7

ASYMPTOTIC CONFIDENCE INTERVALS FOR INDIRECT EFFECTS IN STRUCTURAL EQUATION MODELS

Michael E. Sobel

INDIANA UNIVERSITY

Since its general introduction to the sociological community (Duncan, 1966), path analysis has become one of the most widely used tools in sociological research, and through its use sociologists have become more sophisticated about the general

For comments on an earlier draft of this chapter and for detailed advice I am indebted to Robert M. Hauser, Halliman H. Winsborough, and Toni Richards, several anonymous reviewers, and the editor of this volume. I also wish to thank John Raisian, Nancy Rytina, and Barbara Mann for their comments and Mark Wilson for able research assistance. The opinions expressed here are the sole responsibility of the author.

topic of structural equation models. (For a thorough review of the literature, see Bielby and Hauser, 1977.)

An intriguing aspect of path analysis, as traditionally propounded, is that "it makes explicit both the direct and the indirect effects of causal variables on dependent variables" (Duncan, Featherman, and Duncan, 1972, p. 23) and thereby allows for a detailed substantive accounting of the sociological process under investigation. Despite the enormous interest in indirect effects displayed by sociologists (Finney, 1972; Heise, 1975; Land, 1969; Lewis-Beck, 1974; Blalock, 1971; Duncan, Featherman, and Duncan, 1972; Blau and Duncan, 1967; Alwin and Hauser, 1975; Duncan, 1975), the distribution of these effects has been ignored. Thus sociologists (and researchers in many other disciplines as well) typically treat the indirect effects they calculate as parameter values and formulate inferences without asking whether the effect itself is statistically significant.

I propose to remedy this situation by deriving the asymptotic distribution of indirect effects and indicating, both formally and by example, how researchers may easily compute confidence intervals for their estimates. Since I assume that the reader is already well acquainted with the literature on indirect effects, I offer no extended discussion, per se, of this literature. Although I focus on recursive models, the results that are stated hold true for functions of the structural coefficients under quite general conditions. Since an indirect effect is defined as such a function, there is no need to proceed on a case by case basis.

In the first section, I show how confidence intervals for the indirect effects of a recursive structural equation model may be obtained. In the second section, I work out the example in Alwin and Hauser (1975), which is taken from Duncan, Featherman, and Duncan (1972), and demonstrate the ease with which the confidence intervals are obtained. In addition, I show how inferences about the relative magnitudes of direct and indirect effects may be formulated. The results of this comparison demonstrate the value of computing asymptotic confidence intervals and suggest that researchers proceed with caution before offering detailed substantive explications of the indirect effects in complex structural models.

THE FORMAL THEORY

In a recursive structural equation model, the indirect effects of an independent variable on a dependent variable may be expressed as a linear combination of products of structural parameters.[1] Similarly, in a nonrecursive model or a model with latent variables the indirect effects may be expressed as a function of products of the structural parameters (Fox, 1980; Schmidt, 1980).[2] In either case, the indirect effects may be regarded as a nonlinear function of the structural parameters.

In order to place confidence intervals around the indirect effects, their distribution must first be determined. Among other things, this requires, for a given model, an assessment of the distribution of the complete coefficient vector under a particular estimation procedure. The techniques that are generally used for estimating structural equation models are either maximum-likelihood procedures (for example, full-information maximum likelihood) or procedures that have similar asymptotic properties (such as three-stage least squares). The justification for considering these techniques derives largely from the fact that under general regularity conditions the complete coefficient vector is consistent, asymptotically normally distributed, and efficient (Theil, 1971, pp. 392–396). That is, whereas the small-sample properties of these estimators are often not well known, the large-sample properties are both attractive and tractable. For this reason, primarily, a general procedure for assessing the distribution of indirect effects will use asymptotic distribution theory as opposed to exact distribution theory.

[1] In a recursive model there are no feedback relationships among the dependent variables and the disturbances are uncorrelated across equations. A more formal definition is presented later in the exposition.

[2] This is true under the regularity conditions stated by Fox (1980, p. 18) and Schmidt (1980, pp. 9–10).

The provision of an asymptotic distribution theory for indirect effects requires two steps. First, the asymptotic distribution of the coefficient vector must be determined for a given structural equation model and a given estimation procedure. Second, because the indirect effects are a nonlinear function of the structural coefficients, an appropriate method for evaluating the asymptotic distribution of such functions must be employed. The multivariate-delta method is utilized for this purpose: It provides a simple, general, and elegant means for evaluating the asymptotic distribution of functions of multinormally distributed random variables. With these results in hand, and with a consistent estimator of the variances of the indirect effects, we find it easy to construct a $(1 - \alpha)100\%$ confidence interval for the indirect effect under consideration.

Asymptotic Distribution of the Coefficient Vector in a Recursive Model

We begin by considering a system of M equations. For the jth structural equation, $j = 1, \ldots, M$, let the model be

$$\mathbf{y}_j = \mathbf{Z}_j\boldsymbol{\delta}_j + \boldsymbol{\epsilon}_j \tag{1}$$

where $\mathbf{y}_j$ is the vector of dependent random variables for the n observations and $\boldsymbol{\epsilon}_j$ is a stochastic term distributed $N(\mathbf{0}, \sigma^2\mathbf{I})$; moreover, $\mathbf{Z}_j = (\mathbf{X}_j, \mathbf{Y}_j)$ is the full-rank matrix of predetermined and endogenous variables, respectively, and $\boldsymbol{\delta}_j$ is the corresponding parameter vector $(\boldsymbol{\gamma}_j'^*, \boldsymbol{\beta}_j'^*)'$. Combining the M equations, we may rewrite the system in the alternative form

$$\mathbf{YB} = \mathbf{X\Gamma} + \mathbf{E} \tag{2}$$

where $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_M)$, $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_K)$, $\mathbf{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_M)$ is the nonsingular matrix of coefficients of the endogenous variables, $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_M)$ is the $K \times M$ matrix of coefficients for the exogenous variables, and $\mathbf{E} = (\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_M)$. We assume that the restrictions on $\boldsymbol{\Gamma}$ include only zero restrictions, whereas for $\mathbf{B}$ the restrictions include zero restrictions as well as the restrictions that the diagonal elements are set to 1. Finally, we assume that each row of $\mathbf{E}$ is independently and identically distributed $N(\mathbf{0}', \boldsymbol{\Psi})$.

System (2), with the assumptions stated here, specifies the complete form of the structural equation model under consideration. If the model is identified, the unrestricted elements $\boldsymbol{\delta} = (\boldsymbol{\delta}_1', \ldots, \boldsymbol{\delta}_M')'$ in the parameter matrices $\mathbf{B}$ and $\boldsymbol{\Gamma}$ may be estimated by means of maximum likelihood (Rothenberg and Leenders, 1964; Theil, 1971; Jöreskog, 1973). Under general regularity conditions (Theil, 1971, p. 395), this estimator has, in large samples, an approximately normal distribution with mean $\boldsymbol{\delta}$ and variance-covariance matrix $n^{-1}[\mathrm{I}(\boldsymbol{\delta})]^{-1}$, where n is the size of the sample and $[\mathrm{I}(\boldsymbol{\delta})]$ is the Fisher information matrix (Rao, 1973, pp. 324–328; Theil, 1971, pp. 384–396). The precise statement is that

$$n^{1/2}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}) \xrightarrow{\mathscr{L}} N(\mathbf{0}, [\mathrm{I}(\boldsymbol{\delta})]^{-1}) \tag{3}$$

That is, $n^{1/2}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta})$ converges in law (distribution) to the quantity on the right-hand side of (3), where $\hat{\boldsymbol{\delta}}_n$ is the maximum-likelihood estimator of $\boldsymbol{\delta}$ based on a sample of size n. In the special case where $\mathbf{B}$ is an upper-triangular matrix and $\boldsymbol{\Psi}$ is diagonal, the system (2) is recursive (Theil, 1971, p. 461). It is well known that for this case (Rothenberg and Leenders, 1964; Malinvaud, 1970; Theil, 1971) maximum-likelihood estimation of the system reduces to equation-by-equation least-squares estimation because the log-likelihood function is merely the sum of the M likelihood functions for the structural relationships specified by (1). That is, the log-likelihood function for the system (2) is given by

$$c - (n/2) \log |(\mathbf{B}')^{-1}\boldsymbol{\Psi}\mathbf{B}^{-1}| - (\tfrac{1}{2}) \sum_{m'=1}^{M} \sum_{m=1}^{M} \Psi^{m'm}(\mathbf{y}_{m'} - \mathbf{Z}_{m'}\boldsymbol{\delta}_{m'})' (\mathbf{y}_m - \mathbf{Z}_m\boldsymbol{\delta}_m)$$

which reduces to

$$L(\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_M, \boldsymbol{\Psi}) = c + (n/2) \sum_{j=1}^{M} \log \Psi_{jj}^{-1} - (\tfrac{1}{2}) \sum_{j=1}^{M} (\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j)' (\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j)\Psi_{jj}^{-1} \tag{4}$$

in the recursive case. For this case, it may also be shown (see

Appendix) that $[\mathrm{I}(\boldsymbol{\delta})]^{-1}$ is a block-diagonal matrix

$$[\mathrm{Q}(\boldsymbol{\delta})]^{-1} = \begin{bmatrix} \mathbf{Q}_1^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_3^{-1} & & \mathbf{0} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Q}_M^{-1} \end{bmatrix}$$

with components $\mathbf{Q}_j^{-1} = \operatorname{plim} \hat{\Psi}_{jj}(\mathbf{Z}_j'\mathbf{Z}_j/n)^{-1}$.[3] That is, for a recursive model the vector $\hat{\boldsymbol{\delta}}_n$ obtained by combining the M least-squares estimators satisfies

$$n^{1/2}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}) \xrightarrow{\mathscr{L}} N(\mathbf{0}, [\mathrm{Q}(\delta)]^{-1}) \tag{5}$$

Hence $\hat{\boldsymbol{\delta}}_n$ is asymptotically normally distributed.

Now that the asymptotic distribution of the coefficient vector has been established, it is necessary to consider the asymptotic distribution of the indirect effects. As previously indicated, the multivariate-delta method is used for this purpose. Because this technique may be unfamiliar to some readers, a description of the procedure follows. Readers who are familiar with this technique may wish to skip the exposition and go directly to the results.

The Delta Method

The multivariate-delta method (Rao, 1973, pp. 385–389; Bishop, Fienberg, and Holland, 1975, pp. 486–500) provides a general method for establishing the asymptotic distribution of a differentiable vector function of a multinormally distributed random vector. The method is an extension of the single-delta method, which is justified by a theorem which states that if a random variable $\hat{\delta}_n$ satisfies

$$n^{1/2}(\hat{\delta}_n - \delta) \xrightarrow{\mathscr{L}} N(0, v^2(\delta)) \tag{6}$$

[3] Note that $\hat{\Psi}_{jj}$ and $\mathbf{Z}_j'\mathbf{Z}_j$ both depend on n. Because it is cumbersome to subscript all the random variables, the n subscript is omitted when there is no risk of confusion.

then any function f that is differentiable in a neighborhood of δ will satisfy

$$n^{1/2}(f(\hat{\delta}_n) - f(\delta)) \xrightarrow{\mathscr{L}} N(0, [f'(\delta)]^2\, v^2(\delta)) \tag{7}$$

provided $f'(\delta)$ does not vanish (Rao, 1973, pp. 386–387). For example, in a single-variable regression model

$$y - \bar{y} = b(x - \bar{x}) + e$$

with conditional variance σ^2 and least-squares estimator $\hat{b}$, the asymptotic distribution of $(\hat{b})^{-1}$ may be of interest. By (5) we know that

$$n^{1/2}(\hat{b}_n - b) \xrightarrow{\mathscr{L}} N(0, \sigma^2 \lim(\Sigma(x - \bar{x})^2/n)^{-1})$$

and application of (7) yields the conclusion that[4]

$$n^{1/2}(\hat{b}_n^{-1} - b^{-1}) \xrightarrow{\mathscr{L}} N(0, (-b^{-2})^2\, \sigma^2 \lim(\Sigma(x - \bar{x})^2/n)^{-1})$$

Note, however, that as b approaches zero, $(\hat{b})^{-1}$ becomes unbounded and the delta method cannot be applied.

To see heuristically why the single-delta method works, we consider a Taylor expansion of f, a function of the parameter set. Suppose that f is twice differentiable in a neighborhood of δ. A Taylor expansion of f about δ is given by

$$f(\hat{\delta}_n) - f(\hat{\delta}) = f'(\hat{\delta})\,(\hat{\delta}_n - \hat{\delta}) + f''(\hat{\delta}^*)\,(\hat{\delta}_n - \hat{\delta})^2/2 \tag{8}$$

where δ^* is some number in the interval (δ_n, δ). For large n, δ^* is close to δ, so that the first term on the right-hand side of (8) dominates the second, which may be considered negligible. Thus

$$f(\hat{\delta}_n) - f(\delta) \approx f'(\delta)\,(\hat{\delta}_n - \delta) \tag{9}$$

That is, $f(\hat{\delta}_n) - f(\delta)$ is an approximately linear function of $\hat{\delta}_n$. Because $\hat{\delta}_n$ is approximately normal in large samples and linear functions of normally distributed random variables are normally distributed, $f(\hat{\delta}_n)$ has an approximate normal distribution

[4] For this result to hold, it is technically necessary to assume the existence of the limit in question (Theil, 1971, p. 363).

in large samples. Taking expectations and variances of both sides of (9), we see that

$$E(f(\hat{\delta}_n)) \approx f(\delta)$$

and

$$V(f(\hat{\delta}_n)) \approx n^{-1}[f'(\delta)]^2 v^2(\delta)$$

That is, for large n the result (7) holds.

To consider the multivariate-delta method, let f be a function that is differentiable in a neighborhood of an S-dimensional vector $\boldsymbol{\delta}$ and let

$$(\boldsymbol{\partial} f/\boldsymbol{\partial\delta})' = (\partial f/\partial\delta_1, \ . \ . \ . \ , \partial f/\partial\delta_S)$$

If

$$n^{1/2}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}) \xrightarrow{\mathscr{L}} N(\mathbf{0},\boldsymbol{\Sigma}(\boldsymbol{\delta})) \tag{10}$$

then

$$n^{1/2}(f(\hat{\boldsymbol{\delta}}_n) - f(\boldsymbol{\delta})) \xrightarrow{\mathscr{L}} N(\mathbf{0}, (\boldsymbol{\partial} f/\boldsymbol{\partial\delta})'\boldsymbol{\Sigma}(\boldsymbol{\delta})(\boldsymbol{\partial} f/\boldsymbol{\partial\delta})) \tag{11}$$

provided the quadratic form $(\boldsymbol{\partial} f/\boldsymbol{\partial\delta})' \ \boldsymbol{\Sigma}(\boldsymbol{\delta}) \ (\boldsymbol{\partial} f/\boldsymbol{\partial\delta})$ does not vanish (Rao, 1973, p. 387).

Heuristically, this result may be justified by an extension of the argument used to establish the plausibility of the single-delta method. Suppose that f has the Taylor expansion

$$f(\hat{\boldsymbol{\delta}}_n) - f(\boldsymbol{\delta}) = (\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta})' \ (\boldsymbol{\partial} f/\boldsymbol{\partial\delta}) + R_n \tag{12}$$

where R_n approaches zero as n increases without bound. Then, for n large,

$$f(\hat{\boldsymbol{\delta}}_n) - f(\boldsymbol{\delta}) \approx (\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta})' \ (\boldsymbol{\partial} f/\boldsymbol{\partial\delta}) \tag{13}$$

This is a linear function of a multinormally distributed random vector; hence its distribution is normal, $E(f(\hat{\boldsymbol{\delta}}_n) \approx f(\boldsymbol{\delta})$ and

$$V(f(\hat{\boldsymbol{\delta}}_n)) \approx n^{-1} \ [(\boldsymbol{\partial} f/\boldsymbol{\partial\delta})' \ \boldsymbol{\Sigma}(\boldsymbol{\delta}) \ (\boldsymbol{\partial} f/\boldsymbol{\partial\delta})]$$

As an example, consider the two-variable regression model

$$y - \bar{y} = b(x - \bar{x}) + c(z - \bar{z}) + e$$

and suppose that the product bc is of interest. Let $\hat{\mathbf{b}}_n = (\hat{b}_n, \hat{c}_n)'$ and let $\mathbf{b} = (b, c)'$. Suppose that

$$n^{1/2}(\hat{\mathbf{b}}_n - \mathbf{b}) \xrightarrow{\mathscr{L}} N\left(\mathbf{0}, \begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{12} & \sigma^{22} \end{pmatrix}\right) \qquad (14)$$

Application of the multivariate-delta method then yields the conclusion that

$$n^{1/2}(\hat{b}_n\hat{c}_n - bc) \xrightarrow{\mathscr{L}} N(0, c^2\sigma^{11} + 2bc\sigma^{12} + b^2\sigma^{22}) \qquad (15)$$

Note, however, that if both b and c are zero the variance term vanishes and the initial conditions for application of the delta method are not met.

Obtaining Asymptotic Confidence Intervals for the Indirect Effects

To obtain asymptotic confidence intervals for the indirect effects, their asymptotic distribution must first be determined; a general form of the multivariate-delta method is used to obtain this result. Let $\mathbf{F} = (f_1, \ldots, f_I)'$ be a vector-valued function that is differentiable in a neighborhood of an S-dimensional parameter vector $\boldsymbol{\delta}$, and let $(\partial\mathbf{F}/\partial\boldsymbol{\delta})$ be the matrix

$$\begin{bmatrix} \partial f_1/\partial\delta_1, & \ldots, & \partial f_1/\partial\delta_S \\ \vdots & & \vdots \\ \partial f_I/\partial\delta_1, & \ldots, & \partial f_I/\partial\delta_S \end{bmatrix}$$

If

$$n^{1/2}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}) \xrightarrow{\mathscr{L}} N(\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\delta})) \qquad (16)$$

then a differentiable function $\mathbf{F}$ satisfies

$$n^{1/2}(\mathbf{F}(\hat{\boldsymbol{\delta}}_n) - \mathbf{F}(\boldsymbol{\delta})) \xrightarrow{\mathscr{L}} N(\mathbf{0}, (\partial\mathbf{F}/\partial\boldsymbol{\delta})\boldsymbol{\Sigma}(\boldsymbol{\delta})(\partial\mathbf{F}/\partial\boldsymbol{\delta})') \qquad (17)$$

where the rank of the covariance matrix in (17) is less than or equal to I (Rao, 1973, p. 388).

Obtaining the asymptotic distribution of the indirect effects in a recursive model is now a simple matter. Let $\mathbf{F} =$

$(f_1, \ldots, f_I)'$ now be a column vector of indirect effects. Because the f_i are sums of products of the structural parameters, they are differentiable in a neighborhood of $\boldsymbol{\delta}$. Thus results (5) and (17) may be combined, yielding the conclusion that

$$n^{1/2}(\mathbf{F}(\hat{\boldsymbol{\delta}}_n) - \mathbf{F}(\boldsymbol{\delta})) \xrightarrow{\mathscr{L}} N(\mathbf{0}, (\boldsymbol{\partial}\mathbf{F}/\boldsymbol{\partial\delta})[\mathbf{Q}(\boldsymbol{\delta})]^{-1}(\boldsymbol{\partial}\mathbf{F}/\boldsymbol{\partial\delta})') \quad (18)$$

That is, for large n the indirect effects in a recursive model have the approximate distribution

$$N(\mathbf{F}(\boldsymbol{\delta}), n^{-1}(\boldsymbol{\partial}\mathbf{F}/\boldsymbol{\partial\delta})\,[\mathbf{Q}(\boldsymbol{\delta})]^{-1}(\boldsymbol{\partial}\mathbf{F}/\boldsymbol{\partial\delta})') \quad (19)$$

To obtain asymptotic $(1 - \alpha)100\%$ confidence intervals for the components of $\mathbf{F}$, let $Z \sim N(0, 1)$ and let $\Phi(z^*) = 1 - \alpha/2$, where Φ is the distribution function of Z. Then, because the elements in $\partial\mathbf{F}/\partial\boldsymbol{\delta}$ and $[\mathbf{Q}(\boldsymbol{\delta})]^{-1}$ are continuous near $\boldsymbol{\delta}$, a $(1 - \alpha)100\%$ confidence interval for the ith component of $\mathbf{F}$, $i = 1, \ldots, I$, is given by $f_i(\hat{\boldsymbol{\delta}}) \pm w_{ii}^{1/2}z^*$; w_{ii} is the ith diagonal element of

$$n^{-1}(\boldsymbol{\partial}\mathbf{F}/\boldsymbol{\partial}\hat{\boldsymbol{\delta}})\,[\mathbf{Q}(\hat{\boldsymbol{\delta}})]^{-1}(\boldsymbol{\partial}\mathbf{F}/\boldsymbol{\partial}\hat{\boldsymbol{\delta}})'$$

where $\hat{\boldsymbol{\delta}}$ is used to indicate that the matrices are evaluated at the solution $\hat{\boldsymbol{\delta}}$ (Rao, 1973, pp. 388–389).[5]

[5] It should now be clear how to extend the results beyond the recursive case. In the more general case, the log-likelihood function for system (2) no longer takes form (4) but may be written as

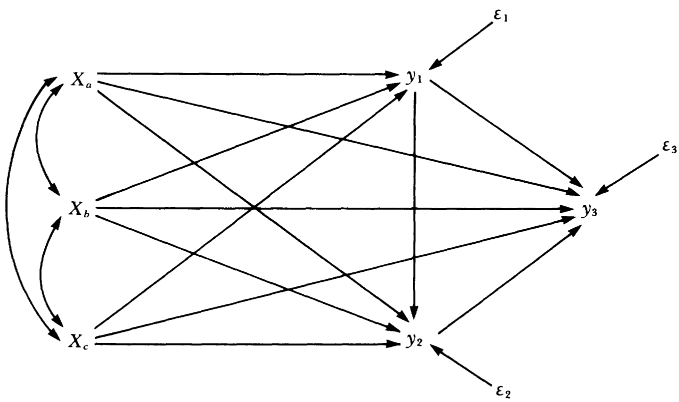
$$c - (n/2) \log |(\mathbf{B}')^{-1}\boldsymbol{\Psi}\mathbf{B}^{-1}| - (\tfrac{1}{2}) \sum_{m'=1}^{M} \sum_{m=1}^{M} \Psi^{m'm}(\mathbf{y}_{m'} - \mathbf{Z}_{m'}\boldsymbol{\delta}_{m'})'\,(\mathbf{y}_m - \mathbf{Z}_m\boldsymbol{\delta}_m)$$

where $\Psi^{m'm}$ is the $(m'm)$th element of $\boldsymbol{\Psi}^{-1}$ (Theil, 1971, p. 525). If $\boldsymbol{\Psi}$ is not diagonal, maximization of the likelihood function with respect to the elements of $\mathbf{B}$, $\boldsymbol{\Gamma}$, and $\boldsymbol{\Psi}$ leads to the full-information maximum-likelihood estimator (Jöreskog, 1973; Rothenberg and Leenders, 1964), which is asymptotically equivalent to the three-stage least-squares estimator (Theil, 1971, p. 526; Sargan, 1964). If $\boldsymbol{\Psi}$ is diagonal, maximization of the likelihood function with respect to the elements of $\mathbf{B}$, $\boldsymbol{\Gamma}$, and $\boldsymbol{\Psi}$ yields the limited-information maximum-likelihood estimator, which is asymptotically equivalent to the two-stage least-squares estimator (Theil, 1971, p. 507). Thus either (1) three-stage least squares or full-information maximum likelihood or (2) two-stage least squares or limited-information maximum likelihood may be used to obtain the estimates and the inverted information matrix of Equation (3). (For mathematical expressions for the elements in the information matrix,

AN EXAMPLE

To illustrate how asymptotic confidence intervals for indirect effects may be obtained, the example from Alwin and Hauser (1975), originally taken from the correlation matrix in Duncan, Featherman, and Duncan (1972, p. 38), is reworked. See Figure 1 for the path diagram. The variables in this model of the achievement process are father's occupational status (X_a), father's education (X_b), number of siblings in the family of ori-

Figure 1. Path diagram for a model of socioeconomic achievement.



see Jöreskog, 1973, p. 112; Theil, 1971, p. 526; and Rothenberg and Leenders, 1964, p. 67.) Provided the indirect effects are differentiable, confidence intervals may be obtained as before.

For models with unobserved variables, LISREL may be used to obtain the maximum-likelihood estimates and the inverse of the information matrix. (See Jöreskog, 1973, p. 109, for formulas.) Provided the indirect effects are differentiable, confidence intervals may be obtained as before. (On the subject of indirect effects in models with unobserved variables, see Schmidt, 1980.)

entation (X_c), respondent's education (y_1), respondent's occupational status in March 1962 (y_2), and respondent's income, expressed in units of thousands, in 1961 (y_3). The results apply to nonblack men with nonfarm background in the experienced civilian labor force, of age 35 to 44 in March 1962.[6]

Table 1 presents the coefficients for the structural model and Table 2 displays the estimated indirect effects in both symbolic and numerical form.[7] In addition, Table 2 contains the asymptotic standard errors of the indirect effects.

The standard errors in Table 2 were obtained as follows. First the f_i, $i = 1, \ldots, 10$, are differentiated with respect to the structural parameters, yielding the 10×12 matrix $(\boldsymbol{\partial}\mathbf{F}/\boldsymbol{\partial\delta})$, with ($is$)th element $(\partial f_i/\partial \delta_s)$, evaluated at the solution $\hat{\boldsymbol{\delta}}$. In symbolic form, Table 3 presents this matrix of partial derivatives; the parameter at the top of each column indicates that the partial derivative is taken with respect to that parameter. Next the estimated variance-covariance matrix of $\hat{\boldsymbol{\delta}}$ is premultiplied by the matrix of Table 3 and postmultiplied by its transpose, yielding the estimated asymptotic variance-covariance matrix of the indirect effects (Table 4).[8] From here it

[6] All the results are presented for the unstandardized form of the model. However, the results that have been established can be modified to hold asymptotically for the standardized solution.

[7] Readers of the Alwin and Hauser paper may wonder why the simpler expressions for the indirect effects they use are not employed here. Their expressions are a function of both structural and reduced-form parameters; in general, to work with these expressions in a distributional framework one needs the joint distribution of the reduced-form parameters and the structural parameters. Furthermore, the formulas given by Alwin and Hauser are valid for only a limited class of recursive models, as the authors note.

Similarly, the expressions given by Fox (1980, p. 11) are not used because the indirect effects he defines are "total" indirect effects—that is, the indirect effect of a variable on a subsequent variable is given as the sum of the indirect effects through all the intervening variables and combinations of the intervening variables. Here I prefer to allow the possibility that users may also find it worthwhile to examine particular components of "total" indirect effects. See, for example, the illustration in this section.)

[8] The estimated asymptotic covariance matrix is 12×12, as opposed to 15×15, because the variables in the analysis were, without loss of generality, deviated about their means. This is also the reason why the matrix of partial derivatives is 10×12 rather than 10×15.

TABLE 1
COEFFICIENTS AND STANDARD ERRORS (IN PARENTHESES) FOR THE STRUCTURAL EQUATION MODEL OF SOCIOECONOMIC ACHIEVEMENT IN FIGURE 1

Independent Variable	Dependent Variable y_1	y_2	y_3
X_a	$\hat{\delta}_{1a} =$ 0.0385	$\hat{\delta}_{2a} =$ 0.1352	$\hat{\delta}_{3a} =$ 0.0114
	(0.0025)	(0.0175)	(0.0045)
X_b	$\hat{\delta}_{1b} =$ 0.1707	$\hat{\delta}_{2b} =$ 0.0490	$\hat{\delta}_{3b} =$ 0.0712
	(0.0156)	(0.1082)	(0.0275)
X_c	$\hat{\delta}_{1c} =$ −0.2281	$\hat{\delta}_{2c} =$ −0.4631	$\hat{\delta}_{3c} =$ −0.0373
	(0.0176)	(0.1231)	(0.0314)
y_1		$\hat{\delta}_{21} =$ 4.3767	$\hat{\delta}_{31} =$ 0.1998
		(0.1202)	(0.0364)
y_2			$\hat{\delta}_{32} =$ 0.0704
			(0.0045)

NOTE: The sample size is 3,214; the population is nonblack, nonfarm, U.S. men in the experienced civilian labor force, aged 35–44 in March 1962. Symbols are as described in the text and Figure 1. The analysis is based on the correlation matrix and the standard deviations reported by Duncan, Featherman, and Duncan (1972, p. 38).

TABLE 2
INDIRECT EFFECTS AND ASYMPTOTIC STANDARD ERRORS (IN PARENTHESES) FOR THE STRUCTURAL EQUATION MODEL OF FIGURE 1

Dependent Variable	Indirect Effects of Independent Variables	Indirect Effect Through y_1 Alone	y_2
y_2	X_a	$\hat{f}_1 = \hat{\delta}_{21}\hat{\delta}_{1a} =$ 0.1685	
		(0.0118)	
	X_b	$\hat{f}_2 = \hat{\delta}_{21}\hat{\delta}_{1b} =$ 0.7471	
		(0.0712)	
	X_c	$\hat{f}_3 = \hat{\delta}_{21}\hat{\delta}_{1c} =$ −0.9983	
		(0.0818)	
y_3	X_a	$\hat{f}_4 = \hat{\delta}_{31}\hat{\delta}_{1a} =$ 0.0077	$\hat{f}_7 = \hat{\delta}_{32}(\hat{\delta}_{2a} + \hat{\delta}_{21}\hat{\delta}_{1a}) =$ 0.0214
		(0.0015)	(0.0020)
	X_b	$\hat{f}_5 = \hat{\delta}_{31}\hat{\delta}_{1b} =$ 0.0341	$\hat{f}_8 = \hat{\delta}_{32}(\hat{\delta}_{2b} + \hat{\delta}_{21}\hat{\delta}_{1b}) =$ 0.0560
		(0.0069)	(0.0096)
	X_c	$\hat{f}_6 = \hat{\delta}_{31}\hat{\delta}_{1c} =$ −0.0456	$\hat{f}_9 = \hat{\delta}_{32}(\hat{\delta}_{2c} + \hat{\delta}_{21}\hat{\delta}_{1c}) =$ −0.1028
		(0.0090)	(0.0111)
	y_1		$\hat{f}_{10} = \hat{\delta}_{32}\hat{\delta}_{21} =$ 0.3081
			(0.0214)

TABLE 3
SYMBOLIC FORM FOR $\partial\mathbf{F}/\partial\boldsymbol{\delta}$

	δ_{1a}	δ_{1b}	δ_{1c}	δ_{2a}	δ_{2b}	δ_{2c}	δ_{21}	δ_{3a}	δ_{3b}	δ_{3c}	δ_{31}	δ_{32}
f_1	δ_{21}	0	0	0	0	0	δ_{1a}	0	0	0	0	0
f_2	0	δ_{21}	0	0	0	0	δ_{1b}	0	0	0	0	0
f_3	0	0	δ_{21}	0	0	0	δ_{1c}	0	0	0	0	0
f_4	δ_{31}	0	0	0	0	0	0	0	0	0	δ_{1a}	0
f_5	0	δ_{31}	0	0	0	0	0	0	0	0	δ_{1b}	0
f_6	0	0	δ_{31}	0	0	0	0	0	0	0	δ_{1c}	0
f_7	$\delta_{32}\delta_{21}$	0	0	δ_{32}	0	0	$\delta_{32}\delta_{1a}$	0	0	0	0	$\delta_{2a} + \delta_{21}\delta_{1a}$
f_8	0	$\delta_{32}\delta_{21}$	0	0	δ_{32}	0	$\delta_{32}\delta_{1b}$	0	0	0	0	$\delta_{2b} + \delta_{21}\delta_{1b}$
f_9	0	0	$\delta_{32}\delta_{21}$	0	0	δ_{32}	$\delta_{32}\delta_{1c}$	0	0	0	0	$\delta_{2c} + \delta_{21}\delta_{1c}$
f_{10}	0	0	0	0	0	0	δ_{32}	0	0	0	0	δ_{21}

is a simple matter to pick out the diagonal elements and compute asymptotic standard errors, using these to create the $(1 - \alpha)100\%$ confidence intervals.

Table 2 indicates that while several of the indirect effects are small, in no case would a 95 or 99 percent confidence interval for any indirect effect cover the value zero.[9] Hence it may be assumed that the indirect effects are nonzero. Thus, although the direct effect of father's education on respondent's occupational status is insignificant, the indirect effect on occupational status through respondent's education is greater than zero. In other words, educational advantages in the family of orientation favorably influence occupational status by augmenting respondent's education, which in turn has a positive direct effect on occupational status. Similarly, although the direct effect of the sibling variable on income is not statistically different from zero, increasing the number of siblings in the family of orientation decreases income both by decreasing respondent's educational attainment (f_6) and by decreasing respondent's educational attainment and respondent's occupational status (f_9).

It is often of considerable theoretical and pragmatic value to compare the relative magnitudes of appropriate direct

[9] Alternatively, we could say, with 90 percent confidence, that none of the indirect effects, considered jointly, are zero (Miller, 1966).

TABLE 4

ASYMPTOTIC VARIANCE-COVARIANCE MATRIX (MULTIPLIED BY 10^4) FOR $\mathbf{F}(\hat{\boldsymbol{\delta}})$

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
f_1	1.3912									
f_2	−2.7139	50.6970								
f_3	−0.2800	4.3676	66.9629							
f_4	0.0538	−0.1671	0.0449	0.0220						
f_5	−0.1671	2.1225	0.4554	0.0792	0.4825					
f_6	0.0449	0.4554	2.7148	−0.1140	−0.4942	0.8119				
f_7	0.0830	−0.2579	0.0695	−0.0058	−0.0545	0.0603	0.0385			
f_8	−0.2579	3.2740	0.7037	−0.0371	0.0373	0.1819	−0.0134	0.9177		
f_9	0.0695	0.7039	4.1868	0.0363	0.1792	−0.0053	−0.0472	0.0019	1.2297	
f_{10}	0.3902	1.7322	−2.3137	−0.1389	−0.6166	0.8236	0.2677	0.7023	−0.9210	4.5763

and indirect effects (Alwin and Hauser, 1975; Duncan, Featherman, and Duncan, 1972). There are a variety of ways in which such comparisons may be made; for example, one could assess the proportion of the total effect that is direct (indirect). Alternatively, the ratio of the direct (indirect) effect of interest to the relevant indirect (direct) effect may be assessed, yielding information on both the relative magnitude and direction of influence of the direct (indirect) effects.[10] To implement such comparisons, we define the ratios r_i, $i = 1, \ldots, I$, as $g_i(\boldsymbol{\delta})/f_i(\boldsymbol{\delta})$, where $f_i(\boldsymbol{\delta})$ is as previously defined and $g_i(\boldsymbol{\delta})$ is the direct effect of the appropriate independent variable on the dependent variable under consideration. Provided $f_i(\boldsymbol{\delta}) \neq 0$, the r_i are differentiable and the multivariate-delta method may be used to obtain asymptotic confidence intervals for the ratios of interest. For example, let $r_1 = \delta_{2a}/f_1(\boldsymbol{\delta})$. From Table 2 it is clear that the null hypothesis $f_1(\boldsymbol{\delta}) = 0$ can be rejected at the 0.01 level; thus r_1 may be assumed to be differentiable. Differentiating r_1 with respect to the elements in $\boldsymbol{\delta}$ yields the 1×12 row vector with

$$\partial r_1/\partial\delta_{1a} = -\delta_{2a}/\delta_{21}\delta_{1a}^2$$

$$\partial r_1/\partial\delta^{2a} = (\delta_{21}\delta_{1a})^{-1}$$

$$\partial r_1/\partial\delta_{21} = -\delta_{2a}/\delta_{1a}\delta_{21}^2$$

and all other partials equal to zero. Premultiplication of this row vector into the estimated asymptotic variance-covariance matrix of the structural coefficients and postmultiplication by

[10] Alternatively, one could examine differences between direct and indirect effects. In general, the manner in which the comparisons are defined should depend both on substantive context and statistical tractability. To appreciate the latter point, suppose that we had defined the $r_i, i = 1, \ldots, I$, as the ratio of the indirect effect to the direct effect. Since the null hypotheses that $\delta_{2b} = 0$ and $\delta_{3c} = 0$ have not been rejected at the 0.05 level (see Table 1), it is not reasonable to allow the comparisons that use these expressions in the denominator: If δ_{2b} and δ_{3c} are zero, the differentiability conditions are violated. However, because it has been established that the indirect effects may be assumed nonzero, it is permissible to formulate all comparisons that use these expressions in the denominator of the ratios, as we have done. Of course, when the direct effect is zero, the ratio of the direct to the indirect effect should not deviate significantly from zero.

the transpose of the row vector yields the estimated asymptotic variance of $\hat{r}_1$. It is now a simple matter to construct $(1 - \alpha)100\%$ confidence intervals for the r_i, $i = 1, \ldots, I$.

Table 5 presents values of the $\hat{r}_i$ and their asymptotic standard errors for $i = 1, \ldots, 10$; Table 6 presents 95 and 99 percent confidence intervals for the r_i. Moreover, Table 6 presents confidence intervals for three additional functions. First, the ratio of the direct effect δ_{3a} to the sum of the indirect effects f_4 and f_7 is defined as r_{11}. Clearly r_{11} is simply the ratio of the direct effect of X_a on y_3 to the "total" indirect effect of X_a on y_3—that is, the indirect effect that operates jointly through y_1 alone, through y_2 alone, and through y_1 and y_2. Similarly, r_{12} is defined as $\delta_{3b}/f_5 + f_8$ and r_{13} is defined as $\delta_{3c}/f_6 + f_9$.

Inspection of Table 6 reveals that r_2, r_6, r_9, and r_{13} may be assumed to equal zero; this result recapitulates the earlier observation that δ_{2b} and δ_{3c} are not statistically different from zero. In addition, the null hypotheses that r_3 and r_{11} are less than 1 are not rejected, at either the 0.05 or 0.01 level, on the

TABLE 5
RATIOS OF DIRECT TO INDIRECT EFFECTS AND ASYMPTOTIC STANDARD ERRORS (IN PARENTHESES) FOR THE STRUCTURAL EQUATION MODEL OF FIGURE 1

Dependent Variable	Indirect Effects of Independent Variables	Indirect Effect Through y_1 Alone	Indirect Effect Through y_2
y_2	X_a	$\hat{r}_1 = \hat{\delta}_{2a}/\hat{\delta}_{21}\hat{\delta}_{1a} = 0.8024$ (0.1231)	
	X_b	$\hat{r}_2 = \hat{\delta}_{2b}/\hat{\delta}_{21}\hat{\delta}_{1b} = 0.0656$ (0.1452)	
	X_c	$\hat{r}_3 = \hat{\delta}_{2c}/\hat{\delta}_{21}\hat{\delta}_{1c} = 0.4639$ (0.1318)	
y_3	X_a	$\hat{r}_4 = \hat{\delta}_{3a}/\hat{\delta}_{31}\hat{\delta}_{1a} = 1.4820$ (0.6860)	$\hat{r}_7 = \hat{\delta}_{3a}/\hat{\delta}_{32}(\hat{\delta}_{2a} + \hat{\delta}_{21}\hat{\delta}_{1a}) = 0.5332$ (0.2205)
	X_b	$\hat{r}_5 = \hat{\delta}_{3b}/\hat{\delta}_{31}\hat{\delta}_{1b} = 2.0876$ (0.9631)	$\hat{r}_8 = \hat{\delta}_{3b}/\hat{\delta}_{32}(\hat{\delta}_{2b} + \hat{\delta}_{21}\hat{\delta}_{1b}) = 1.2704$ (0.5376)
	X_c	$\hat{r}_6 = \hat{\delta}_{3c}/\hat{\delta}_{31}\hat{\delta}_{1c} = 0.8184$ (0.7298)	$\hat{r}_9 = \hat{\delta}_{3c}/\hat{\delta}_{32}(\hat{\delta}_{2c} + \hat{\delta}_{21}\hat{\delta}_{1c}) = 0.3625$ (0.3098)
	y_1		$\hat{r}_{10} = \hat{\delta}_{31}/\hat{\delta}_{32}\hat{\delta}_{21} = 0.6485$ (0.1451)

TABLE 6
ASYMPTOTIC CONFIDENCE INTERVALS FOR RATIOS OF DIRECT TO INDIRECT EFFECTS IN A MODEL OF SOCIOECONOMIC ACHIEVEMENT

	95% CI	99% CI
r_1	(0.5827,1.0221)	(0.4848,1.1200)
r_2	(−0.2190,0.3502)	(−0.3090,0.4402)
r_3	(0.2056,0.7222)	(0.1239,0.8039)
r_4	(0.1374,2.8266)	(−0.2879,3.2519)
r_5	(0.1999,3.9753)	(−0.4839,4.5724)
r_6	(−0.6120,2.2488)	(−1.0645,1.8828)
r_7	(0.1010,0.9654)	(−0.0357,1.1021)
r_8	(0.2167,2.3241)	(−0.1166,2.6574)
r_9	(−0.2447,0.9697)	(−0.4368,1.2164)
r_{10}	(0.3641,0.9329)	(0.2741,1.0229)
r_{11}	(0.0729,0.7107)	(−0.0280,0.8116)
r_{12}	(0.1407,1.4397)	(−0.0648,1.6452)
r_{13}	(−0.1723,0.6745)	(−0.3062,0.8084)

basis of these data. Thus the indirect effect of the sibling variable on respondent's occupational status is more negative than the direct effect, indicating that additional siblings in the family of orientation detract from subsequent occupational attainments predominantly by attenuating the educational experiences of individuals reared in larger families. Similarly, the effect of father's occupational status on respondent's income is predominantly indirect: Respondent's income is enhanced by high-status origins, primarily because high-status origins increment subsequent educational and occupational achievements.

The widths of the confidence intervals do not appear to allow strong inferences about the magnitudes of the other comparisons in Table 6. Nevertheless, it is not unreasonable to suggest that r_1, r_7, and r_{10} are not greater than 1. In other words, the direct effects of father's status on son's status and income are not larger than the effects that intervene through son's education and son's occupational status, respectively; similarly, the direct effect of respondent's education on income is not larger than the indirect effect of education on income.

On balance, the analysis suggests that the indirect effects in the model under consideration are nonzero, but the confidence intervals for the ratios in Table 6 are too large to permit precise inferences about the relative magnitudes of the direct and indirect effects. In turn, the example illustrates the utility of computing the confidence intervals and suggests that researchers proceed with both caution and thoughtfulness before offering detailed substantive interpretations of the various effects in complex causal models.

SUMMARY

In the preceding pages I have proposed a method for assessing the significance of indirect effects in structural equation models and given an example. The method itself is quite simple, and it is easy to perform the necessary computations. All one needs is a regression program that computes the variance-covariance matrix of the coefficients (for example, BMDP or LISREL), a calculator for computing the elements of the matrix of partial derivatives, and a matrix multiplication program for computing the estimated asymptotic variance-covariance matrix of the indirect effects. Alternatively, the computations can be performed in one step with a good matrix algebra program (for example, PROC MATRIX in SAS).

One caveat: The confidence intervals derived here are valid for large samples. Since one seldom knows when a sample is large enough, the application of these methods may be inappropriate in particularly small samples.

APPENDIX

In this appendix we derive the asymptotic distribution of the complete coefficient vector in a recursive model. We begin with the log-likelihood function

$$L(\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_M, \boldsymbol{\Psi}) = c + (n/2) \sum_{j=1}^{M} \log \Psi_{jj}^{-1} - (\tfrac{1}{2}) \sum_{j=1}^{M} (\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j)' (\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j) \Psi_{jj}^{-1} \tag{A-1}$$

which is identical to (4) in the text.

Following Rothenberg and Leenders (1964, pp. 60–61), we differentiate (A-1) with respect to $\Psi_{jj}^{-1}, j = 1, \ldots, M$, obtaining the conclusion that in order for (A-1) to be maximized, it must be the case that

$$\Psi_{jj} = n^{-1}(\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j)'(\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j) = \hat{\Psi}_{jj} \qquad j = 1, \ldots, M$$

Substitution of this result into (A-1) yields the concentrated log-likelihood function

$$L(\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_M) = c' - (n/2)\sum_{j=1}^{M} \log \hat{\Psi}_{jj} \tag{A-2}$$

which may be maximized in lieu of (A-1).[11] Differentiating (A-2) with respect to $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_M$ yields the first-order derivatives

$$\partial L/\partial\boldsymbol{\delta}_j = \hat{\Psi}_{jj}^{-1}(\mathbf{Z}_j'(\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j)) \qquad j = 1, \ldots, M \tag{A-3}$$

and repeated differentiation yields

$$\partial^2 L/\partial\boldsymbol{\delta}_j\partial\boldsymbol{\delta}_h' = \begin{cases} \mathbf{0} & \text{if } j \neq h \\ -\hat{\Psi}_{jj}^{-1}\mathbf{Z}_j'\mathbf{Z}_j + 2(n\hat{\Psi}_{ji}^2)^{-1}(\mathbf{Z}_j'(\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j) & \\ \qquad (\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j)'\mathbf{Z}_j) = \mathbf{G}_j & \text{if } j = h \end{cases} \tag{A-4}$$

Let $\boldsymbol{\delta} = (\boldsymbol{\delta}_1', \ldots, \boldsymbol{\delta}_M')'$ as in the text; then combining the elements of (A-4) gives

$$\boldsymbol{\partial}^2\mathbf{L}/\boldsymbol{\partial\delta\partial\delta}' = \begin{bmatrix} \mathbf{G}_1 & \cdot & \cdot & \cdot & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_2 & \cdot & \cdot\ \cdot & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_3 & \cdots & \mathbf{0} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_M \end{bmatrix}$$

It is well known that, under suitable regularity conditions,

$$\operatorname{plim}(n^{-1}(\boldsymbol{\partial}^2\mathbf{L}/\boldsymbol{\partial\delta\partial\delta}')) = -\mathbf{I}(\boldsymbol{\delta})$$

where **I** is the Fisher information matrix (Rao, 1973, p. 366).

[11] In effect we are solving for a subset of the parameters and using the solution to simplify the maximization problem. For a complete treatment of concentrated likelihood functions, see Dhrymes (1974, pp. 324–334).

Suppose that $\hat{\Psi}_{jj}^{-1}(\mathbf{Z}_j'\mathbf{Z}_j/n)$ has plim $Q_j \neq \mathbf{0}$, $j = 1, \ldots, M$. Since

$$\text{plim } n^{-1}(\mathbf{Z}_j'(\mathbf{y}_j - \mathbf{Z}_j\boldsymbol{\delta}_j)) = \mathbf{0}$$

from (A-4) we readily obtain

$$\mathbf{I}(\boldsymbol{\delta}) = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_3 & \cdots & \mathbf{0} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Q}_M \end{bmatrix}$$

Next we use the fact that under general regularity conditions (Theil, 1971, p. 395)

$$n^{1/2}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta}) \xrightarrow{\mathscr{L}} N(\mathbf{0}, [I(\boldsymbol{\delta})]^{-1}) \tag{A-5}$$

That is, $n^{1/2}(\hat{\boldsymbol{\delta}}_n - \boldsymbol{\delta})$ converges in law (distribution) to the quantity on the right-hand side of (A-5), where $[\mathbf{I}(\boldsymbol{\delta})]^{-1}$ is the block-diagonal matrix

$$\begin{bmatrix} \mathbf{Q}_1^{-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_3^{-1} & & \mathbf{0} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Q}_M^{-1} \end{bmatrix}$$

This suffices to show the result.

REFERENCES

ALWIN, D. F., AND HAUSER, R. M.

1975 "The decomposition of effects in path analysis." *American Sociological Review* 40:37–47.

BIELBY, W. T., AND HAUSER, R. M.

1977 "Structural equation models." *Annual Review of Sociology* 3:137–161.

BISHOP, Y. M., FIENBERG, S. E., AND HOLLAND, P. W.
1975 *Discrete Multivariate Analysis: Theory and Practice.* Cambridge, Mass.: M.I.T. Press.

BLALOCK, H. M. (ED.)
1971 *Causal Models in the Social Sciences.* Chicago: Aldine.

BLAU, P. M., AND DUNCAN, O. D.
1967 *The American Occupational Structure.* New York: Wiley.

DHRYMES, P. J.
1974 *Econometrics.* New York: Springer-Verlag.

DUNCAN, O. D.
1966 "Path analysis: sociological examples." *American Journal of Sociology* 72:1–16.
1975 *Introduction to Structural Equation Models.* New York: Academic Press.

DUNCAN, O. D., FEATHERMAN, D. L., AND DUNCAN, B.
1972 *Socioeconomic Background and Achievement.* New York: Seminar Press.

FINNEY, J. M.
1972 "Indirect effects in path analysis." *Sociological Methods and Research* 1:175–186.

FOX, J.
1980 "Effect analysis in structural equation models." *Sociological Methods and Research* 9:3–28.

HEISE, D. R.
1975 *Causal Analysis.* New York: Wiley.

JÖRESKOG, K. G.
1973 "A general method for estimating a linear structural equation system." In A. S. Goldberger and O. D. Duncan (eds.), *Structural Equation Models in the Social Sciences.* New York: Academic Press.

LAND, K. C.
1969 "Principles of path analysis." In E. F. Borgatta (ed.), *Sociological Methodology 1969.* San Francisco: Jossey-Bass.

LEWIS-BECK, M. S.
1974 "Determining the importance of an independent variable: A path analytic solution." *Social Science Research* 3:95–108.

MALINVAUD, E.
1970 *Statistical Methods of Econometrics.* Amsterdam: North-Holland.

MILLER, R.
1966 *Simultaneous Statistical Inference.* New York: McGraw-Hill.

RAO, C. R.
1973 *Linear Statistical Inference and Its Applications.* New York: Wiley.

ROTHENBERG, T. J., AND LEENDERS, C. T.
1964 "Efficient estimation of simultaneous equatión systems." *Econometrica* 32:57–76.

SARGAN, J. D.
1964 "Three-stage least-squares and full maximum-likelihood estimates." *Econometrica* 32:77–81.

SCHMIDT, P.
1980 "On decomposition of effects in causal models." Mannheim: Zentrum für Umfragen, Methoden und Analysen. Unpublished manuscript.

THEIL, H.
1971 *Principles of Econometrics.* New York: Wiley.