 1934–1950. Am J Sociol 98(2):308–339

Amenta E, Halfmann D (2000) Wage wars: institutional politics, WPA wages, and the struggle for U.S. social policy. Am Sociol Rev 67:506–528

Bishop YMM, Fienberg SE, Holland PW (1975) Discrete multivariate analysis. The MIT Press, Cambridge, MA

Blumstein A, Cohen J, Martin S, Tonry M (1983) Research on sentencing: the search for reform, vol 1. National Academy Press, Washington, DC

Charemza WW, Deadman DF (1997) New directions in econometric practice: general to specific modelling, cointegration, and vector autoregression, 2nd edn. Edward Elgar Publishing, Cheltenham, UK

Chiricos T, Crawford C (1995) Race and imprisonment: a contextual assessment of the evidence. In: Hawkins D (ed) Ethnicity, race, and crime. State University of New York Press, Albany, NY

Drass KA, Ragin CC (1992) QCA: qualitative comparative analysis. Institute for Policy Research, Northwestern University, Evanston, IL

Goodman LA (1972) A modified multiple regression approach to the analysis of dichotomous variables. Am Sociol Rev 37:28–46

Hagan J (1974) Extra-legal attributes and criminal sentencing: an assessment of a sociological viewpoint. Law Soc Rev 8:357–383

Johnson B (2005) Contextual disparities in guideline departures: courtroom social contexts, guideline compliance, and extralegal disparities in criminal sentencing. Criminology 43(3):761–796

LaFree GD, Birkbeck C (1991) The neglected situation: a cross-national study of the situational characteristics of crime. Criminology 29:73–98

Myers M, Talarico S (1987) The social context of criminal sentencing. Springer-Verlag, New York, NY

Miethe TD, Drass KA (1999) Exploring the social context of instrumental and expressive homicides: an application of qualitative comparative analysis. J Quant Criminol 15(1):1–21

Miethe TD, Lu H, Deibert G (2005) Cross-national variability in capital punishment: exploring the sociopolitical sources of its differential legal status. Int Crim Justice Rev 15:115–131

Miethe TD, Moore CA (1986) Racial differences in criminal processing: the consequence of model selection for conclusions about differential treatment. Sociol Q 27(2):217–237

Miethe TD, Regoeczi WC (2004) Rethinking homicide: exploring the structure and process underlying deadly situations. Cambridge University Press, New York, NY

Peterson R, Hagan J (1984) Changing conceptions of race: toward an account of anomalous findings of sentencing research. Am Sociol Rev 49:56–70

Ragin CC (1987) The comparative method. The University of California Press, Berkeley, CA

Ragin CC (2000) Fuzzy-set social science. The University of Chicago Press, Chicago, IL

Steffensmeier D, Ulmer JT, Kramer J (1998) The interaction of race, gender, and age in criminal sentencing: the punishment cost of being young, black, and male. Criminology 36:763–798

Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading, MA

Ulmer J (1997) Social worlds of sentencing: court communities under sentencing guidelines. State University of New York Press, Albany, NY

von Eye A (2002) Configural frequency analysis: methods, models, and application. Lawrence Erlbaum Associates, New York, NY

von Eye A, Bogat GA, Rhodes JE (2006) Variable-oriented and person-oriented perspectives of analysis: the example of alcohol consumption in adolescence. J Adolesc 29:981–1004

Zatz M (1987) The changing form of racial/ethnic biases in sentencing. J Res Crime Delinq 25:69–92