7

LOGLINEAR MODELS FOR RECIPROCAL AND OTHER SIMULTANEOUS EFFECTS

*Robert D. Mare**
Christopher Winship†

This paper presents new models for simultaneous relationships among endogenous categorical variables. Previous investigators have argued that the loglinear/logit framework is insufficiently rich for the development of simultaneous equation models and that only models that postulate latent continuous variables (e.g. multivariate probit models) can represent simultaneous relationships among categorical variables. This paper shows that by using latent class methods, we can develop loglinear models for simultaneous effects that are analogous to linear models in simultaneous equation theory. These models, which are extensions of conventional loglinear and logit models for cross-classified data, are suitable when an independent variable is jointly determined with the dependent variable in a single-equation logit model or when there are reciprocal effects between two endogenous categorical variables. The models proposed here are extensions of recently developed

This research was supported by grants from the National Institute of Child Health and Human Development (HD-25748-01 and HD-25749-01), the National Science Foundation, and the Graduate School of the University of Wisconsin—Madison. Computations were performed using the facilities of the Center for Demography and Ecology of the University of Wisconsin—Madison, which are supported by the Center for Population Research of the NICHD (HD-5876). The authors are grateful to Charles Halaby, Yu Xie, Peter Marsden, and anonymous reviewers for helpful comments on an earlier draft of this paper.

*University of Wisconsin, Madison
†Northwestern University

loglinear models for missing and other partially observed data. They have several advantages over the multivariate probit approach, including avoidance of the assumption of latent multivariate normally distributed variables, a closer link between units of measurement and structural parameters, and relative ease of computation.

1. INTRODUCTION

1.1. *The Problem of Simultaneity*

A standard tool for the analysis of complex social phenomena is the structural equation model, which specifies the relationships between dependent variables and independent variables. This model is particularly valuable when several outcomes are jointly (simultaneously) determined, that is, when each endogenous variable depends on the other endogenous variables under investigation. For example, in the interaction between spouses, the behavior of one spouse may affect the behavior of the other spouse, and vice versa (e.g., Duncan 1974; Duncan and Duncan 1978). The behavior of each spouse is both an independent variable and a dependent variable in a model of reciprocal effects. In other instances the joint determination of endogenous variables is more subtle. For example, one may wish to examine the effects of participation in a job-training program on the probability of employment at a subsequent date. Ideally, program participation is an exogenous variable that affects a single endogenous variable; but in the absence of random assignment of persons to the program, the structural relationship between participation and employment may be obscured by systematic selection of individuals into (or out of) the program. Although employment status is the endogenous variable of primary interest, program participation is also endogenous and is jointly determined with employment (e.g., Heckman and Hotz 1989).

For continuous endogenous variables, simultaneous equation models are well-established extensions of the general linear model (e.g., Goldberger and Duncan 1973; Amemiya 1985; Duncan 1975). When one or more endogenous variables are discrete, however, more complex methods are used. Typically, simultaneous equation models for discrete endogenous variables are multivariate probit models (e.g., Mallar 1977; Heckman 1978; Muthén 1984). These simultaneous models rely critically on the assumption that discrete endogenous variables are realizations of latent *continuous* variables. That is, structural relations are specified in terms of continuous variables, thereby allowing interpretations that are similar to those from conventional structural equation models for observed continuous variables. In these models, the link between the latent continuous variables and the observed discrete variables is specified in auxiliary measurement equations, and estimation usually requires the assumption that conditional on the exogenous variables, the latent continuous variables follow a multivariate normal distribution.

1.2. *Simultaneous Equations: A Loglinear Approach*

An alternative strategy for analysis of simultaneous relations among discrete variables is to extend standard loglinear and logit models for categorical data (e.g., Bishop, Fienberg, and Holland 1975; Goodman 1978; Haberman 1978–79; Fienberg 1980). On the surface, this strategy seems attractive: (a) It expresses structural relations among variables in a way more closely tied to the way that variables are measured; (b) it avoids the analytic fiction that discrete endogenous variables always arise from latent continuous variables; (c) it avoids the assumption of (conditional) multivariate normality of endogenous variables; and (d) it avoids the computational burden that arises in probit models for polytomous outcomes or for more than two or three dichotomous outcomes (e.g. Daganzo 1979).

This strategy for simultaneous equation modeling for discrete endogenous variables, however, has *not* been followed. Indeed, many analysts have concluded that loglinear models cannot be used to analyze simultaneous relationships among endogenous variables.

Goodman (1973) presents methods for analysis of recursive causal systems using loglinear models but represents relationships between jointly determined variables by only their partial association. By this formulation, the reciprocal effects between pairs of endogenous variables are not identified, and the structural relations between variables are not distinguished from their partial associations.

Brier (1978) shows that elementary loglinear models for two endogenous variables imply two logit equations in which the reciprocal effects of the endogenous variables are equal. He concludes that "generally, reciprocal effects can never be separated in systems of logistic models. . . . The only techniques that allow simultaneous estimation of reciprocal effects involve the concept of latent continuous variables" (1978, pp. 124, 126).

In a widely used text, Fienberg states: "Can we set up nonrecursive systems of logit models for categorical variables, with properties resembling those of the nonrecursive systems of linear structural equations? The answer to this question is no" (1980, p. 134).

Heckman (1978, p. 950) asserts that "the loglinear model is not sufficiently rich in parameters to distinguish structural association among discrete random variables from purely statistical association among discrete random variables. The distinction between structural and statistical association is at the heart of simultaneous equation theory." Heckman argues that the error structure of the logit model is too restrictive to allow the model to represent simultaneous relationships. Specifically, because the logit model does not allow for correlated errors across equations, it is inappropriate for estimating simultaneous effects.

This paper shows that models for simultaneous effects among endogenous variables can in fact be specified and estimated by extending loglinear models for cross-classified data. In particular, we develop models that expand the standard loglinear model by incorporating partially observed variables. These variables are observed for some cases but are unobserved for others. The use of partially observed variables allows for a sufficiently rich parametric structure to model simultaneity within the framework of loglinear models. These simultaneity models do not rely on the assumption of latent continuous variables; nor do they make distributional assumptions beyond the usual multinomial sampling assumptions of the loglinear model. The models are analogous to standard simultaneous equation models for continuous variables in that they permit the separation of structural relations of variables from their statistical associations and, in models of simultaneity between two endogenous variables, the isolation of distinct reciprocal effects. At the same time, this approach to simultaneous equation modeling retains the conceptual and practical advantages of loglinear and logit approaches.

The models presented in this paper build on recently developed models for cross-classified data in which some variables are missing for some observations. Fay (1986), Little and Rubin (1987), Baker and Laird (1988), and Winship and Mare (1989) present loglinear models for tables with missing data, including data that are not missing at random. Haberman (1988) presents a general computational algorithm for estimating loglinear models on indirectly or partially observed contingency tables.

Section 2 of this paper describes data that we will use to illustrate our models and discusses alternative structural relationships that may be investigated with the data. Section 3 presents a model for a single structural relationship that is potentially confounded by simultaneity between the dependent variable and one of the regressors. This model is analogous to the dummy-endogenous-variable model that is based on extensions of multivariate probit analysis (Heckman 1978; Maddala 1983). Section 4 presents a model for reciprocal effects between two endogenous variables. This model is analogous to the simultaneous equation model that has been commonly applied in sociology (e.g., Duncan, Haller, and Portes 1968; Stolzenberg and Waite 1977; Marini 1984). We show how each of these two models can be formulated and estimated on a partially observed contingency table and present illustrative empirical results. Section 5 discusses the identifiability of the simultaneous equation models presented here. Section 6 discusses some limitations of the proposed models and problems for further research.

2. AN EXAMPLE

2.1. *Simultaneous Equation Models for Reciprocal Effects*

Table 1 cross classifies two-wave panel data on whether or not high school boys perceive themselves to be members of their school's leading crowd and whether their attitude toward the leading crowd is favorable or unfavorable. These data are from a study by Coleman (1961), in which high school students were interviewed in October 1957 and in May 1958. Membership is measured by a response to the question, "Are you a member of the leading crowd?" Attitude is measured by the respondent's agreement or disagreement with the statement, "If a fellow wants to be part of the leading crowd around here, he sometimes has to go against his principles" (Coleman 1964, p. 168). These data have been analyzed by Coleman (1964), Good-

TABLE 1

Cross-Classification of Panel Data on Membership In and Attitude Toward the Leading Crowd in High School

			Wave 2			
	Membership (C)		Member		Nonmember	
	Attitude (D)		Favorable	Unfavorable	Favorable	Unfavorable
	Membership (A)	Attitude (B)				
Wave 1	Member	Favorable	458	110	140	49
	Member	Unfavorable	171	56	182	87
	Nonmember	Favorable	184	531	75	281
	Nonmember	Unfavorable	85	338	97	554

Source. Coleman 1964.

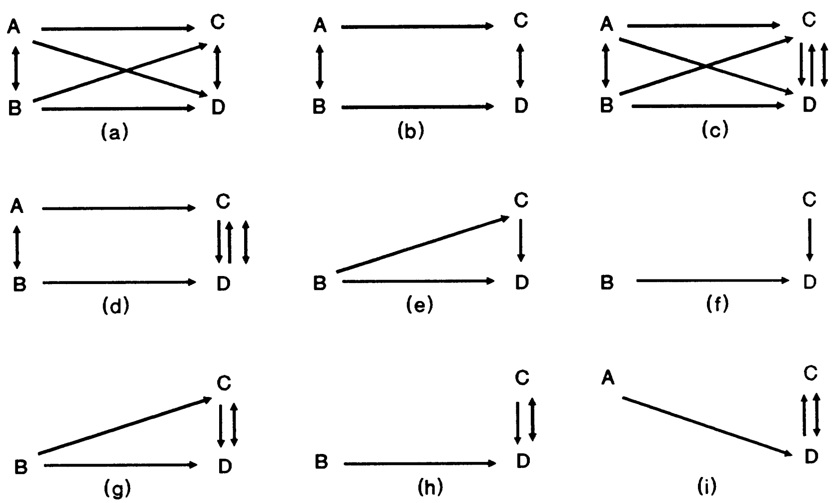

FIGURE 1. Models for two endogenous variables.

man (1973, 1974), Fienberg (1980), Duncan (1985), and Haberman (1988).

Although many models can be applied to these data, we emphasize those that represent simultaneous relationships between membership in and attitude toward the leading crowd. Figure 1 represents alternative models for the leading-crowd data, where A, B, C, and D denote membership in the leading crowd at wave 1, attitude toward the leading crowd at wave 1, membership at wave 2, and attitude at wave 2, respectively. In model *a,* both membership and attitude at wave 1 affect both membership and attitude at wave 2. Within each wave, membership and attitude are associated, but the direction of the effect is unspecified. Thus, although model *a* allows for mutual causation of membership and attitude *over time,* it does not represent their simultaneous effects on each other. This model can be estimated as a loglinear model with the terms AB, AC, AD, BC, BD, and CD. Model *b* is similar to model *a,* except that it omits the cross-lagged associations between membership and attitude. It can be estimated as a simple loglinear model with terms AB, AC, BD, and CD.

Unlike models *a* and *b,* models *c* and *d* include the reciprocal effects of membership and attitude on each other at wave 2. These effects are represented by the single-headed arrows that connect *C* and *D*. In addition, models *c* and *d* include residual association between *C* and *D* that remains once the dependence of *C* and *D* on each other and on *A* and *B* are taken into account. This residual association is represented by the double-headed arrows connecting *C* and *D* in models *c* and *d*. Whereas model *c* also allows for cross-lagged effects, model *d* omits these effects. These models could be specified as pairs of simple logit models. For example, model *c* could be represented as a loglinear model or as two logit models for the probability that a boy is in the leading crowd (given *A*, *B*, and *D*) and that he has a favorable attitude toward the leading crowd (given *A*, *C*, and *D*). By this formulation, however, the partial effects of *D* on *C* and of *C* on *D* are necessarily equal and thus are no more informative than the *CD* partial association in model *a* (Brier 1978). The same result holds for model *d*. The logit models, moreover, cannot distinguish the reciprocal effects between *C* and *D* from the association that remains once their dependence on *A*, *B*, and each other is taken into account. In contrast, if *C* and *D* were continuous variables, a conventional simultaneous equation model (or the analogous multivariate probit model [Heckman 1978]) for model *d* would yield distinct estimates of the effects of *C* on *D* and *D* on *C*. In this approach, variables *B* and *A* are instrumental variables for *C* and *D*, respectively. Moreover, by this approach it is possible to identify both the reciprocal effects of *C* and *D* and also the residual correlation between *C* and *D* net of their dependence on *A*, *B*, and each other. Model *c*, however, is not identified in a conventional simultaneous equation approach.

Although conventional loglinear and logit models cannot isolate the reciprocal effects of *C* and *D* in model *d*, extensions of the loglinear model can. These extensions, which we shall term *structural loglinear models*, also enable one to distinguish the two reciprocal effects of *C* and *D* from their remaining partial association once causal relationships are taken into account. Whereas extensions of loglinear models can contain all of the parameters for model *d*, they cannot do so for model *c*. These new models, which are presented in section 4, enable one to isolate all of the effects that can be obtained in conventional simultaneous equation models. Before we present these models, however, we consider simpler simultaneous equation models.

2.2. *Simultaneous Equation Models for One Equation with Endogenous Regressor*

A simpler model is required when one has a single dependent variable but one or more independent variables can be jointly determined with the dependent variable. This problem arises in the study of job-program effects on employment mentioned above, but it can also be illustrated with the leading-crowd data in Table 1. Suppose that variable A is not observed, that instead of Table 1 we have a 2^3 table of attitude at wave 1 (B) by membership at wave 2 (C) by attitude at wave 2 (D), and that we are mainly interested in the effects of membership on attitude. Models *e–i* in Figure 1 apply to 2^3 tables. Models *e* and *f* are simple models for the effects of B and C on D and can be specified and estimated as elementary loglinear models. Models *g, h,* and *i,* in contrast, represent joint determination of C and D. The single-headed arrow between C and D denotes the effect of C on D. The double-headed arrow represents additional association between the variables not due to the effect of C on D. This association may occur because D affects C as in models *c* and *d;* because other variables, not included in the table, affect both C and D; or because C and D have measurement errors that are correlated.

As in the four-variable model with reciprocal effects, the parameters of models *g, h,* and *i* cannot be retrieved from conventional loglinear models. In contrast, if C and D were continuous variables, conventional simultaneous equation methods could be used to estimate model *i,* in which A serves as an instrumental variable for D. In the conventional simultaneous equation approach, models *g* and *h* are not identified. The structural loglinear models proposed in this paper, however, can isolate the parameters of models *h* and *i,* but not model *g*. Models *h* and *i* are analogous to the dummy-endogenous-variable model proposed by Heckman (1978) within the multivariate probit framework. We discuss this model in the next section.

3. STRUCTURAL LOGLINEAR AND LOGIT MODELS FOR A SINGLE EQUATION WITH SIMULTANEITY

In this section we describe single-equation models for categorical variables in which one independent variable is jointly determined with the dependent variable. We begin by outlining a general approach to distinguishing between structural and spurious association in categorical variables. Then we apply this approach to the simultaneous equation model.

3.1. *Structural Effects and Partial Observability*

In nonexperimental data, observations on the joint distributions of dependent and independent variables almost always confound the structural relationships between variables with spurious association. Spurious associations arise because unmeasured variables may affect both the independent and the dependent variables; because respondents may be selected (or select themselves) into categories of the independent variable on the basis of their expected outcomes on the dependent variable, creating "feedback" between the dependent and independent variables; or because of correlated errors of measurement in the dependent and independent variables. Spurious association between observed variables arises because observations are not randomly assigned to levels of the independent variable.[1]

Another way of viewing spuriousness is that it results from *incomplete observation* on the dependent variable. For each respondent, we observe the dependent variable for a single level of an independent variable but do not observe what the dependent variable would have been had the respondent been assigned to other levels of that independent variable (Rubin 1978). In most nonexperimental studies, one must infer effects from differences in the dependent variable across levels of an independent variable that are observed for different respondents; that is, one must infer from comparisons *between persons* who are not necessarily identical on unmeasured variables. If, on the other hand, one could observe *distinct* dependent variables for each individual for each value of an independent variable, then one could make much stronger causal inferences from comparisons *within persons* who are, by definition, identical on unmeasured variables across levels of the independent variables. In the absence of repeated observations on the same respondent across levels of an independent variable, one can nonetheless *model* the par-

[1]For concreteness our discussion refers to "respondents," "persons," and "individuals" throughout. Obviously, our models apply to other units of analysis as well.

tially observed data. Although one cannot estimate effects for each person, one can estimate the average effect of a variable across persons. Rubin (1978) provides a general model for causal inference in which outcomes on the dependent variable on unobserved levels of the independent variables are regarded as "missing data." Winship and Mare (1989) show how loglinear models for missing data enable one to make inferences about respondents' behavior when it is not observed. Before extending these models to simultaneous equations, we describe their formulation for the more elementary case of a single endogenous variable. In practice, if the dependent variable is the only endogenous variable (that is, if all independent variables are exogenous), then elementary loglinear and logit models and the models for partially observed data presented here provide identical estimates of effects. We begin with this case, however, to illustrate the approach in its simplest form, before going on to the more complex case of jointly determined variables.

Suppose that we have a single endogenous variable D and a single exogenous variable B, each of which takes the values 1 or -1. For example, B and D may denote attitude toward the leading crowd at waves 1 and 2, respectively, and thus be observed in a collapsed version of Table 1. The observed data, therefore, are a 2×2 frequency table. To assess the effect of B on D, we can use an elementary logit model,

$$\text{logit}[p(D = 1 \mid B)] = \beta + \beta_j^B, \tag{1}$$

where the subscript j indexes levels of B and $\Sigma_j \beta_j^B = 0$.

An alternative formulation recognizes that for each respondent, we observe D for only one level of B and we do not observe D for the level of B that the respondent did not experience. Define two additional variables, D_1 and D_0, that denote respondents' values on D when B equals 1 and -1, respectively. For respondents for whom $B = 1$, $D_1 = D$ and D_0 is unobserved; for respondents for whom $B = -1$, $D_0 = D$ and D_1 is unobserved. Here, B is an indicator for whether D_1 or D_0 is observed. In the language of experiments it indicates the treatment to which an individual is assigned. Since B is exogenous, it is independent of D_1 and D_0. Whether $B = 1$ or -1 is not related to the outcome on D_1 and D_0. Assignment to levels of B is at "random."

We can represent the relationship between our three variables in a partially observed 2^3 table with dimensions B, D_1, and D_0. This

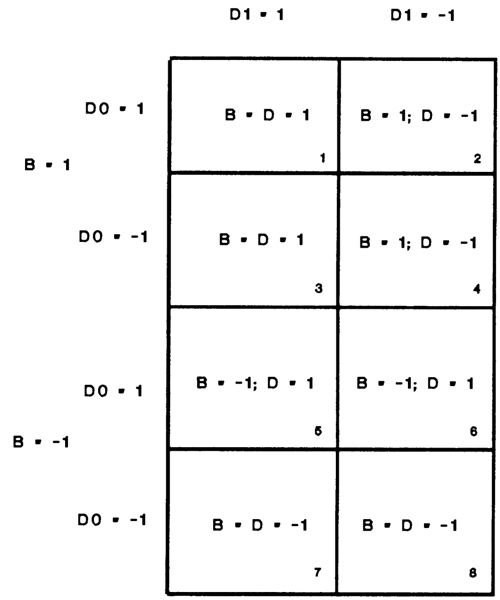

FIGURE 2. Expanded form of table with one exogenous and one endogenous variable.

table is illustrated in Figure 2, which shows the mapping between the observed data on B and D and the partially observed relations among B, D_1, and D_0. This table can be modeled as a latent class/loglinear model using methods described by Winship and Mare (1989) and below. The loglinear model that is equivalent to the logit model (1) is the model of one-way effects (independence): BD_1D_0. This equivalence is explained below. The assumption that B is independent of D_1 and D_0 (that is, the assumption that B is exogenous) is essential to the identification of the model. The D_1D_0 interaction is not identified. For convenience it is set to zero.

To understand the relationship between the above loglinear model and the logit model (1), consider two logit equations that the above loglinear model implies:

$$\text{logit}[p(D_1 = 1)] = \beta^{D1}, \tag{2}$$

$$\text{logit}[p(D_0 = 1)] = \beta^{D0}. \tag{3}$$

The effect of B on D is the difference in levels between D_1 and D_0, that is, $\beta^{D1} - \beta^{D0}$. The average level of D in the population is $(\beta^{D1} + \beta^{D0})/2$. Equations (2) and (3) cannot be estimated using standard methods, since D_1 and D_0 are not fully observed. That B is independent of D_1 and D_0, however, implies that

$$\text{logit}[p(D_1 = 1) \mid B = 1] = \beta^{D1}, \tag{4}$$

$$\text{logit}[p(D_0 = 1) \mid B = 0] = \beta^{D0}. \tag{5}$$

These two equations can be estimated using standard methods, since D_1 and D_0 are fully observed for the specified value of B in each question. The estimation of (4) and (5) is equivalent to the estimation of (1). Combining the dependent variables in equations (4) and (5) gives the dependent variable in equation (1). The relationships between the parameters in the three equations are as follows: $\beta = (\beta^{D1} + \beta^{D0})/2$, $\beta_1^B = (\beta^{D1} - \beta^{D0})/2$, and $\beta_0^B = (\beta^{D0} - \beta^{D1})/2$.

If B is exogenous, estimation of equation (1) and estimation of a loglinear model with only the one-way effects in Figure 2 yield identical results. Nothing more about the relationship between B and D is learned by using the latent class approach. One can always postulate a separate dependent variable for each combination of levels of the independent variables, but this is unnecessary whenever one assumes that all of the independent variables are exogenous. In this case, by assumption, the effects that derive from the between-respondent comparisons on the dependent variable across levels of the independent variable are satisfactory.

As discussed in detail below, when an independent variable is jointly determined with the dependent variable, it becomes fruitful to distinguish between respondents' observed outcomes on endogenous variables and those that they would have obtained if their values on the independent variables were different from those observed. It is possible to model these relationships by using latent class loglinear models that are more complicated than the one-way effects model above. This is the key to our approach.

3.2. *General Form of the Single-Equation Model with Simultaneity*

Our approach is applicable to models *h* and *i* in Figure 1, although our initial discussion will be confined to model *h*. Model *i*, which can be estimated by a similar approach, is discussed in section 5.2.

To separate the structural effect of C on D from the statistical association between these variables, we use a model in which each respondent has two outcomes on variable C and two outcomes on variable D. That is, we distinguish between respondents' actual responses on D and their hypothetical responses if they had a value of C different from the one that we observe; and we distinguish between respondents' actual responses on C and their hypothetical responses if they had a value of D different from the one that we observe. Let B, C, and D take on values of 1 or -1. Now define four new variables: D_1 and D_0, which denote respondents' outcomes for alternative values of C, and C_1 and C_0, which denote respondents' outcomes for alternative values of D.

D_1, D_0, C_1, and C_0 are partially observed variables inasmuch as whether we observe them depends on the particular value of C and D that we observe. The logical relations among C, D, C_1, C_0, D_1, and D_0 are as follows: (a) if $C = D = 1$, then $C_1 = 1$, $D_1 = 1$, and C_0 and D_0 are unobserved; (b) if $C = 1$ and $D = -1$, then $C_0 = 1$, $D_1 = -1$, and C_1 and D_0 are unobserved; (c) if $C = -1$ and $D = 1$, then $C_1 = -1$, $D_0 = 1$, and C_0 and D_1 are unobserved; (d) if $C = D = -1$, then $C_0 = -1$, $D_0 = -1$, and C_1 and D_1 are unobserved. This implies that $C = [(D + 1)C_1 + (1 - D)C_0]/2$ and that $D = [(C + 1)D_1 + (1 - C)D_0]/2$.

Despite the partial observability of C_1, C_0, D_1, and D_0, a general model for the dependence of D on C is

$$p(D_1 = 1 \mid B,C_1,C_0) = F_1(B,C_1,C_0), \tag{6}$$

$$p(D_0 = 1 \mid B,C_1,C_0) = F_0(B,C_1,C_0), \tag{7}$$

where F_1 and F_0 denote functions that will be specified more fully below. Equations (6) and (7) are a structural model for D inasmuch as they represent the effects of B and C on D apart from the actual sample values of B, C, and D. As we will show more explicitly below, the model separates the structural effect of C on D from the sample association between C and D. In these respects, this model is analogous to the endogenous switching regression model (Maddala 1983; Mare and Winship 1988), which separates the structural effects of a categorical variable from systematic, sample-specific selection of observations into levels of the categorical variable that are potentially correlated with the dependent variable.

		D1 = 1		D1 = -1	
		D0 = 1	D0 = -1	D0 = 1	D0 = -1
C1 = 1	C0 = 1	C = D = 1 1	C = D = 1 2	C = 1; D = -1 3	C = 1; D = -1 4
	C0 = -1	C = D = 1 5	C = D = 1 C = D = -1 6	7	C = D = -1 8
C1 = -1	C0 = 1	C = -1; D = 1 9	10	C = 1; D = -1 C = -1; D = 1 11	C = 1; D = -1 12
	C0 = -1	C = -1; D = 1 13	C = D = -1 14	C = -1; D = 1 15	C = D = -1 16

FIGURE 3. Expanded form of table with two endogenous variables (conditional on exogenous variables).

3.3. *Structural Loglinear Model for the Expanded Table*

We specify and estimate structural loglinear models for an expanded, partially observed contingency table with dimensions B, C_1, C_0, D_1, and D_0.[2] Figure 3 illustrates the expanded table for the jth category of B. We observe none of the individual cells of the expanded table. Instead, each observed combination of C and D for a given level of B can fall into four cells in the expanded table. For example, if $C = -1$ and $D = 1$ for an observation, then it is potentially a member of cell 9, 11, 13, or 15 of the expanded table. Two

[2]An alternative approach to the single-equation model with an endogenous independent variable is to specify a model for an expanded table with dimensions B, C, D_1, and D_0, that is, to consider alternative outcomes on D for different values of C but to assume that $C_1 = C_0 = C$. This approach yields a model that is similar but not identical to the one presented here. The relationship between these two models is a subject for further research. The model presented here generalizes more easily to the simultaneous equation model presented in section 4.

cells of the expanded table are logically impossible (7 and 10), and for two other cells of the expanded table (6 and 11), two distinct cells of the observed table provide information.

Since B has two categories, a general loglinear model for the expanded table would have 28 parameters ($32 - 4 = 28$, because of the four logically impossible cells), which would not be estimable because the observed table has only eight cells. Thus, we use a more restricted model. In keeping with model h in Figure 1, we assume that B does not directly affect C. This is analogous to the assumption in standard simultaneous equation theory that B is an instrument for C. Thus, all interactions that include the BC_1 and BC_0 terms are set to zero. In addition, we assume that all interactions that include the D_1D_0 and C_1C_0 terms are zero. This implies that conditional on B, D_1 and D_0 are independent and C_1 and C_0 are independent.[3] Assuming that observations are obtained under multinomial sampling, a restricted loglinear model is

$$\log p_{jklmn} = \lambda + \lambda_j^{B} + \lambda_k^{C1} + \lambda_l^{C0} + \lambda_m^{D1} + \lambda_n^{D0} + \lambda_{jm}^{BD1} + \lambda_{jn}^{BD0} + \lambda_{km}^{C1D0} + \lambda_{kn}^{C1D0} + \lambda_{lm}^{C0D1} + \lambda_{ln}^{C0D0}, \tag{8}$$

where p_{jklmn} denotes the probability that an individual falls into the jth category of B ($j = -1,1$), the kth category of C_1 ($k = -1,1$), the lth category of C_0 ($l = -1,1$), the mth category of D_1 ($m = -1,1$), and the nth category of D_0; the intercept λ is determined by the constraint that $\Sigma_{jklmn}P_{jklmn} = 1$; the remaining λ's are parameters; and $\Sigma_j\lambda_j^{B} = \Sigma_k\lambda_k^{C1} = \Sigma_l\lambda_l^{C0} = \Sigma_m\lambda_m^{D1} = \Sigma_n\lambda_n^{D0} = \Sigma_j\lambda_{jm}^{BD1} = \Sigma_m\lambda_{jm}^{BD1} = \Sigma_j\lambda_{jn}^{BD0} = \Sigma_n\lambda_{jn}^{BD0} = \Sigma_k\lambda_{km}^{C1D1} = S_m\lambda_{km}^{C1D1} = \Sigma_k\lambda_{kn}^{C1D0} = \Sigma_n\lambda_{kn}^{C1D0} = \Sigma_l\lambda_{lm}^{C0D1} = \Sigma_m\lambda_{lm}^{C0D1} = \Sigma_l\lambda_{ln}^{C0D0} = \Sigma_n\lambda_{ln}^{C0D0} = 0$.

Model (8) is not identified inasmuch as it contains 12 parameters, four more than the number of observed cells. However, a more restrictive version of (8), which is identified, captures the essential features of model h in Figure 1. In particular, we assume that D does not affect C, i.e., that $\lambda_k^{C1} = \lambda_l^{C0}$. We also assume that $\lambda_{km}^{C1D1} = \lambda_{kn}^{C1D0}$

[3]Unlike the interactions in the endogenous switching model based on linear and probit equations (Maddala 1983; Mare and Winship 1988), these interactions are identified in structural loglinear models under certain conditions. The assumption that they are zero is usually not realistic. In the examples presented in this paper, however, the models fit well when these interactions are omitted. Footnotes 4 and 6 report results for models in which these interactions are included. For the most part, the estimates do not change, although standard errors are larger.

$= \lambda_{lm}^{C0D1} = \lambda_{ln}^{C0D0}$, a restriction explained below. Equation (8) includes distinct effects of B on D_1 and D_0. This amounts to a three-way interaction between B, C, and D. Model h in Figure 1, however, assumes a single effect of B on D (irrespective of the value of C). Thus, we assume that $\lambda_{jm}^{BD1} = \lambda_{jn}^{BD0}$, i.e., that B does not interact with C in affecting D.

We impose the above constraints by using two new variables, C^* and D^*. These variables help to simplify the equations below and to provide a direct connection to the structure of the design matrices discussed below. C^* is indexed by s and has three levels, 1, 0, -1. The index s of C^* is the sum of the indexes of C_1 and C_0, k and l respectively; that is, $s = (k + l)/2$. The parameters for C^*, for example λ^{C^*}, are the sums of the parameters for C_1 and C_0, for example $\lambda_s^{C^*} = \lambda_k^{C1} + \lambda_l^{C0} = 2\lambda_k^{C1}$. D^* is defined analogously for the variables D_1 and D_0. D^* has index t, where $t = (n + m)/2$, and $\lambda_t^{D^*} = \lambda_m^{D1} + \lambda_n^{D0} = 2\lambda_m^{D1}$. When C^* and D^* are used in higher-way effects, the above relations generalize. Thus, including the term $\lambda_{jt}^{BD^*}$ implies that $\lambda_{jm}^{BD1} = \lambda_{jn}^{BD0}$, with $\lambda_{jt}^{BD^*} = \lambda_{jm}^{BD1} + \lambda_{jn}^{BD0} = 2\lambda_{jm}^{BD1}$. Using C^* and D^* implies that the terms involving C_1, C_0, D_1, and D_0 are implicitly included in an expression and that they have been constrained to be equal in the way noted above: $\lambda_{st}^{C^*D^*} = \lambda_{km}^{C1D1} + \lambda_{kn}^{C1D0} + \lambda_{lm}^{C0D1} + \lambda_{ln}^{C0D0} = 4\lambda_{km}^{C1D1}$. This term represents the residual association of C and D net of the structural relationship(s) between the two variables. Inclusion of this term implies that all terms to which it sums are included and constrained to be equal. As we show below, the definitions of C^* and D^* are consistent with the way design effects are defined when equality constraints are imposed.

Under these assumptions and notation, the model becomes

$$\log p_{jklmn} = \lambda + \lambda_j^B + \lambda_{kl}^{C^*} + \lambda_m^{D1} + \lambda_n^{D0} + \lambda_{jmn}^{BD^*} + \lambda_{st}^{C^*D^*}. \quad (9)$$

Model (9) is a structural loglinear model for the effects of C on D, taking account of the residual association between these two variables. The *structural* coefficient for the effect of C on D is $2(\lambda_m^{D1} - \lambda_n^{D0})$. To see this, note that λ_m^{D1} denotes the average level of D when $C = 1$, whereas λ_n^{D0} denotes the average level of D when $C = -1$. Thus, the effect of C is proportional to the difference between these two parameters. The *partial association* between C and D is $2\lambda_{st}^{C^*D^*}$. The latter parameter corresponds to the double-headed arrow in model h of Figure 1 and measures the tendency of respondents for whom $C_1 =$

1 and $C_0 = 1$ to be more (or less) likely to have $D_1 = 1$ and $D_0 = 1$. If the model contained separate parameters $\lambda_{st}^{C^*D1}$ and $\lambda_{st}^{C^*D0}$ (instead of $\lambda_{st}^{C^*D^*}$), then these two parameters would correspond to the effects of nonrandom selection from the levels of D_1 and D_0 into categories of C and would be analogous to correlated disturbances in endogenous switching regression models (e.g., Maddala 1983; Mare and Winship 1988). In this model, however, we estimate only a single parameter for nonrandom selection, yielding a model that is analogous to the dummy endogenous variable model (Heckman 1978, p. 938).

3.4. *Logit Form of the Model*

To show the link between the structural loglinear model and the general model presented in section 3.2, we write (9) as two logit models, one for D_1 and one for D_0. That is,

$$\text{logit}[p(D_1 = 1 \mid B,C_1,C_0)] = \beta^{D1} + \beta_j^{BD^*} + \beta^{C^*D^*}, \quad (10)$$

$$\text{logit}[p(D_0 = 1 \mid B,C_1,C_0)] = \beta^{D0} + \beta_j^{BD^*} + \beta^{C^*D^*}, \quad (11)$$

where $\beta^{D1} = 2\lambda_m^{D1}$, $\beta^{D0} = 2\lambda_n^{D0}$, $\beta_j^{BD^*} = 2\lambda_{jt}^{BD^*}$, and $\beta^{C^*D^*} = 2\lambda_{st}^{C^*D^*}$. Equations (10) and (11) correspond to the general model of (6) and (7), but we can also write the model as a single equation:

$$\text{logit}[p(D_q = 1 \mid B,C_1,C_0)] = \beta^{D} + \beta_q^{DC} + \beta_j^{BD^*} + \beta^{C^*D^*}, \quad (12)$$

where D_q is a variable taking the value 1 or -1 for the qth category of C ($q = 0,1$); β^D is the grand mean of the logit of the probability that $D_q = 1$ [$\beta^D = (\beta^{D1} + \beta^{D0})/2$]; β_q^{DC} is the structural effect of C on D (that is, $\beta_1^{DC} = \beta^{D1} - \beta^{D0}$, and $\beta_0^{DC} = \beta^{D0} - \beta^{D1}$); and the remaining notation is as defined above. Model (12), therefore, contains parameters for the structural effects of B and C on D, as well as for the association between C and D that remains once the dependence of D on C and B is taken into account.

3.5. *Estimation with an Empirical Illustration*

We can obtain the parameters of (12) from the corresponding loglinear model (9), which we estimate by treating the unobserved cells in Figure 3 as "missing" data (Winship and Mare 1989). We get maximum likelihood estimates of expected cell frequencies and parameters of (9) using Haberman's (1988) DNEWTON program for loglinear models estimated from indirectly observed contingency tables. Haberman (1988) describes the computational algorithm, and Winship and Mare (1989) provide an example of its application to simple contingency tables with missing data. The estimation procedure requires that one specify (a) a design matrix for the relationship between the parameters to be estimated and the expected frequencies of the expanded table and (b) a mapping between the observed cell frequencies and the cells of the expanded table.

For model (9) applied to the three-variable *BCD* version of the leading-crowd data in Table 1, we present the design matrix and cell mapping in Table 2. The rows of Table 2 refer to cells in the expanded table. The columns of the design matrix labeled *Model Terms* correspond to parameters in equation (9). The columns for B, C^*, D_0, and D_1 are contrast-coded indicator variables. These variables are d^B, d^{C^*}, d^{D0}, and d^{D1}, respectively. (The variable d^{C^*} is the average of the separate indicator variables for C_1 and C_0, say d^{C1} and d^{C0}.) Then the columns for BD^* and C^*D^* are $d^B(d^{D0} + d^{D1})/2$ and $d^{C^*}(d^{D0} + d^{D1})/2$, respectively. This coding imposes the restrictions $\lambda_{jm}^{BD1} = \lambda_{jn}^{BD0}$, $\lambda_k^{C1} = \lambda_l^{C0}$, and $\lambda_{km}^{C1D1} = \lambda_{kn}^{C1D0} = \lambda_{lm}^{C0D1} = \lambda_{ln}^{C0D0}$ on (8) above.

The first column shows how the observed frequencies are mapped into 28 of the cells in the expanded table. The four cells corresponding to the two levels of cells 7 and 10 in Figure 3 are not included in this list because they are logically impossible. Thus, their frequencies are constrained to be zero.

As noted above there are potentially 28 parameters associated with the model for the expanded table in Figure 3. But Table 2 includes 32 rows because cells 6 and 7 in Figure 3 receive observations under two outcomes for each of the two levels of B. We have listed the contribution of each outcome separately. If we had combined the contributions, which would lead to the same results, Table 2 would have 28 rows. The last column of the table contains expected frequencies for a model to be discussed below.

The top panel of Table 3 presents likelihood-ratio chi-square (G^2) statistics for model (9) plus several simpler models fit to the *BCD* version of Table 1. Model I includes the structural effect of C on D (as implied by the terms D_0 and D_1), but no partial association between these variables. Model II includes the partial association (as implied by the term C^*D^*), but no structural effect. Model III includes both the structural effect and the partial association. Models I

TABLE 2
Design Matrix, Cell Mapping, and Expected Frequencies for Single-Equation Model with Endogenous Regressor for Collapsed Version of Leading-Crowd Data

Observed Frequency	C	D	Model Terms B	C^*	D_1	D_0	BD^*	C^*D^*	Expected Frequencies (Expanded Table) Model III
642	1	1	1	1	1	1	1	2	338.248
642	1	1	1	1	1	−1	0	0	34.175
642	1	1	1	0	1	1	1	0	209.986
642	1	1	1	0	1	−1	0	0	57.722
256	1	1	−1	1	1	1	−1	2	105.465
256	1	1	−1	1	1	−1	0	0	32.310
256	1	1	−1	0	1	1	−1	0	65.473
256	1	1	−1	0	1	−1	0	0	54.572
215	1	−1	1	1	−1	1	0	0	66.609
215	1	−1	1	1	−1	−1	−1	−2	6.730
215	1	−1	1	0	−1	1	0	0	112.503
215	1	−1	1	0	−1	−1	−1	0	30.925
279	1	−1	−1	1	−1	1	0	0	62.974
279	1	−1	−1	1	−1	−1	1	−2	19.293
279	1	−1	−1	0	−1	1	0	0	106.363
279	1	−1	−1	0	−1	−1	1	0	88.654
641	−1	1	1	0	1	1	1	0	209.986
641	−1	1	1	0	−1	1	0	0	112.503
641	−1	1	1	−1	1	1	1	−2	130.360
641	−1	1	1	−1	−1	1	0	0	190.018
394	−1	1	−1	0	1	1	−1	0	65.473
394	−1	1	−1	0	−1	1	0	0	106.363
394	−1	1	−1	−1	1	1	−1	−2	40.646
394	−1	1	−1	−1	−1	1	0	0	179.648
330	−1	−1	1	0	1	−1	0	0	57.722
330	−1	−1	1	0	−1	−1	−1	0	30.925
330	−1	−1	1	−1	1	−1	0	0	97.493
330	−1	−1	1	−1	−1	−1	−1	2	142.110
641	−1	−1	−1	0	1	−1	0	0	54.572
641	−1	−1	−1	0	−1	−1	1	0	88.654
641	−1	−1	−1	−1	1	−1	0	0	92.172
641	−1	−1	−1	−1	−1	−1	1	2	407.386

TABLE 3
Likelihood-Ratio Chi-Square Statistics for Three-Variable Models for Leading-Crowd Data

Model	G^2	*df*
Structural loglinear models		
I. B, C^*, D_1, D_0, BD	57.64	2
II. $B, C^*, D^*, BD^*, C^*D^*$	14.02	2
III. $B, C^*, D_1, D_0, BD^*, C^*D^*$	0.07	1
Elementary loglinear models		
IV. BD, CD	32.44	2
V. BC, CD	263.46	2
VI. BC, BD	31.18	2
VII BC, BD, CD	0.04	1

TABLE 4
Parameter Estimates in Logit Form of Structural Loglinear Models for the Three-Way Table

	Model I		Model II		Model III	
Variable[a]	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Intercept	.297	.037	.364	.040	.403	.090
$D_1 - D_0$	.416	.075			−.667	.218
B	1.162	.072	1.192	.070	1.110	.076
C^*D^*			1.080	.124	2.000	.324

[a] *B* is attitude toward the leading crowd in wave 1; *C* is membership in the leading crowd in wave 2; and *D* is attitude toward the leading crowd in wave 2.

and II are both nested within III, and as the G^2 statistics indicate, both the structural and the partial associations are statistically significant and are needed to provide an adequate model for these data.

Table 4 presents logit parameter estimates for these models. As indicated by the parameter estimates for *B* in all three models, the odds of a favorable attitude toward the leading crowd in wave 2 are higher for persons having a favorable attitude in wave 1 than for persons having an unfavorable attitude. The estimated effect of membership in the leading crowd in wave 2 (C), however, depends on which model is estimated. Models I and II imply that members of the leading crowd in wave 2 have more favorable attitudes toward the leading crowd in wave 2, whereas model III reveals a more complex relationship. Membership in the leading crowd *reduces* the probability of a favorable attitude in wave 2, but the partial association of membership and attitude is positive. This suggests that although belonging to the leading crowd makes one less likely to view the leading crowd favorably, unmeasured common causes of membership and attitude induce a positive association between these two variables. Although this example is simple, it illustrates that one can distinguish structural from residual associations using our methods.[4]

The lower panel of Table 3 reports G^2 statistics for several elementary loglinear models. The fit of the model of no three-way interaction, model VII, is close but not identical to that of model III. Whereas model VII includes all two-way interactions, model III uses the two-way interaction of B and C to identify the two parameters governing the relationship between C and D. Although models III and VII fit the *observed* data similarly, they yield different expected frequencies for the *expanded* table. Unlike the expected frequencies for model VII, the expected frequencies for model III vary across cells of the expanded table that correspond to the same cells of the observed table. The choice between models III and VII should be determined in part by the assumption that is made about the structural relationship between the variables, that is, whether B directly affects C. This choice is analogous to the problem of making assumptions about instrumental variables in standard simultaneous equation models.

4. STRUCTURAL LOGLINEAR AND LOGIT MODELS FOR RECIPROCAL EFFECTS

4.1. *General Form of the Model*

We now extend the approach described in section 3 to the model for reciprocal effects. This corresponds to model *d* in Figure 1 and will be illustrated with the four-way table for the leading-crowd data. Our model again views each respondent as having *two pairs* of outcomes, a pair for each of variables C and D. Defining all notation as above, let A be an additional exogenous variable, taking values of

[4]We also estimated model III allowing for a D_1D_0 association. Under this model, $G^2 = 0$ (0 *df*). The estimated association parameter is -0.376 (SE = 1.212), other parameter estimates for the model change somewhat, and their standard errors are larger than when the D_1D_0 association is omitted. The parameter for $D_1 - D_0$ becomes -0.422 (SE = 0.338).

1 or -1, that affects C but not D. A general model for the mutual dependence of C and D is

$$p(C_1 = 1 \mid A,B,D = 1) = G_1(A,D_0,D_1), \tag{13}$$

$$p(C_0 = 1 \mid A,B,D = -1) = G_2(A,D_0,D_1), \tag{14}$$

$$p(D_1 = 1 \mid A,B,C = 1) = G_3(B,C_0,C_1), \tag{15}$$

$$p(D_0 = 1 \mid A,B,C = -1) = G_4(B,C_0,C_1), \tag{16}$$

where the G_i denote functions that will be specified more fully below. Equations (13)–(16) are a structural model for C and D and include not only the effects of A and D on C and of B and C on D, but also the residual association between C and D. These effects are described below.

4.2. *Structural Loglinear Model for the Expanded Table*

We specify and estimate a loglinear model for an expanded, partially observable table with dimensions A, B, C_1, C_0, D_1, and D_0. The full expanded table has $2^6 = 64$ cells. Figure 3 now shows the cross-classification of C_1, C_0, D_1, and D_0 given a combination of values of A and B ($A = i$, $B = j$). The relationships between the observed values of C and D and the cells of the expanded table are the same as for the single-equation model.

A general loglinear model for the expanded table has 56 parameters (since there are eight structural zeros due to cells 7 and 10 in Figure 3), whereas the observed table has only 16 cells. Thus, we impose a number of restrictions. In keeping with model *d* of Figure 1, we assume that partial associations between A and D_0, A and D_1, B and C_0, B and C_1, and higher-way interactions involving these pairs of terms are zero. These are analogous to instrumental-variable assumptions. As before, we assume that all interactions that include the C_1C_0 or D_1D_0 terms are zero. Additional discussion of these two assumptions is given in section 6 below. Assuming that observations are obtained under multinomial sampling, a restricted loglinear model is

$$\begin{aligned}\log P_{ijklmn} = {} & \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^{C1} + \lambda_l^{C0} + \lambda_m^{D1} + \lambda_n^{D0} + \lambda_{ij}^{AB} \\ & + \lambda_{ik}^{AC1} + \lambda_{il}^{AC0} + \lambda_{jm}^{BD1} + \lambda_{jn}^{BD0} + \lambda_{km}^{C1D1} + \lambda_{kn}^{C1D0} + \lambda_{lm}^{C0D1} \\ & + \lambda_{ln}^{C0D0},\end{aligned} \tag{17}$$

where P_{ijklmn} denotes the probability that an individual falls into the ith category of A ($i = -1,1$), the jth category of B ($j = -1,1$), the kth category of C_1 ($k = -1,1$), the lth category of C_0 ($l = -1,1$), the mth category of D_1 ($m = -1,1$), and the nth category of D_0; the intercept λ is determined by the constraint that $\Sigma_{ijklmn}p_{ijklmn} = 1$; the remaining λ's are parameters; and $\Sigma_i\lambda_i^A = \Sigma_j\lambda_j^B = \Sigma_k\lambda_k^{C1} = \Sigma_l\lambda_l^{C0} = \Sigma_m\lambda_m^{D1} = \Sigma_n\lambda_n^{D0} = \Sigma_i\lambda_{ij}^{AB} = \Sigma_j\lambda_{ij}^{AB} = \Sigma_i\lambda_{ik}^{AC1} = \Sigma_k\lambda_{ik}^{AC1} = \Sigma_i\lambda_{il}^{AC0} = \Sigma_l\lambda_{il}^{AC0} = \Sigma_j\lambda_{jm}^{BD1} = \Sigma_m\lambda_{jm}^{BD1} = \Sigma_j\lambda_{jn}^{BD0} = \Sigma_n\lambda_{jn}^{BD0} = \Sigma_k\lambda_{km}^{C1D1} = \Sigma_m\lambda_{km}^{C1D1} = \Sigma_k\lambda_{kn}^{C1D0} = \Sigma_n\lambda_{kn}^{C1D0} = \Sigma_l\lambda_{lm}^{C0D1} = \Sigma_m\lambda_{lm}^{C0D1} = \Sigma_l\lambda_{ln}^{C0D0} = \Sigma_n\lambda_{ln}^{C0D0} = 0$.

Equation (17) contains more parameters than are required by the reciprocal-effects model. Thus, we impose some further restrictions. Equation (17) includes distinct effects of A on C_1 and C_0 (that is, the effect of A on C interacts with D) and effects of B on D_1 and D_0 (that is, the effect of B on D interacts with D). In contrast, model d in Figure 1 assumes a single effect of A on C and B on D. Thus, we also assume that $\lambda_{ik}^{AC1} = \lambda_{il}^{AC0}$ and that $\lambda_{jm}^{BD1} = \lambda_{jn}^{BD0}$. These restrictions imply that

$$\log p_{ijklmn} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^{C1} + \lambda_l^{C0} + \lambda_m^{D1} + \lambda_n^{D0} + \lambda_{ij}^{AB} + \lambda_{jt}^{BD^*} + \lambda_{is}^{AC^*} + \lambda_{st}^{C^*D^*}, \tag{18}$$

where C^* and D^* are as defined above.

Equation (18) is a structural loglinear model for the reciprocal effects of C and D. The effect of C on D is $2(\lambda_m^{D1} - \lambda_n^{D0})$, and the effect of D on C is $2(\lambda_k^{C1} - \lambda_l^{C0})$. Note that λ_m^{D1} denotes the average level of D when $C = 1$ and that λ_n^{D0} denotes the average level of D when $C = -1$. Thus, the effect of C on D is proportional to the difference between these two parameters. Similarly, λ_k^{C1} and λ_l^{C0} denote the average levels of C when $D = 1$ and $D = -1$, respectively, and the effect of D on C is proportional to the difference between these two parameters. The partial association between C and D is $\lambda_{st}^{C^*D^*}$. It measures the association between C and D that remains once their reciprocal effects and the effects of A and B are taken into account. Whereas model (17) includes four parameters for associations among C_1, C_0, D_1, and D_0, model (18) includes only a single parameter, which is analogous to the residual correlation in the structural form of a conventional simultaneous equation model.

4.3. *Logit Form of the Model*

To show the link between the structural loglinear model and the general model presented in section 4.1, we write (18) as four logit equations, one each for C_1, C_0, D_1, and D_0. That is,

$$\text{logit}[p(C_1 = 1 \mid A,D_1,D_0)] = \beta^{C1} + \beta_i^{AC^*} + \beta^{C^*D^*}, \qquad (19)$$

$$\text{logit}[p(C_0 = 1 \mid A,D_1,D_0)] = \beta^{C0} + \beta_i^{AC^*} + \beta^{C^*D^*}, \qquad (20)$$

$$\text{logit}[p(D_1 = 1 \mid B,C_1,C_0)] = \beta^{D1} + \beta_j^{BD^*} + \beta^{C^*D^*}, \qquad (21)$$

$$\text{logit}[p(D_0 = 1 \mid B,C_1,C_0)] = \beta^{D0} + \beta_j^{BD^*} + \beta^{C^*D^*}, \qquad (22)$$

where $\beta^{C1} = 2\lambda_k^{C1}, \beta^{C0} = 2\lambda_l^{C0}, \beta^{D1} = 2\lambda_m^{D1}, \beta^{D0} = 2\lambda_n^{D0}, \beta_i^{AC^*} = 2\lambda_{is}^{AC}, \beta_j^{BD^*} = 2\lambda_{jt}^{BD^*}$, and $\beta^{C^*D^*} = 2\lambda_{st}^{C^*D^*}$. Equations (19–22) correspond to the general model given by (13)–(16), but we can also write the model as two logit equations, one for C and one for D:

$$\text{logit}[p(C_r = 1 \mid A,D_1,D_0)] = \beta^{C} + \beta_r^{CD} + \beta_i^{AC^*} + \beta^{C^*D^*}, \quad (23)$$

$$\text{logit}[p(D_q = 1 \mid B,C_1,C_0)] = \beta^{D} + \beta_q^{DC} + \beta_j^{BD^*} + \beta^{C^*D^*}, \quad (24)$$

where C_r is a variable taking the value 1 or -1 for the rth category of D ($r = 0,1$); D_q is a variable taking the value 1 or -1 for the qth category of C ($q = 0,1$); β^C is the grand mean of the logit of the probability that $C_r = 1$ $[\beta^C = (\beta^{C1} + \beta^{C0})/2]$; β^D is the grand mean of the logit of the probability that $D_q = 1$ $[\beta^D = (\beta^{D1} + \beta^{D0})/2]$; β_r^{CD} is the structural effect of D on C (that is, $\beta_1^{CD} = \beta^{C1} - \beta^{C0}$, and $\beta_0^{CD} = \beta^{C0} - \beta^{C1}$); β_q^{DC} is the structural effect of C on D (that is, $\beta_1^{DC} = \beta^{D1} - \beta^{D0}$, and $\beta_0^{DC} = \beta^{D0} - \beta^{D1}$); and the remaining notation is as defined above.

Equations (23) and (24), therefore, contain structural parameters for the effects of B and D on C and the effects of A and C on D, as well as the association between C and D that remains once their dependence on each other and on A and B is taken into account. Note that the structural parameters are unrestricted and that in general, $\beta_r^{CD} \neq \beta_q^{DC}$. That $\beta^{C^*D^*}$ enters both (23) and (24) implies that the two logit models should be estimated jointly. This contrasts with Brier's (1978) formulation in which the logit model for each dependent variable is estimated separately and yields identical parameters for the reciprocal relationship between C and D.

4.4. *Empirical Illustration*

As in the case of the single-equation logit model, we obtain the parameters for the two-equation logit model (23)–(24) from its corresponding structural loglinear model for the expanded table (18). Table 5 presents the design matrix and cell mapping for (18). As in Table 2, the rows in Table 5 refer to cells in the expanded table.

TABLE 5
Design Matrix, Cell Mapping, and Expected Frequencies for Reciprocal-Effects Model for Four-Variable Leading-Crowd Data

Observed Frequency	C	D	C_1	C_0	D_1	D_0	A	B	AB	AC^*	BD^*	C^*D^*	Expected Frequencies (Expanded Table) Model III
			Model Terms										
458	1	1	1	1	1	1	1	1	1	1	1	2	308.1
458	1	1	1	1	1	-1	1	1	1	1	0	0	47.9
458	1	1	1	-1	1	1	1	1	1	0	1	0	74.8
458	1	1	1	-1	1	-1	1	1	1	0	0	0	22.0
140	1	-1	1	1	-1	1	1	1	1	1	0	0	92.5
140	1	-1	1	1	-1	-1	1	1	1	1	-1	-2	14.4
140	1	-1	-1	1	-1	1	1	1	1	0	0	0	30.0
140	1	-1	-1	1	-1	-1	1	1	1	0	-1	0	8.8
110	-1	1	-1	1	1	1	1	1	1	0	1	0	52.8
110	-1	1	-1	1	-1	1	1	1	1	0	0	0	30.0
110	-1	1	-1	-1	1	1	1	1	1	-1	1	-2	12.8
110	-1	1	-1	-1	-1	1	1	1	1	-1	0	0	13.8
49	-1	-1	1	-1	1	-1	1	1	1	0	0	0	22.0
49	-1	-1	1	-1	-1	-1	1	1	1	0	-1	0	12.5
49	-1	-1	-1	-1	1	-1	1	1	1	-1	0	0	7.1
49	-1	-1	-1	-1	-1	-1	1	1	1	-1	-1	2	7.7
171	1	1	1	1	1	1	1	-1	-1	1	-1	2	88.9
171	1	1	1	1	1	-1	1	-1	-1	1	0	0	43.6
171	1	1	1	-1	1	1	1	-1	-1	0	-1	0	21.6
171	1	1	1	-1	1	-1	1	-1	-1	0	0	0	20.0
182	1	-1	1	1	-1	1	1	-1	-1	1	0	0	84.3
182	1	-1	1	1	-1	-1	1	-1	-1	1	1	-2	41.4
182	1	-1	-1	1	-1	1	1	-1	-1	0	0	0	27.3
182	1	-1	-1	1	-1	-1	1	-1	-1	0	1	0	25.4
56	-1	1	-1	1	1	1	1	-1	-1	0	-1	0	15.2
56	-1	1	-1	1	-1	1	1	-1	-1	0	0	0	27.3
56	-1	1	-1	-1	1	1	1	-1	-1	-1	-1	-2	3.7
56	-1	1	-1	-1	-1	1	1	-1	-1	-1	0	0	12.6

87	-1	-1	1	-1	1	-1	1	-1	-1	0	0	0	20.0
87	-1	-1	1	-1	-1	-1	1	-1	-1	0	1	0	35.9
87	-1	-1	-1	-1	1	-1	1	-1	-1	-1	0	0	6.5
87	-1	-1	-1	-1	-1	-1	1	-1	-1	-1	1	2	22.1
184	1	1	1	1	1	1	-1	1	-1	-1	1	2	41.6
184	1	1	1	1	1	-1	-1	1	-1	-1	0	0	6.5
184	1	1	1	-1	1	1	-1	1	-1	0	1	0	108.2
184	1	1	1	-1	1	-1	-1	1	-1	0	0	0	31.8
75	1	-1	1	1	-1	1	-1	1	-1	-1	0	0	12.5
75	1	-1	1	1	-1	-1	-1	1	-1	-1	-1	-2	1.9
75	1	-1	-1	1	-1	1	-1	1	-1	0	0	0	43.4
75	1	-1	-1	1	-1	-1	-1	1	-1	0	-1	0	12.8
531	-1	1	-1	1	1	1	-1	1	-1	0	1	0	76.5
531	-1	1	-1	1	-1	1	-1	1	-1	0	0	0	43.4
531	-1	1	-1	-1	1	1	-1	1	-1	1	1	-2	199.1
531	-1	1	-1	-1	-1	1	-1	1	-1	1	0	0	213.8
281	-1	-1	1	-1	1	-1	-1	1	-1	0	0	0	31.8
281	-1	-1	1	-1	-1	-1	-1	1	-1	0	-1	0	18.1
281	-1	-1	-1	-1	1	-1	-1	1	-1	1	0	0	110.7
281	-1	-1	-1	-1	-1	-1	-1	1	-1	1	-1	2	118.9
85	1	1	1	1	1	1	-1	-1	1	-1	-1	2	12.8
85	1	1	1	1	1	-1	-1	-1	1	-1	0	0	6.3
85	1	1	1	-1	1	1	-1	-1	1	0	-1	0	33.2
85	1	1	1	-1	1	-1	-1	-1	1	0	0	0	30.8
97	1	-1	1	1	-1	1	-1	-1	1	-1	0	0	12.1
97	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-2	5.9
97	1	-1	-1	1	-1	1	-1	-1	1	0	0	0	42.1
97	1	-1	-1	1	-1	-1	-1	-1	1	0	1	0	39.1
338	-1	1	-1	1	1	1	-1	-1	1	0	-1	0	23.5
338	-1	1	-1	1	-1	1	-1	-1	1	0	0	0	42.1
338	-1	1	-1	-1	1	1	-1	-1	1	1	-1	-2	61.1
338	-1	1	-1	-1	-1	1	-1	-1	1	1	0	0	207.3
554	-1	-1	1	-1	1	-1	-1	-1	1	0	0	0	30.8
554	-1	-1	1	-1	-1	-1	-1	-1	1	0	1	0	55.3
554	-1	-1	-1	-1	1	-1	-1	-1	1	1	0	0	107.3
554	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	2	364.2

TABLE 6
Likelihood-Ratio Chi-Square Statistics for Four-Variable Models for Leading-Crowd Data

Model	G^2	*df*
Structural loglinear models		
I. AB, AC^*, BD^*, C^*D^*	9.78	7
II. $AB, C_1, C_0, D_1, D_0, AC^*, BD^*$	30.28	6
III. $AB, C_1, C_0, D_1, D_0, AC^*, BD^*, C^*D^*$	1.17	5
Elementary loglinear models		
IV. AB, AC, BD, CD	17.91	7
V. AB, AC, AD, BC, BD	15.72	6
VI. AB, AC, AD, BC, BD, CD	1.21	5

Except for the C and D columns, each column of the design matrix corresponds to a parameter in (18). The columns for C_1, C_0, D_1, and D_0 are contrast-coded indicator variables that denote positions in Figure 3. If we denote these variables as d^A, d^B, d^{C1}, d^{C0}, d^{D1}, and d^{D0}, respectively, then the columns for higher-way interaction are formed as follows: $d^{AB} = d^A d^B$; $d^{AC^*} = d^A(d^{C1} + d^{C0})/2$; $d^{BD^*} = d^B(d^{D1} + d^{D0})/2$; and $d^{C^*D^*} = (d^{C1} + d^{C0})(d^{D1} + d^{D0})/4$. This coding reflects the equality restrictions that we impose on (17). The first column shows the observed frequency that is mapped into each cell of the expanded table. The final column reports the expected frequencies for one of the models that is discussed below.[5]

The top panel of Table 6 presents G^2 statistics for (18) plus several simpler models fit to Table 1. Model I includes the partial association between C and D but no structural relationships between these two variables. This model fits the data well. Model II includes the structural effects of C on D and D on C but not the partial association between the two variables. This model fits the data

[5]The structure of Table 5 is analogous to that of Table 2. Table 5 has 64 rows, 16 for each of the four combinations of A and B. As shown in Figure 3, for each combination of A and B, only 14 of the 16 combinations of C_1, C_0, D_1, and D_0 are logically possible. Within levels of A and B, Table 5 contains a row for each of the 14 combinations, for a total of 56 rows. Figure 3 contains eight additional rows for cells 6 and 11 because these cells receive observations from more than one source. This leads to a total of 64 rows in Table 5. See the discussion of Table 3 for further details.

TABLE 7
Parameter Estimates in Logit Form of Structural Loglinear Models for the Four-Way Table

Variable[a]	Model I $\hat{\beta}$	Model I SE($\hat{\beta}$)	Model II $\hat{\beta}$	Model II SE($\hat{\beta}$)	Model III $\hat{\beta}$	Model III SE($\hat{\beta}$)
Effects on C						
Intercept	−.168	.044	−.115	.044	−.234	.051
$C_1 - C_0$			.142	.146	.347	.160
A	2.444	.086	2.498	.086	2.372	.094
Effects on D						
Intercept	.329	.039	.287	.039	.320	.041
$D_1 - D_0$			.314	.064	−.658	.247
B	1.176	.072	1.174	.074	1.150	.074
C^*D^*	.344	.104			1.276	.264

[a] A is membership in the leading crowd in wave 1; B is attitude toward the leading crowd in wave 1; C is membership in the leading crowd in wave 2; and D is attitude toward the leading crowd in wave 2.

poorly. Model III includes both the structural effect and the partial association between C and D. This model fits the data extremely well, both absolutely and relative to models I and II. The expected frequencies of the expanded table under model III are reported in the final column of Table 5.

Table 7 presents estimates of the logit parameters for these models. In all three models, membership in the leading crowd in wave 1 substantially increases the odds of membership in wave 2. Likewise, holding a favorable attitude toward the leading crowd in wave 1 substantially increases the odds of a favorable attitude in wave 2. The three models, however, yield different results about the relationships between membership in wave 2 and attitude in wave 2. Model I indicates a net positive association between membership and favorable attitude. Model II indicates that the reciprocal effects of membership and attitude are both positive, but only the effect of membership on attitude is statistically significant. The estimate for the latter effect, moreover, is more than twice the estimate for the effect of attitude on membership. Model III, in contrast, indicates that a favorable attitude toward the leading crowd in wave 2 significantly raises the odds of joining the leading crowd in wave 2. But membership in the leading crowd reduces the chances of holding a favorable attitude. Model III also shows a positive partial association between membership and attitude. These results are consistent with those reported in Table 4 for the three-variable table. The simultaneous equation model in the latter table also indicated a negative effect of membership on attitude once the joint determination of the two variables was taken into account. These calculations illustrate that the structural loglinear model enables us to isolate the separate reciprocal effects of a pair of endogenous variables and to distinguish these effects from their residual association.[6]

The lower panel of Table 6 reports G^2 statistics for several elementary loglinear models. The fit of the model of no three-way (all two-way) interactions, model VI, to the observed data is very close to the fit for model III, but it is not identical. The structural and elementary loglinear models are distinct, and the parameters of one model cannot be derived from those of the other. The choice between models VI and III should be determined by one's analytic goals and the plausibility of the assumption that A does not directly affect D and that B does not directly affect C.

5. IDENTIFICATION

General rules for the identification of structural loglinear models have not yet been developed, but some guidance is available from results on the identifiability of models for missing data in loglinear models. In this section we show the link between the identifiability of our models and that of certain models for missing data. This enables us to show that the models presented in sections 3 and 4 are in fact identified and to suggest some general guidelines for identification of structural loglinear models. The identification conditions presented in this section are sufficient but not necessary conditions.

5.1. *Identification of Models for Missing Categorical Data*

The structural loglinear models presented in this paper are extensions of models for categorical data in which some variables are

[6]We also estimated model III allowing for C_1C_0 and D_1D_0 associations. For this model, $G^2 = 0.806$ (3 *df*). The estimated parameters for these terms are, respectively, -0.931 (SE $= 4.60$) and 0.756 (SE $= 0.938$). Under this specification the estimate for $D_1 - D_0$ is -0.390 (SE $= 0.2115$) and for $C_1 - C_0$ is 0.332 (SE $= 0.1743$). In general, standard errors are larger under this model.

not fully observed (Little and Rubin 1987; Winship and Mare 1989). In (9) and (18), C_1, C_0, D_1, and D_0 are each partially observed, and for each of them, a fully observed variable denotes whether or not they have "missing data." Variable C_1 is observed when $D = 1$ and missing when $D = -1$, whereas C_0 is observed and missing under the opposite conditions. Likewise, D_1 is observed when $C = 1$ and missing when $C = -1$, whereas D_0 is observed and missing under the opposite conditions. Each of these partially observed variables, moreover, is potentially subject to *nonignorable nonresponse* (NINR); that is, whether or not the variables are missing is associated with the level of the variable itself (e.g., Winship and Mare 1989). For example, D_1 in (8) and (17) is associated with C_1 and C_0 and thereby with C, which determines whether or not D_1 is observed. This implies that the identifiability of model terms involving C_1, C_0, D_1, and D_0 should be governed by rules similar to those that apply to other, simpler contingency tables in which a variable is subject to NINR.

These rules can be summarized as follows. Consider a two-way ($J \times K$) table with dimensions X and Y. Let all observations be present for X, but let some observations be missing for Y. A third variable, M, denotes whether data are missing on Y. A potentially identifiable NINR model is $(XY)(YM)$, that is, a model in which the fully observed variable X is associated with the partially observed variable Y but is conditionally independent of whether data are missing on Y. This model is identified if $J \geq K$, that is, if the number of categories of the fully observed variable is at least as large as the number of categories of the partially observed variable (Little and Rubin 1987, pp. 238–39). More generally, NINR models are identified if (a) for every partially observed variable, there exists a fully observed variable that is conditionally independent of whether data are missing on the partially observed variable and (b) the number of categories of the partially observed variable does not exceed the number of categories of the fully observed variable (Winship and Mare 1989).

5.2. *Identification of Structural Loglinear Models*

To establish the identifiability of structural loglinear models (9) and (18), we rely on the results for models with missing data stated above plus the fact that some parameters for models of the expanded table can be identified without any overidentifying restrictions on the structural model. Consider a simplified version of (18),

$$\log p_{klmn} = \lambda + \lambda_s^{C^*} + \lambda_t^{D^*} + \lambda_{st}^{C^*D^*}, \quad (25)$$

which implies the inclusion of the terms $\lambda_k^{C1} = \lambda_l^{C0}$, $\lambda_m^{D1} = \lambda_n^{D0}$, and $\lambda_{km}^{C1D1} = \lambda_{kn}^{C1D0} = \lambda_{lm}^{C0D1} = \lambda_{ln}^{C0D0}$ and where p_{klmn} denotes the probability that an individual falls into the kth category of C_1 ($k = -1,1$), the lth category of C_0 ($l = -1,1$), the mth category of D_1 ($m = -1,1$), and the nth category of D_0; the intercept λ is determined by the constraint that $\Sigma_{klmn} p_{klmn} = 1$; and all other notation is as defined above. Model (25) represents the two-way association between the average values of C_1 and C_0 and of D_1 and D_0. This model can be identified directly from the observed 2×2 CD table without other identifying restrictions. The four parameters of the model are nonlinear functions of the four frequencies in the observed CD table.[7] Because $\lambda_s^{C^*}$, $\lambda_t^{D^*}$, and $\lambda_{st}^{C^*D^*}$ are always identified from the CD table alone, to establish the identifiability of (9) and (18) it suffices to show that parameters for C_1, C_0, D_1, and D_0 and for the relationships between these variables and the exogenous variables are identified.

In the structural loglinear model for a single equation with an endogenous independent variable, model (9), we can identify $\lambda_s^{C^*}$ and $\lambda_{st}^{C^*D^*}$ from the observed CD table, as noted above. We identify λ_j^B and $\lambda_{jt}^{BD^*}$ directly from observed data on the joint distribution of B, C, and D. (The associations between B and D_1 and between B and D_0 are directly observed when $C = 1$ and $C = -1$, respectively. Model (9) constrains these two associations to be equal.) We identify λ_m^{D1} and λ_n^{D0} using the rules for identification of missing-data models discussed above. Under the model, B is conditionally independent of C_1 and C_0, given D_1 and D_0, and is thus conditionally independent of C, which determines whether D_1 and D_0 are missing (since C is determined by C_1, C_0, D_1, and D_0). Thus, B is fully observed and conditionally independent of whether data are missing on D_1 and D_0. Since the number of categories of B equals the number of categories in D_1 and D_0, the inclusion of B in the model identifies λ_m^{D1} and λ_n^{D0}. Thus, model (9) and its corresponding logit form, model (12), are identified.

[7]To see this, write the expected frequencies of the *expanded* table in terms of the parameters of (24) and collapse the expanded table to the observed table. If x_{uv} denotes the observed frequency in the 2×2 table for the uth level of C and the vth level of D ($u = 1,2$; $v = 1,2$), then

$$x_{uv} = \sum_{klmn \in uv} [\exp(\lambda + \lambda_{kl}^{C^*} + \lambda_{mn}^{D^*} + \lambda_{st}^{C^*D^*})].$$

In the structural loglinear model for reciprocal effects, model (18), we can identify $\lambda_{st}^{C^*D^*}$ from the observed CD table, as noted above. We identify λ_i^A, λ_j^B, λ_{ij}^{AB}, $\lambda_{jt}^{BD^*}$, and $\lambda_{is}^{AC^*}$ directly from the observed data on the joint distributions of A, B, C, and D. We identify λ_m^{D1} and λ_n^{D0} from the rules for identification of missing-data models. The inclusion of B, which is associated with D_1 and D_0 but conditionally independent of C, identifies these two parameters. Likewise, we identify λ_k^{C1} and λ_l^{C0} by the inclusion of A, which is associated with C_1 and C_0 but conditionally independent of D. Thus, model (18) and its corresponding logit form, equations (23)–(24), are identified.

Our analysis of the identifiability of models (9) and (18) suggests that parameters of structural loglinear models can be identified by restrictions on the parameters for the effects of the exogenous variables. In particular, a sufficient condition for the identification of the effect of one endogenous variable on another is that the model include an exogenous variable that affects the dependent endogenous variable but not the independent endogenous variable and that the exogenous variable have at least as many categories as the dependent endogenous variable. This principle is illustrated by the use of B to identify λ_m^{D1} and λ_n^{D0} in (9) and λ_m^{D1} and λ_n^{D0} and by the use of A to identify λ_k^{C1} and λ_l^{C0} in (18).

These identification conditions resemble those that govern the use of instrumental variables to identify conventional simultaneous equation models, but they differ in one key respect. In conventional simultaneous equation models, instrumental variables are *excluded* from the structural equation of interest but affect the endogenous independent variable in the equation. In the models described above, however, the exogenous variables are *included* in the structural equation of interest but are conditionally independent of the endogenous independent variable. This distinction is operationally significant only for the single-equation model. For the two-equation model, the same pattern of restrictions on the relationships among the four variables applies in the conventional simultaneous equation and the structural loglinear models. A single-equation model that follows the more usual conventions of simultaneous equation model estimation is model *i* in Figure 1. In this model the exogenous variable A affects the endogenous independent variable D but not the dependent endogenous variable C. It is possible to show that the structural loglinear model for *i* is also identified, using modified versions of the arguments presented above. For the sake of brevity, we do not present these arguments here.

6. CONCLUSION

We have proposed models for jointly determined categorical variables that are extensions of loglinear and logit models for cross-classified data. We have shown that it is possible to develop loglinear models that are analogous to standard linear simultaneous equation models. The models proposed here can be estimated using the same algorithms and software that are available for loglinear models of partially or indirectly observed contingency tables.

An issue for further research on these models concerns the robustness of their results to alternative identifying restrictions. In estimating structural loglinear models, we have made a number of assumptions. The two most important are (a) that the C_1C_0 and D_1D_0 interactions are zero (that is, $\lambda_{lk}^{C1C0} = \lambda_{mn}^{D1D0} = 0$) and (b) that all the nonstructural associations between C and D are equal (that is, $\lambda_{km}^{C1D1} = \lambda_{kn}^{C1D0} = \lambda_{lm}^{C0D1} = \lambda_{ln}^{C0D0}$). Some of these restrictions can be relaxed, but the model would be underidentified if all were relaxed simultaneously. Thus, it is not possible to test individual restrictions when all other parameters are unrestricted. As noted in footnotes 4 and 6, our empirical results change slightly when C_1C_0 and D_1D_0 are not restricted to zero in the examples presented. Analyses of hypothetical data not reported here, however, suggest that different specifications may produce different estimates of $\beta^{C1} - \beta^{C0}$ and $\beta^{D1} - \beta^{D0}$. Thus, our models may not always be robust to the specification of the error structure. This issue awaits further investigation.

Our discussion has been confined to models with two dichotomous endogenous variables, but these methods generalize to models with polytomous variables and more than two outcomes. Models with more complex dependent variables may require much larger expanded tables than the simple case presented here. For example, a table that contains one dichotomous and one four-category dependent variable has $2 \times 4 = 8$ cells for each unique combination of the exogenous variables and $2^4 \times 4^2 = 256$ cells in the corresponding expanded table. Larger models obviously require more computation than the elementary models presented here.

The models presented in this paper apply when one is concerned only with simultaneous relationships between discrete variables. If a discrete variable exerts its effect not only discretely but also as an indicator of an underlying continuous variable, then models other than those presented here are more suitable (Heckman 1978; Maddala 1983; Winship and Mare 1983).

REFERENCES

Amemiya, Takeshi. 1985. *Advanced Econometrics.* Cambridge, MA.: Harvard University Press.

Baker, Stuart G., and Nan M. Laird. 1988. "Regression Analysis for Categorical Variables with Outcome Subject to Nonignorable Response." *Journal of the American Statistical Association* 83:62–69.

Bishop, Yvonne M., Stephen E. Fienberg, and Paul W. Holland. 1975. *Discrete Multivariate Analysis.* Cambridge, MA: MIT Press.

Brier, Stephen S. 1978. "The Utility of Systems of Simultaneous Logistic Response Equations." Pp. 119–29 in *Sociological Methodology 1979,* edited by K. F. Schuessler. San Francisco: Jossey-Bass.

Coleman, James S. 1961. *The Adolescent Society: The Social Life of the Teenager and its Impact on Education.* Glencoe, IL: Free Press.

———. 1964. *Introduction to Mathematical Sociology.* New York: Free Press.

Daganzo, Carlos. 1979. *Multinomial Probit: The Theory and its Application to Demand Forecasting.* New York: Academic Press.

Duncan, Beverly, and Otis Dudley Duncan. 1978. *Sex Typing and Social Roles.* New York: Academic Press.

Duncan, Otis Dudley. 1974. "A Model of Interaction Between Spouses." Unpublished manuscript. University of Arizona, Department of Sociology.

———. 1975. *Introduction to Structural Equation Models.* New York: Academic Press.

———. 1985. "New Light on the 16-fold Table." *American Journal of Sociology* 91:88–128.

Duncan, Otis Dudley, Archibald O. Haller, and Alejandro Portes. 1968. "Peer Influences on Aspirations: A Reinterpretation." *American Journal of Sociology* 74:119–37.

Fay, Robert E. 1986. "Causal Models for Patterns of Nonresponse." *Journal of the American Statistical Association* 81:354–65.

Fienberg, Stephen E. 1980. *The Analysis of Cross-Classified Categorical Data.* Cambridge, MA: MIT Press.

Goldberger, Arthur S., and Otis Dudley Duncan. 1973. *Structural Equation Models in the Social Sciences.* New York: Seminar Press.

Goodman, Leo A. 1973. "Causal Analysis of Data from Panel Studies and Other Kinds of Surveys. *American Journal of Sociology* 78:1135–91.

———. 1974. "The Analysis of Systems of Qualitative Variables When Some of the Variables are Unobservable. I. A Modified Latent Structure Approach." *American Journal of Sociology* 79:1179–1259.

———. 1978. *Analyzing Qualitative/Categorical Data.* Cambridge, MA: Abt Books.

Haberman, Shelby J. 1978–79. *Analysis of Qualitative Data.* 2 vols. New York: Academic Press.

———. 1988. "A Stabilized Newton-Raphson Algorithm for Log-Linear Models for Frequency Tables Derived by Indirect Observation." Pp. 193–211 in *Sociological Methodology 1988,* edited by C. C. Clogg. Washington, DC: American Sociological Association.

Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46:931–59.

———. 1978. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153–61.

Heckman, James J., and V. Joseph Hotz. 1989. "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association* 84:862–74.

Little, Roderick J. A., and Donald B. Rubin. 1987. *Statistical Analysis with Missing Data.* New York: Wiley.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Mallar, Charles D. 1977. "The Estimation of Simultaneous Probability Models." *Econometrica* 45:1717–22.

Mare, Robert D., and Christopher Winship. 1988. "Endogenous Switching Regression Models for the Causes and Effects of Discrete Variables." Pp. 132–60 in *Common Problems/Proper Solutions: Avoiding Error in Quantitative Research,* edited by J. S. Long. Beverly Hills: Sage University Press.

Marini, Margaret M. 1984. "Women's Educational Attainment and the Timing of Entry into Parenthood." *American Sociological Review* 49:491–511.

Muthén, Bengt. 1984. "A General Structural Equation Model with Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators." *Psychometrika* 49:115–32.

Rubin, Donald B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *Annals of Statistics* 6:34–58.

Stolzenberg, Ross M., and Linda J. Waite. 1977. "Age, Fertility Expectations, and Plans for Employment." *American Sociological Review* 42:769–82.

Winship, Christopher, and Robert D. Mare. 1983. "Structural Equations and Path Analysis for Discrete Data." *American Journal of Sociology* 89: 54–110.

———. 1989. "Loglinear Models for Missing Data: A Latent Class Approach." Pp. 331–67 in *Sociological Methodology 1989,* edited by C. C. Clogg. Oxford: Basil Blackwell.