Discussion Paper No. 142

QUALITATIVE VARIABLES AND SIMULTANEOUS EQUATION ECONOMETRIC MODELS*

by

Jay Magidson

May 1975

Presented at the NBER-NSF Conference on Decision Making Under Uncertainty, University of Chicago, May 16, 1975.

---

*The author is indebted to discussions with Jim Swan, Richard Bagozzi, Professors Richard Berk, Robert Boruch, Donald Campbell, Marc Nerlove and Avinoam Perry, and especially to Professor Leo Goodman for his excellent course in qualitative variables at the University of Chicago. This research was supported by Research Contract No. NIE-C-74-0115 from the National Institute of Education.

1. Introduction

Many research studies, particularly in the social sciences, deal with qualitative variables in addition to quantitative variables. Such variables may be ordered or unordered. For example, sex is an unordered (nominal) dichotomous variable since it consists of two unordered categories, namely, male and female. Social class is an example of an ordered (ordinal) qualitative variable since it classifies the population into ordered categories, say upper, middle and lower social classes. Similarly, attitudes are either unfavorable, neutral or favorable.

Working toward the ultimate goal of a unified approach to data analysis, qualitative and quantitative variables are tied together in sections 2 and 3, the connections being established through the use of dummy variables and the concept of variance. The variance of a qualitative variable is defined in section 2. Quantitative variance is shown to be a weighted version of qualitative variance.

Via joint variables and the concept of joint variance introduced at the end of section 2, the generalization from regression analysis to simultaneous equations is immediate in the case of qualitative or discrete endogenous variables.

Correlation ratios are derived in section 3. In the case that interaction variables are not necessary, the correlation ratio reduces to the square of the (linear) correlation coefficient. It is shown how the interaction variables can be defined in section 4.

A formulation for the analysis of qualitative variance, MANQUOVA, is developed in section 4. The effect parameters are (partial) regression coefficients from a "collapsed" regression model. For a hypothetical example, it is shown that these measures of effect have many very "intuitive" properties.

In section 5 an estimation procedure is suggested and it is shown how missing data can be handled in a straightforward manner.

This research is the result of pursuing various similarities between Goodman's qualitative regression model (see e.g. Goodman (1972a, 1972b)) and the standard quantitative model as characterized in econometrics (see e.g. Goldberger (1964)).

2. The Variance of a Qualitative Variable

Variance is defined for quantitative data in terms of deviations from the mean. For qualitative (nominal or ordinal) variables however, the mean is an undefined concept. In this section dummy variables will be used to define the "variation" associated with each category of a qualitative polytomy. The variance is then defined as the sum of these variations less the excess due to the "overlap variance" among the categories. It is then shown that quantitative variance is a weighted version of qualitative variance, the weights being a function of the values

assigned to the variable.

The concept of joint variance is then introduced for both qualitative and quantitative discrete variables. Joint variance can be used in simultaneous equations models in the same manner that variance is analyzed in regression models.

Let A represent a polytomous variable classifying the population into I distinct categories. The distribution of A is then given by the vector
$P^A = (P_1^A, P_2^A, \ldots, P_i^A, \ldots, P_I^A)$ where

$P_i^A \equiv$ the proportion of the population classified as $A = A_i$

Next, define the random dummy vector
$Z^A = (Z_1^A, Z_2^A, \ldots, Z_i^A, \ldots, Z_I^A)$ as follows

$$Z_i^A \equiv \begin{cases} 1 \text{ if } A = A_i \\ 0 \text{ otherwise} \end{cases}$$

Thus, Z_1 for example indicates whether or not an observation is classified as $A = A_1$. (For convenience, at times the superscripts will be omitted when there is no chance of confusion.)

Each Z_i has a Bernoulli distribution so that its variance is

$$VAR(Z_i) = P_i(1 - P_i) \tag{1}$$

$VAR(Z_i)$ is a measure of the dispersion associated with category A_i. For A dichotomous ($I = 2$), Z_1^A and Z_2^A provide identical information, namely, each variable pinpoints which category of A an individual is in, A_1 or A_2. Thus, Z_1 and Z_2 are each information-equivalents for A. For $I = 2$ we therefore define VAR(A) as $VAR(Z_1^A)$ or equivalently as $VAR(Z_2^A)$

$$\begin{aligned} VAR(A) &= P_1(1 - P_1) \\ &= P_2(1 - P_2) \\ &= P_1 P_2 \qquad (I = 2) \end{aligned} \tag{2}$$

Notice that this definition is consistent with the variance of A attained by the usual methods of assigning the values 1 and 0 or 1 and 2 to the two categories.

In general, for $I \geq 2$ categories, since any individual is in one and only one category of A, the Z_i's must sum to one so that any Z_i, say Z_{i*}, can be expressed in terms of the other Z_i's as

$$Z_{i*} = 1 - \sum_{i=1}^{i*-1} Z_i - \sum_{i=i*+1}^{I} Z_i \tag{3}$$

Since each Z_{i*} is perfectly predicted from the others, the variance of Z_{i*} can be partitioned into a portion explainable by the categories preceding category A_{i*} and a portion explainable by the categories following category

A_{i*} respectively as follows

$$\mathrm{VAR}(Z_{i*}) = \mathrm{COV}(Z_{i*},\ 1 - \sum_{i=1}^{i*-1} Z_i - \sum_{i=i*+1}^{I} Z_i)$$

$$= P_{i*} \sum_{i=1}^{i*-1} P_i + P_{i*} \sum_{i=i*+1}^{I} P_i \tag{4}$$

$$= \left(\sum_{i=1}^{i*} P_i\right)^2 \left(\frac{P_{i*}}{\sum_{i=1}^{i*} P_i}\right)\left(1 - \frac{P_{i*}}{\sum_{i=1}^{i*} P_i}\right)$$

$$+ \left(\sum_{i=i*}^{I} P_i\right)^2 \left(\frac{P_{i*}}{\sum_{i=i*}^{I} P_i}\right)\left(1 - \frac{P_{i*}}{\sum_{i=i*}^{I} P_i}\right)$$

$$= \left(\sum_{i=1}^{i*} P_i\right)^2 \mathrm{VAR}(Z_{i*} | Z_{i*+1} = Z_{i*+2} = \ldots = Z_I = 0)$$

$$+ \left(\sum_{i=i*}^{I} P_i\right)^2 \mathrm{VAR}(Z_{i*} | Z_1 = Z_2 = \ldots = Z_{i*-1} = 0) \tag{5}$$

The total variance of A is defined here as the following composition of these categorical components. It is the variation associated with A_1 plus the variation due to A_2 but unexplainable by A_1, plus the variation due to A_3 but not explainable by the preceding categories A_1 and A_2, etc. By summing over the second term of (4) for all categories, VAR(A) is thus

$$\mathrm{VAR}(A) = \sum_{i=1}^{I-1} \sum_{k>i} P_i P_k = \sum_{i=2}^{I} \sum_{k<i} P_i P_k$$

$$= \frac{1}{2} \sum_{i=1}^{I} P_i \sum_{\substack{k=1 \\ k \neq i}}^{I} P_k \tag{6}$$

$$= \frac{1}{2} \sum_{i=1}^{I} P_i (1 - P_i)$$

$$= \frac{1}{2} \sum_{i=1}^{I} \mathrm{VAR}(Z_i) \tag{7}$$

$$= \frac{1}{2}(1 - \sum_{i=1}^{I} P_i^2) \tag{8}$$

Notice that VAR(A) can be computed by summing the $\mathrm{VAR}(Z_i)$ and dividing by 2. Thus, this definition does not depend upon the ordering of the categories. For ordinal variables, although the ordering is not explicitly taken into account, VAR(A) still provides a **meaningful measure** of variation. In section 4 it is shown how order can be included in the model. Also, various hypotheses regarding order can either be tested or imposed on the data using the estimation procedure given in section 5.

This definition for the variance of A has the following desirable properties.

1. For a fixed number of categories, VAR(A) is maximized when the probability of being in each category is equally likely.

2. Under the condition of equally likely probabilities, VAR(A) approaches its maximum value of .5 as the number of categories approaches infinity, i.e.,

$$\lim_{I \to \infty} \frac{1}{2}[1 - \sum_{i=1}^{I} (\frac{1}{I})^2] = \frac{1}{2}$$

3. VAR(A) approaches its minimum value of zero as the probability of being in any one particular category approaches one.

Before examining the relationship between qualitative and quantitative variables, consider partitioning VAR(A) in the case of I = 4 categories. From (4) and (6) we have

$$VAR(Z_1^A) = P_1P_2 + P_1P_3 + P_1P_4$$

$$VAR(Z_2^A) = P_1P_2 + P_2P_3 + P_2P_4$$

$$VAR(Z_3^A) = P_1P_3 + P_2P_3 + P_3P_4$$

$$VAR(Z_4^A) = P_1P_4 + P_2P_4 + P_3P_4$$

and

$$VAR(A) = P_1P_2 + P_1P_3 + P_1P_4 + P_2P_3 + P_2P_4 + P_3P_4 \quad (9)$$

Category A_1 accounts for the first 3 terms of VAR(A) via $VAR(Z_1)$. Next, category A_2 contributes $P_2P_3 + P_2P_4$, P_1P_2 having been already taken into account by category A_1. P_1P_2 is that portion of $VAR(Z_2)$ explained by the category preceding A_2, while $P_2P_3 + P_2P_4$ is explained by the categories following A_2. Category A_3 contributes the

final term P_3P_4, the only part of VAR(Z_3) not explained by the preceding categories. Category A_4 adds no new information, which follows from (3) with $i^* = 4$. It can easily be verified that VAR(A) remains unchanged under any reordering of the 4 categories.

Gini (1912) proposed precisely the same formula for VAR(A) using a very different approach. He first noted that quantitative variance can be expressed solely as a function of pairwise differences. Specifically, for a discrete random variable Y, taking on I possible values, the variance is

$$\mathrm{VAR}(Y) = \frac{1}{2} \sum_{i=1}^{I} \sum_{k=1}^{I} P_i P_k (Y_i - Y_k)^2 \tag{10}$$

$$= \frac{1}{2} \sum_{i=1}^{I} \sum_{k=1}^{I} P_i P_k d_{ik}^2 \tag{11}$$

where

$$d_{ik} \equiv |Y_i - Y_k| \tag{12}$$

Reasoning by analogy, for each of the $\binom{I}{2}$ possible pairs (A_i, A_k) selected with replacement, define d_{ik} as

$$d_{ik} \equiv \begin{cases} 1 \text{ if } i \neq k \\ 0 \text{ if } i = k \end{cases} \tag{13}$$

Analogous to (11), VAR(A) is then

$$\mathrm{VAR}(A) = \frac{1}{2} \sum_{i=1}^{I} \sum_{\substack{k=1 \\ k \neq i}}^{I} P_i P_k$$

which is precisely the same as eq. (6). Thus, there are at least two different justifications for calculating the variance of a qualitative variable by this formula.

In order to see the relationship between these approaches, define Z_i^Y as

$$Z_i^Y \equiv \begin{cases} 1 \text{ if } Y = Y_i \\ 0 \text{ otherwise} \end{cases} \tag{14}$$

and notice that Y can be expressed in the following "algebraic-tree" form

$$Y = Y_1 Z_1^Y + Y_2 Z_2^Y + \ldots + Y_I Z_I^Y \tag{15}$$

Calculating VAR(Y) directly from (15) we have

$$\begin{aligned}
\mathrm{VAR}(Y) &= \sum_{i=1}^{I} \mathrm{COV}(Y, Y_i Z_i^Y) \\
&= \sum_{i=1}^{I} [Y_i^2 P_i (1 - P_i) - \sum_{\substack{k=1 \\ k \neq i}}^{I} Y_i Y_k P_i P_k] \\
&= \frac{1}{2}[\sum_{i=1}^{I} Y_i^2 P_i (1 - P_i) - 2 \sum_{i=1}^{I} \sum_{\substack{k=1 \\ k \neq i}}^{I} Y_i Y_k P_i P_k + \sum_{k=1}^{I} Y_k^2 P_k (1 - P_k)] \\
&= \frac{1}{2}[\sum_{i=1}^{I} Y_i^2 P_i \sum_{\substack{k=1 \\ k \neq i}}^{I} P_k - 2 \sum_{i=1}^{I} \sum_{\substack{k=1 \\ k \neq i}}^{I} Y_i Y_k P_i P_k + \sum_{k=1}^{I} Y_k^2 P_k \sum_{\substack{i=1 \\ i \neq k}}^{I} P_i]
\end{aligned}$$

$$= \frac{1}{2} \sum_{k=1}^{I} \sum_{i=1}^{I} (Y_k - Y_i)^2 P_i P_k$$

$$= \frac{1}{2} \sum_{i=1}^{I} P_i(1 - P_i) E[(Y - Y_i)^2 | Y \neq Y_i] \qquad (16)$$

Comparing (16) with (7) it is seen that qualitative variance is the basic ingredient of quantitative variance. Factoring out the qualitative portion from each term of the quantitative variance converts from a fixed mean formulation to a floating mean formulation where Y_i serves as the floating mean in the ith term.

Another convenient form for VAR(Y) is

$$\mathrm{VAR}(Y) = \sum_{i=1}^{I} \sum_{k=i+1}^{I} P_i P_k (Y_k - Y_i)^2 \qquad (17)$$

For Y continuous, letting g represent the density for Y, VAR(Y) can also be expressed in terms of the mean difference (for a proof see Kendall and Stuart (1963, p. 47).)

$$\mathrm{VAR}(Y) = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (Y - X)^2 g(X) g(Y) \, dX dY \qquad (18)$$

Next, we introduce another polytomous variable, B, classifying the population in J groups. The marginal distribution for B is given by $P^B = (P_1^B, P_2^B, \ldots, P_j^B, \ldots, P_J^B)$ and the joint distribution for A and B is given by $P^{AB} = (P_{11}^{AB}, P_{12}^{AB}, \ldots, P_{IJ}^{AB})$.

The marginal and joint distributions can be conveniently summarized in a 2-way table. The general I x J table is illustrated in Table 1.

Define the dummy vector $Z^B = (Z_1^B, Z_2^B, \ldots, Z_j^B, \ldots, Z_J^B)$ and the joint vector $Z^{AB} = (Z_{11}^{AB}, Z_{12}^{AB}, \ldots, Z_{ij}^{AB}, \ldots, Z_{IJ}^{AB})$ as follows

$$Z_j^B = \begin{cases} 1 \text{ if } B = B_j \\ 0 \text{ otherwise} \end{cases}$$

and

$$Z_{ij}^{AB} = \begin{cases} 1 \text{ if } A = A_i \text{ and } B = B_j \\ 0 \text{ otherwise} \end{cases}$$

Notice that the joint variable equals the product of its marginal components

$$Z_{ij}^{AB} = Z_i^A Z_j^B \tag{19}$$

The joint distribution of A and B can be equivalently thought of as the marginal distribution of the single joint variable AB. Viewing AB as a single polytomy with IJ categories we then have from (1) and (9)

$$\mathrm{VAR}(Z_{ij}^{AB}) = P_{ij}(1 - P_{ij}) \tag{20}$$

and

$$\mathrm{VAR}(AB) = \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} P_{ij}(1 - P_{ij}) \tag{21}$$

Table 1. The marginal and joint distribution for A and B.

A\B	B_1	B_2	...	B_J	TOTAL
A_1	P_{11}	P_{12}	...	P_{1J}	P_1^A
A_2	P_{21}	P_{22}	...	P_{2J}	P_2^A
$\vdots$				$\vdots$	$\vdots$
A_I	P_{I1}	P_{I2}	...	P_{IJ}	P_I^A
TOTAL	P_1^B	P_2^B	...	P_J^B	

Similarly, we introduce the discrete quantitative variable X taking on J possible values $X_1, X_2, \ldots, X_J$. Analogous to (19) define the joint variable YX as the product of Y and X. We then have

$$YX = \sum_{i=1}^{I} \sum_{j=1}^{J} Y_i X_j Z_{ij}^{YX} \tag{22}$$

and as in (16), the variance of the joint variable YX can be expressed as

$$VAR(YX) = \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} P_{ij}(1 - P_{ij}) E[(YX - Y_i X_j)^2 | (Y,X) \neq (Y_i, X_j)] \tag{23}$$

3. Derivation of the Correlation Ratios

The correlation ratio η^2 measures the proportion of the variance of the dependent variable explainable by the independent variables. When the population regression of the dependent variable is linear in the independent variables, the correlation ratio equals the square of the (linear) correlation coefficient ρ^2.

In this section, the multiple correlation ratios are derived when the independent variables are other than continuous. In the case of a qualitative or discrete (quantitative) dependent variable, since this variable can be viewed as a joint variable, the generalization for simultaneous equations to the "joint correlation ratio" is

straightforward.

The variance of A will be partitioned into a portion unexplainable by B and a portion explainable by B by partitioning $VAR(Z_i^A)$ and accumulating these two portions via (7).

$$
\begin{aligned}
VAR(Z_i^A) &= P_i(1 - P_i) \\
&= \sum_{j=1}^{J} P_{ij}(1 - P_{ij}|P_j + P_{ij}|P_j - P_i) \\
&= \sum_{j=1}^{J} P_j \frac{P_{ij}}{P_j}\left(1 - \frac{P_{ij}}{P_j}\right) + \sum_{j=1}^{J} P_j \frac{P_{ij}}{P_j}\left(\frac{P_{ij}}{P_j} - P_i\right) \\
&= \sum_{j=1}^{J} P_j VAR(Z_i|B = B_j) + \sum_{j=1}^{J} P_j[VAR(Z_i) - VAR(Z_i|B = B_j)]
\end{aligned}
\tag{24}
$$

The first sum above, representing the portion of variance of Z_i^A unexplainable by B, equals zero corresponding to perfect prediction of Z_i from B if and only if the group probability of being in category A_i equals one or zero for each of the J groups. The second sum represents the

explainable portion. It equals zero if and only if the probability that $A = A_i$ is the same for each B_j group.

From (7) and (8), the correlation ratio is thus

$$\eta^2_{A.B} = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{P_{ij}}{P_j} (P_{ij} - P_i P_j)}{1 - \sum_{i=1}^{I} P_i^2} \tag{25}$$

Light and Margolin (1971) developed this same measure of association for samples (calling it R^2) by partitioning Gini's variance in a manner similar to that above using conventional analysis of variance techniques for this case one one dependent and one independent variable.

This correlation ratio is identical to the Goodman-Kruskal Tau-b, developed under somewhat different conditions. Regarding this coefficient, Goodman and Kruskal (1954, p. 760) write

> It is clear that τ_b takes on values between 0 and 1; it is 0 if and only if there is independence, and 1 if and only if knowledge of A_a completely determines B_b. Finally, τ_b is indeterminate if and only if both independence and determinism silultaneously hold, that is if all $\rho_{.b}$'s but one are zero.

Next suppose that there are M dependent (endogenous) variables, the mth of which has I_m categories and G independent (exogenous) variables, the gth of which has J_g categories. Then a 2 way table can be formed whose $I = I_1 \times I_2 \times \ldots \times I_M$ rows represent all possible groupings

of the dependent variables and whose $J = J_1 \times J_2 \times \ldots \times J_G$ columns represent all possible groupings of the independent variables. Hence, any joint correlation ratio can be calculated from the usual formulae treating the dependent variables as one joint dependent variable with I categories and the independent variables as one joint independent variable with J categories.

For example, Table 2 illustrates the 2-way table for the case of two dichotomous dependent variables and two dichotomous independent variables.

For this case we have

$$\eta^2_{AB.CD} = \frac{\sum_{(i,j)} \sum_{(k,\ell)} \frac{P_{ijk\ell}}{P_{k\ell}} (P_{ijk\ell} - P_{ij}P_{k\ell})}{1 - \sum_{(i,j)} P^2_{ij}} \tag{26}$$

Now consider the case of a quantitative dependent variable Y and either a qualitative or quantitative discrete independent variable X. The regression of Y on X is

$$E(Y|X) = \sum_{j=1}^{J} \mu_j Z_j^X \tag{27}$$

where

$$\mu_j \equiv E(Y|X = X_j) \tag{28}$$

Table 2. The marginal and joint distribution for AB and CD where A, B, C and D are dichotomies

AB\CD	C_1D_1	C_1D_2	C_2D_1	C_2D_2	TOTAL
A_1B_1	P_{1111}	P_{1112}	P_{1121}	P_{1122}	P^{AB}_{11}
A_1B_2	P_{1211}	P_{1212}	P_{1221}	P_{1222}	P^{AB}_{12}
A_2B_1	P_{2111}	P_{2112}	P_{2121}	P_{2122}	P^{AB}_{21}
A_2B_2	P_{2211}	P_{2212}	P_{2221}	P_{2222}	P^{AB}_{22}
TOTAL	P^{CD}_{11}	P^{CD}_{12}	P^{CD}_{21}	P^{CD}_{22}	

From (17), with μ_j in place of Y_i we then have

$$\mathrm{VAR}[E(Y|X)] = \sum_{j=1}^{J} \sum_{k=j+1}^{J} P_j P_k (\mu_j - \mu_k)^2 \tag{29}$$

and the correlation ratio is calculated from the usual formula

$$\eta^2_{Y.X} = \frac{\mathrm{VAR}[E(Y|X)]}{\mathrm{VAR}(Y)} \tag{30}$$

The correlation ratio is zero if and only if the conditional expectation of Y given X is the same for all values of X. It is unity if and only if Y is an exact function of X. It can easily be verified that η^2 is invariant with respect to any linear transformation of Y and does not depend on X at all.

For Y discrete, μ_j can be written in terms of the probabilities

$$Y = \sum_{i=1}^{I} Y_i Z_i^Y$$

so that

$$E(Y|X) = \sum_{i=1}^{I} Y_i \sum_{j=1}^{J} \frac{P_{ij}}{P_j} Z_j^X$$

$$= \sum_{j=1}^{J} \sum_{i=1}^{I} Y_i \frac{P_{ij}}{P_j} Z_j^X$$

and thus from (27)

$$\mu_j = \sum_{i=1}^{I} Y_i \frac{P_{ij}}{P_j} \tag{31}$$

For this case (29) becomes

$$\mathrm{VAR}[E(Y|X)] = \sum_{j=1}^{J} \sum_{k=j+1}^{J} P_j P_k \sum_{i=1}^{I} \left[Y_i \left(\frac{P_{ij}}{P_j} - \frac{P_{ik}}{P_k} \right) \right]^2 \tag{32}$$

In the case that X is discrete (as opposed to qualitative) define α and β_j as

$$\alpha \equiv \mu_{j*} \quad \text{where } X_{j*} = 0$$

and

$$\beta_j \equiv \mu_j - \alpha \quad \text{for } X_j \neq 0$$

Then (27) becomes

$$E(Y|X) = \alpha + \beta_1 X_1 Z_1^X + \beta_2 X_2 Z_2^X + \ldots + \beta_J X_J Z_J^X \tag{33}$$

In this case, the correlation ratio reduces to the square of the coefficient of determination ρ^2 in the following 2 cases

1) $\beta_1 = \beta_2 = \ldots = \beta_J$
2) $I = J = 2$ (i.e., both Y and X are dichotomous)

4. The Additive and Multiplicative Effects Models

The general simultaneous equation models developed in Magidson (1975) consist of $2^M - 1$ endogenous variables and $2^G - 1$ exogenous variables. These represent all interaction variables in addition to the M first order endogenous and the G first order exogenous variables. The endogenous variables are either 1) all quantitative and discrete, 2) all qualitative or 3) one continuous variable. The exogenous variables are either 1) quantitative and discrete, 2) qualitative or 3) a mixture of these two.

In this section we consider the case of G = 2 first order exogenous variables. The extension to the general case for $G \geq 1$ is direct. We first suppose that the endogenous variables are qualitative.

Let A represent the joint variable composed of all endogenous variables and let BC represent the joint variable consisting of the 2 exogenous variables B and C. Defining the regression of A on BC in terms of the information equivalents Z^A and Z^{BC} we have

$$E(A|BC) \equiv E(Z^A|Z^{BC}) = E\begin{pmatrix} Z_1^A|Z^{BC} \\ Z_2^A|Z^{BC} \\ \vdots \\ Z_I^A|Z^{BC} \end{pmatrix} \tag{34}$$

Thus, the regression of A on BC can be decomposed into I separate sub-regressions. The ith sub-regression

can be expressed in a form where the regression coefficients are the group probabilities, as follows

$$E(Z_i^A | Z^{BC}) = \sum_{j=1}^{J} \sum_{k=1}^{K} \beta_{ijk} Z_{jk}^{BC} \tag{35}$$

where

$$\beta_{ijk} \equiv \frac{P_{ijk}}{P_{jk}} = \text{Prob}(A = A_i | Z_{jk}^{BC} = 1) \tag{36}$$

We wish to arrive at separate measures for the effect of B, of C and their interaction on A. To accomplish this,[1] consider having information regarding the values of the disjoint variables Z_{jk}^{BC}, $Z_{j\bar{k}}^{BC}$, and $Z_{\bar{j}k}^{BC}$ where

$$Z_{j\bar{k}}^{BC} \equiv \begin{cases} 1 \text{ if } B = B_j \text{ and } C \neq C_k \\ 0 \text{ otherwise} \end{cases} \tag{37}$$

and

$$Z_{\bar{j}k}^{BC} \equiv \begin{cases} 1 \text{ if } B \neq B_j \text{ and } C = C_k \\ 0 \text{ otherwise} \end{cases} \tag{38}$$

Similarly, define $Z_{\bar{j}\bar{k}}^{BC}$ by

$$Z_{\bar{j}\bar{k}}^{BC} \equiv \begin{cases} 1 \text{ if } B \neq B_j \text{ and } C \neq C_k \\ 0 \text{ otherwise} \end{cases} \tag{39}$$

We then have the following "collapsed" model

---

[1]The approach taken in Magidson (1975) is different from the approach given here and is consistent with the regression approach to ANOVA. When all variables are dichotomous, the two approaches coincide.

$$E(Z_i^A | Z_{jk}^{BC}, Z_{j\bar{k}}^{BC}, Z_{\bar{j}k}^{BC}) = \beta_{ijk} Z_{jk}^{BC} + \beta_{ij\bar{k}} Z_{j\bar{k}}^{BC} + \beta_{i\bar{j}k} Z_{\bar{j}k}^{BC} + \beta_{i\bar{j}\bar{k}} Z_{\bar{j}k}^{BC} \tag{40}$$

where

$$\beta_{ij\bar{k}} \equiv \frac{P_{ij} - P_{ijk}}{P_j - P_{jk}} \equiv \frac{P_{ij\bar{k}}}{P_{j\bar{k}}} \tag{41}$$

$$\beta_{i\bar{j}k} \equiv \frac{P_{ik} - P_{ijk}}{P_k - P_{jk}} \equiv \frac{P_{i\bar{j}k}}{P_{\bar{j}k}} \tag{42}$$

and

$$\beta_{i\bar{j}\bar{k}} \equiv \frac{P_i - P_{ij} - P_{ik} + P_{ijk}}{1 - P_j - P_k + P_{jk}} \equiv \frac{P_{i\bar{j}\bar{k}}}{P_{\bar{j}\bar{k}}} \tag{43}$$

The 4 independent variables in (40) can be decomposed symmetrically as follows

$$Z_{jk}^{BC} = \frac{1}{4}(1 + X_j^B + X_k^C + X_{jk}^{BC}) \tag{44}$$

$$Z_{j\bar{k}}^{BC} = \frac{1}{4}(1 + X_j^B - X_k^C - X_{jk}^{BC}) \tag{45}$$

$$Z_{\bar{j}k}^{BC} = \frac{1}{4}(1 - X_j^B + X_k^C - X_{jk}^{BC}) \tag{46}$$

and

$$Z_{\bar{j}\bar{k}}^{BC} = \frac{1}{4}(1 - X_j^B - X_k^C + X_{jk}^{BC}) \tag{47}$$

where

$$X_j^B \equiv \begin{cases} 1 \text{ if } Z_j^B = 1 \\ -1 \text{ otherwise} \end{cases} \tag{48}$$

$$X_k^C \equiv \begin{cases} 1 \text{ if } Z_k^C = 1 \\ -1 \text{ otherwise} \end{cases} \tag{49}$$

and

$$X_{jk}^{BC} \equiv X_j^B X_k^C \tag{50}$$

For an example with j = k = 1 for 2 trichotomies see table 3. Substituting eqs. (44) - (47) in (40) and collecting terms yields

$$E(Z_i^A | Z_{jk}^{BC}, Z_{j\bar{k}}^{BC}, Z_{\bar{j}k}^{BC}) = \lambda_i^{\bar{A}} + \lambda_{ij}^{\bar{A}B} X_j^B + \lambda_{ik}^{\bar{A}C} X_k^C + \lambda_{ijk}^{\bar{A}BC} X_{jk}^{BC} \tag{51}$$

where

$$\lambda_i^{\bar{A}} \equiv \frac{1}{4}(\beta_{ijk} + \beta_{ij\bar{k}} + \beta_{i\bar{j}k} + \beta_{i\bar{j}\bar{k}}) \tag{52}$$

$$\lambda_{ij}^{\bar{A}B} \equiv \frac{1}{4}(\beta_{ijk} - \beta_{i\bar{j}k} + \beta_{ij\bar{k}} - \beta_{i\bar{j}\bar{k}}) \tag{53}$$

$$\lambda_{ik}^{\bar{A}C} \equiv \frac{1}{4}(\beta_{ijk} - \beta_{ij\bar{k}} + \beta_{i\bar{j}k} - \beta_{i\bar{j}\bar{k}}) \tag{54}$$

and

$$\lambda_{ijk}^{\bar{A}BC} \equiv \frac{1}{4}(\beta_{ijk} - \beta_{ij\bar{k}} - \beta_{i\bar{j}k} + \beta_{i\bar{j}\bar{k}}) \tag{55}$$

The analysis of variance formulation induced by model (51), called the collapsed version of MANQUOVA, is

Table 3. Illustration of the decomposition of the collapsed joint variable

j^*k^*	Z_{11}^{BC}	$Z_{1\bar{1}}^{BC}$	$Z_{\bar{1}1}^{BC}$	$Z_{\bar{1}\bar{1}}^{BC}$	X_1^B	X_1^C	X_{11}^{BC}
1 1	1	0	0	0	1	1	1
1 2	0	1	0	0	1	-1	-1
1 3	0	1	0	0	1	-1	-1
2 1	0	0	1	0	-1	1	-1
2 2	0	0	0	1	-1	-1	1
2 3	0	0	0	1	-1	-1	1
3 1	0	0	1	0	-1	1	-1
3 2	0	0	0	1	-1	-1	1
3 3	0	0	0	1	-1	-1	1

$$E(Z_i^A | Z_{jk}^{BC} = 1) = \lambda_i^A + \lambda_{ij}^{\overline{A}B} + \lambda_{ik}^{\overline{A}C} + \lambda_{ijk}^{\overline{A}BC} \tag{56}$$

Let us consider the example in table 4 where it is desired to measure the effect of an advertisement on the probability of buying a product. Since A is a single first order variable, the simultaneous equations model reduces to a regression model. Also, since A is dichotomous (I = 2), model (34) simplifies to a single equation, the second equation being a restatement of the first. We have

$$E(Z_1^A | Z^{BC}) = \beta_{111} Z_{11}^{BC} + \beta_{112} Z_{12}^{BC} + \beta_{121} Z_{21}^{BC} + \beta_{122} Z_{22}^{BC} \tag{57}$$

and

$$\lambda_1^{\overline{A}} = \frac{1}{4}(\beta_{111} + \beta_{112} + \beta_{121} + \beta_{122}) = .6 \tag{58}$$

$$\lambda_{11}^{\overline{A}B} = \frac{1}{4}(\beta_{111} - \beta_{121} + \beta_{112} - \beta_{122}) = .325 \tag{59}$$

$$\lambda_{11}^{\overline{A}C} = \frac{1}{4}(\beta_{111} - \beta_{112} + \beta_{121} - \beta_{122}) = .125 \tag{60}$$

and

$$\lambda_{111}^{\overline{A}BC} = \frac{1}{4}(\beta_{111} - \beta_{112} - \beta_{121} + \beta_{122}) = -.1 \tag{61}$$

$$\lambda_{12}^{\overline{A}B} = -\lambda_{11}^{\overline{A}B} \quad \lambda_{12}^{\overline{A}C} = -\lambda_{11}^{\overline{A}C} \text{ and } \lambda_{122}^{\overline{A}BC} = -\lambda_{112}^{\overline{A}BC} = -\lambda_{121}^{\overline{A}BC} = \lambda_{111}^{\overline{A}BC} \tag{62}$$

Table 4. The 3-way table summarizing the effects of an advertisement

		MALES (B_1)				FEMALES (B_2)			
		(B_1,C_1) EXPERIMENTALS		(B_1,C_2) CONTROLS		(B_2,C_1) EXPERIMENTALS		(B_2,C_2) CONTROLS	
(A_1)	BUY	950	(95%)	9000	(90%)	5000	(50%)	50	(5%)
(A_2)	NO BUY	50	(5%)	1000	(10%)	5000	(50%)	950	(95%)
		1000	(100%)	10000	(100%)	10000	(100%)	1000	(100%)

The main (overall) effect of that ad is given by $\lambda_{11}^{\overline{A}C}$. It equals the total difference in the purchase behavior of the experimentals over the controls for the two sexes divided by 4, i.e., $\lambda_{11}^{\overline{A}C} = \frac{1}{4}(.05 + .45)$.

The interaction effect is given by $\lambda_{111}^{\overline{A}BC}$. A zero interaction effect indicates that the difference in purchase behavior between the experimentals and controls is the same for both sexes. The interaction effect thus measures the extent to which sex is relevant for measuring the effect of the ad, $\lambda_{111}^{\overline{A}BC} = \frac{1}{4}(.05 - .45)$.

Notice that if the main effect $\lambda_{11}^{\overline{A}C}$ and the interaction effect $\lambda_{111}^{\overline{A}BC}$ were both zero, from (60) and (61) it follows that the probability of buying for the experimentals and controls is the same for each sex.

$$(\lambda_{11}^{\overline{A}C} = \lambda_{111}^{\overline{A}BC} = 0) \Longrightarrow (\beta_{111} = \beta_{112} \text{ and } \beta_{121} = \beta_{122}) \quad (63)$$

The main or overall effect of the ad will be zero if the ad increased the probability of buying for (say) males the same amount as it decreased the probability of buying for females. In this case, the interaction effect will be nonzero, indicating that one needs to look at the effects for males and females separately.

The classical formulation for the analysis of variance is defined for a quantitative dependent variable Y. Taking $Y = Z_1^A$, the classical ANOVA (see e.g., Scheffé (1959, p. 91)) yields the identical effect parameters as MANQUOVA for this example. In general, the two formulations do not coincide even when all variables are dichotomous.

A major difference between the two formulations occurs when some of the variables have more than 2 categories. For example, suppose that B represents marital status where B_1 = married, B_2 = single and B_3 = widowed/divorced. Equation (54) leads to 3 measures for the main effect of the ad corresponding to the 3 levels of B. If the effect of the ad were positive for the married group and negative by the same amount for the unmarried, the first measure would equal zero. Similarly, if the effect were positive for singles and negative by the same amount for nonsingles, the second measure would equal zero. The main effect of the ad is said to be zero if and only if all 3 measures are zero.

Notice that no arbitrary restrictions are imposed on the MANQUOVA parameters such as certain parameters summing to zero. For this reason, the MANQUOVA parameters are more easily interpretable than those of ANOVA.

In table 5, the 2-way table is given in which the sex variable is omitted. In this case, it appears as if the ad has had a <u>negative</u> effect, reducing the probability of buying from .82 to .54.

This spurious result is due to an omitted variable in a quasi-experimental or nonexperimental design. When randomization is possible, spurious results such as these can be avoided (see Campbell and Boruch).

In the case that A is ordinal, one might also wish

Table 5. The 2-way table

		C_1 EXPERIMENTALS		C_2 CONTROLS	
(A_1)	BUY	5950	(54%)	9050	(82%)
(A_2)	NO BUY	5050	(46%)	1950	(18%)
		11000	(100%)	11000	(100%)

EXPERIMENTALS $\frac{.95(1000) + .50(10000)}{11000} = .54$

CONTROLS $\frac{.90(10000) + .05(1000)}{11000} = .82$

to use the following cumulative formulation along with (56)

$$E(\sum_{i=1}^{i*} Z_i^A | Z_{jk}^{BC} = 1) = \text{Prob}(A \leq A_{i*} | Z_{jk}^{BC} = 1)$$

$$= \sum_{i=1}^{i*} \lambda_i^{\bar{A}} + \sum_{i=1}^{i*} \lambda_{ij}^{\bar{A}B} + \sum_{i=1}^{i*} \lambda_{ik}^{\bar{A}C} + \sum_{i=1}^{i*} \lambda_{ijk}^{\bar{A}BC} \tag{65}$$

where the categories of A are ordered say from low to high.

Analogous to the additive effect models, a multiplicative effects model can be developed. The multiplicative version of (51) is

$$E(Z_i^A | Z_{jk}^{BC}, Z_{j\bar{k}}^{BC}, Z_{\bar{j}k}^{BC}) = \tau_i (\tau_{ij})^{X_j} (\tau_{ik})^{X_k} (\tau_{ijk})^{X_{jk}} \tag{66}$$

where

$$\tau_i \equiv (\beta_{ijk}\beta_{ij\bar{k}}\beta_{i\bar{j}k}\beta_{i\bar{j}\bar{k}})^{1/4} \tag{67}$$

$$\tau_{ij} \equiv (\beta_{ijk}\beta_{ij\bar{k}}/\beta_{i\bar{j}k}\beta_{i\bar{j}\bar{k}})^{1/4} \tag{68}$$

$$\tau_{ik} \equiv (\beta_{ijk}\beta_{i\bar{j}k}/\beta_{ij\bar{k}}\beta_{i\bar{j}\bar{k}})^{1/4} \tag{69}$$

$$\tau_{ijk} \equiv (\beta_{ijk}\beta_{i\bar{j}\bar{k}}/\beta_{ij\bar{k}}\beta_{i\bar{j}k}) \tag{70}$$

A multiplicative effects model <u>for the odds</u> can then be defined as follows

For A nominal, the model is

$$\Omega_{ijk}^{\overline{A}BC} = \frac{E(Z_i^A | Z_{jk}^{BC}, Z_{j\overline{k}}^{BC}, Z_{\overline{j}k}^{BC})}{1 - E(Z_i^A | Z_{jk}^{BC}, Z_{j\overline{k}}^{BC}, Z_{\overline{j}k}^{BC})}$$

$$= \gamma_i (\gamma_{ij})^{X_j} (\gamma_{ik})^{X_k} (\gamma_{ijk})^{X_{jk}} \qquad (71)$$

where

$$\gamma_i \equiv (\omega_{ijk}\omega_{ij\overline{k}}\omega_{i\overline{j}k}\omega_{i\overline{jk}})^{1/4}$$

$$\gamma_{ij} \equiv (\omega_{ijk}\omega_{ij\overline{k}}/\omega_{i\overline{j}k}\omega_{i\overline{jk}})^{1/4}$$

$$\gamma_{ik} \equiv (\omega_{ijk}\omega_{i\overline{j}k}/\omega_{ij\overline{k}}\omega_{i\overline{jk}})^{1/4}$$

$$\gamma_{ijk} \equiv (\omega_{ijk}\omega_{i\overline{jk}}/\omega_{ij\overline{k}}\omega_{i\overline{j}k})^{1/4}$$

$$\omega_{ijk} \equiv \frac{\beta_{ijk}}{1 - \beta_{ijk}}$$

$$\omega_{i\overline{j}k} \equiv \frac{\beta_{i\overline{j}k}}{1 - \beta_{i\overline{j}k}}$$

etc.

In the case that all variables are dichotomous, (71) is identical to Goodman's model (see Goodman (1972b)). In the general case, the formulation is somewhat different (see e.g. Goodman (1971)).

For A ordinal, one can also use the following formulation which utilizes the ordering of A.

$$\Omega^{\overline{A}BC}_{i^*jk} = \frac{E(\sum_{i=1}^{i^*} Z^A_i | Z^{BC}_{jk}, Z^{BC}_{j\overline{k}}, Z^{BC}_{\overline{j}k})}{E(\sum_{i=i^*+1}^{I} Z^A_i | Z^{BC}_{jk}, Z^{BC}_{j\overline{k}}, Z^{BC}_{\overline{j}k})} \tag{72}$$

where the same notation from (71) applies except that

$$\omega_{i^*jk} \equiv \frac{\text{Prob}(A \leq A_{i^*} | Z^{BC}_{jk} = 1)}{\text{Prob}(A > A_{i^*} | Z^{BC}_{jk} = 1)} = \frac{\sum_{i=1}^{i^*} \beta_{ijk}}{\sum_{i=i^*+1}^{I} \beta_{ijk}} \tag{73}$$

$$\omega_{i^*\overline{j}k} = \frac{\sum_{i=1}^{i^*} \beta_{i\overline{j}k}}{\sum_{i=i^*+1}^{I} \beta_{i\overline{j}k}} \tag{74}$$

etc.

For a discussion of the conditions under which an additive effects model will yield results similar to a multiplicative effects model see Goodman (1974). For a comparison of the results for certain data sets see Swan (1975).

When the endogenous variable(s) are quantitative we have the same results, the regression coefficients again being the group means. For Y a single continuous variable or a joint discrete variable we have

$$E(Y|Z^{BC}) = \sum_{j=1}^{J} \sum_{k=1}^{K} \mu_{jk} Z_{jk}^{BC} \tag{75}$$

where

$$\mu_{jk} \equiv E(Y|Z_{jk}^{BC} = 1) \tag{76}$$

Thus, for the MANQUOVA formulation, proceeding as before, we define $\mu_{j\bar{k}}$, $\mu_{\bar{j}k}$, and $\mu_{\overline{jk}}$ as follows

$$\mu_{j\bar{k}} \equiv E(Y|Z_{j\bar{k}}^{BC} = 1) = \frac{P_j\mu_j - P_{jk}\mu_{jk}}{P_j - P_{jk}} \tag{77}$$

$$\mu_{\bar{j}k} \equiv E(Y|Z_{\bar{j}k}^{BC} = 1) = \frac{P_k\mu_k - P_{jk}\mu_{jk}}{P_k - P_{jk}} \tag{78}$$

$$\mu_{\overline{jk}} \equiv E(Y|Z_{\overline{jk}} = 1) = \frac{\mu - P_j\mu_j - P_k\mu_k + P_{jk}\mu_{jk}}{1 - P_j - P_k + P_{jk}} \tag{79}$$

where

$$\mu = \sum_{j=1}^{J} \sum_{k=1}^{K} P_{jk}\mu_{jk}$$

$$= \sum_{j=1}^{J} P_j\mu_j$$

$$= \sum_{k=1}^{K} P_k\mu_k$$

We then have

$$E(Y|Z_{jk} = 1) = \lambda + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC} \tag{80}$$

where

$$\lambda = \frac{1}{4}(\mu_{jk} + \mu_{j\bar{k}} + \mu_{\bar{j}k} + \mu_{\bar{j}\bar{k}})$$

$$\lambda_j^B = \frac{1}{4}(\mu_{jk} - \mu_{\bar{j}k} + \mu_{j\bar{k}} - \mu_{\bar{j}\bar{k}})$$

$$\lambda_k^C = \frac{1}{4}(\mu_{jk} - \mu_{j\bar{k}} + \mu_{\bar{j}k} - \mu_{\bar{j}\bar{k}})$$

and

$$\lambda_{jk}^{BC} = \frac{1}{4}(\mu_{jk} - \mu_{j\bar{k}} - \mu_{\bar{j}k} + \mu_{\bar{j}\bar{k}})$$

5. A Suggested Estimation Procedure

Regarding the multiplicative effects model for the odds, a maximum likelihood estimation algorithm can be found in Goodman (1972a). Regarding additive effects models, various estimation methods have been proposed. See e.g. Cox (1970), Zellner and Lee (1965) and Nerlove and Press (1973) for some discussion of these.

In this section a new estimation procedure is suggested resulting in asympototically unbiased and normal estimators. At the end of this section it is shown how missing data can be handled in a straightforward manner without making any assumptions.

Consider the regression of A on B where for simplicity we assume A and B are qualitative. As in (35), for each category of A we have

$$Z_i^A = \sum_{j=1}^{J} \beta_{ij} Z_j^B + \epsilon_i \tag{81}$$

where ϵ_i is the orthogonal component of Z_i^A with respect to B, i.e., the regression error term.

From (10), the error variance is

$$\mathrm{VAR}(\epsilon) = \frac{1}{2} \sum_{i=1}^{I} \mathrm{VAR}(\epsilon_i) \tag{82}$$

But

$$\begin{aligned} \mathrm{VAR}(\epsilon_i) &= \sum_{j=1}^{J} [P_{ij}(1 - \beta_{ij})^2 + (P_j - P_{ij})(0 - \beta_{ij})^2] \\ &= \sum_{j=1}^{J} [P_{ij}(1 - \beta_{ij})^2 + (P_j - P_{ij})\beta_{ij}^2] \end{aligned} \tag{83}$$

Thus far this discussion has been in terms of populations whereas in practice one deals with samples from generally unknown populations. Given a sample of size N from an unknown population let f_{ij}^{AB} represent the observed frequency of $A = A_i$ and $B = B_j$. The error sum of squares from the sample as a function of β is then

$$\mathrm{SSE}(\beta) = \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} [f_{ij}^{AB}(1 - \beta_{ij})^2 + (f_j^B - f_{ij}^{AB})\beta_{ij}^2] \tag{84}$$

where

$$f_i^A = \sum_{j=1}^{J} f_{ij}^{AB} = \text{observed frequency for } A = A_i.$$

It can easily be verified that SSE(β) is convex in β. Minimizing (84) directly yields the sample proportions

$$\hat{\beta}_{ij} = \frac{f_{ij}^{AB}}{f_j} \tag{85}$$

which in the saturated (unrestricted) model are the maximum likelihood estimates for β.

Now suppose we wish to test the hypotheses given by some unsaturated model H_k. Since the observed proportions are consistent estimates of the true proportions under model H_k, we choose the estimates for β which solve the following constrained minimization problem

$$\min_{\beta} \text{SSE}(\beta)$$

subject to

1) H_k is true

2) $\sum_{j=1}^{J} f_j^B \beta_{ij} = f_i^A$ for each i

3) $\sum_{i=1}^{I} \beta_{ij} = 1$ for each j

and

4) $0 \leq \beta_{ij} \leq 1$ for each (i,j)

The second constraint fixes the marginal distribution of A. f_i^A/N can be considered to be the (prior) probability that $A = A_i$; prior to receiving information B. Under certain circumstances, one might wish to use other values than f_i^A for constraint 2.

The 3rd and 4th constraints incorporate known information into the model, namely, that probabilities are between 0 and 1 and sum to 1. In the case that these constraints are binding, they improve the efficiency of the estimators.

Under additive effects models such as the new ANOVA formulation, most hypotheses correspond to linear restrictions. In this case, the estimates for β will be those estimates which minimize the Lagrangian.

Expected frequencies are then calculated by

$$F_{ij}^{AB} = f_{\ j}\hat{\beta}_{ij} \tag{86}$$

and model H_k is tested using the usual chi-square goodness of fit test or the likelihood ratio chi-square test. For details on these tests see Goodman (1972a).

In the case of missing data (empty cells) we proceed as follows. Suppose for example, the 1st 3 cells in the ith row of the I x J table are missing. This corresponds to no information being provided regarding Z_1^B, Z_2^B and Z_3^B toward the explanation of Z_i^A. We thus have

$$E(Z_i^A | Z_4^B, Z_5^B, \ldots, Z_J^B) = \beta_{ij*}(Z_1^B + Z_2^B + Z_3^B) + \sum_{j=4}^{J} \beta_{ij} Z_j^B \tag{87}$$

where

$$\beta_{ij*} = \frac{\sum_{j=1}^{3} P_{ij}}{\sum_{j=1}^{3} P_j} \tag{88}$$

Thus, missing data fits into the model by collapsing the cells in each row having missing data and taking

$$f_{ij*} = f_i^A - \sum_{j=4}^{J} f_{ij} \tag{89}$$

Of course if other information is used regarding the empty cells, (89) is revised accordingly.

References

Campbell, D. T. and Boruch, R. F., "Making the case for randomized assignment to treatments by considering the alternatives: Six ways in which quasi-experimental evaluations in compensatory education tend to underestimate effects." To appear in Bennett, C.A. and Lumsdaine, A. (Eds.), Central Issues in Social Program Evaluation. New York: Academic Press, (in press).

Cox, D. R. 1970. The analysis of binary data. London: Methuen.

Gini, C. W. 1912. "Variability and mutability," contribution to the study of statistical distributions and relations, Studi Economico - Guiridici della R. Universita de Cagliari.

Goldberger, A.S. 1964. Econometric theory. New York: John Wiley and Sons.

Goodman, L.A. 1971. "The analysis of multidimensional contingency tables: stepwise procedures and direct estimation methods for building models for multiple classifications," Technometrics, 13, 33-61.

_____. 1972a. "A general model for the analysis of surveys," American Journal of Sociology, 77, 1035-1086.

_____. 1972b. "A modified multiple regression approach to the analysis of dichotomous variables," American Sociological Review, 37, 28-46.

_____. 1974. "The relationship between the modified and the more usual multiple regression approach to the analysis of dichotomous variables," Technical report No. 9, Department of Statistics, University of Chicago.

Goodman, L.A. and Kruskal, W.H. 1954. "Measures of association for cross classifications," Journal of the American Statistical Association, 49, 732-764.

Kendall, M.G., and Stuart, A. 1963. The advanced theory of statistics. Vol. 1, 2nd ed., London: Griffin.

Light, R.J. and Margolin, B.H. 1971. "An analysis of variance for categorical data," Journal of the American Statistical Association, 66, 534-544.

Magidson, J. 1975. "The multivariate analysis of qualitative variance: Analyzing the probability of an event as a function of observable variables," Ph.D. dissertation, Northwestern University. (Forthcoming)

Malinvaud, E. 1970. Statistical methods of econometrics, 2nd. ed., New York: American Elsevier.