8 Causal Analysis I

In Chapters 2 through 4, we covered the basic statistical concepts of regression analysis, including both descriptive and inferential statistics. These concepts constitute the core of linear regression analysis. In Chapters 5 through 7, we saw how to deal with three common problems in linear regression—nominal independent variables, nonlinear relationships, and nonadditive relationships. No new statistical concepts were introduced in these chapters. Instead, the focus was on how to use transformations of variables—such as dummy variables, powers of variables, and products between variables—in order to incorporate nominal variables, nonlinearity, and nonadditivity into regression equations.

In Chapters 8 and 9 we move beyond the material that is customarily presented as regression analysis proper. The perspective of this book is that multiple regression is well suited and commonly used for nonexperimental causal analysis—that is, for explanation. When the variables that are included in a multiple regression equation are carefully selected (see the discussion in Chapter 1) the partial slopes may be treated as estimates of the effects of two or more independent variables on a single dependent variable. The final two chapters extend this perspective to systems of equations or *multi-equation* causal models. The multi-equation models with which we will be involved consist of two or more dependent variables (*endogenous variables*), at least one of which must be an independent variable with respect to one or more of the other endogenous variables (see Figure 1.2). The analysis will be restricted to linear and additive equations.

In this chapter we will see how to use multi-equation models to calculate a new kind of causal effect, an *indirect effect*. In single-equation models only

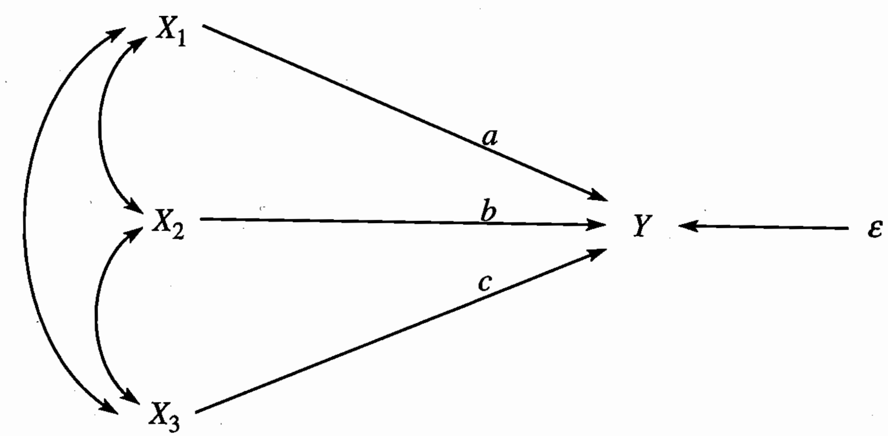
direct effects exist. The distinction between direct and indirect effects also allows us to calculate a *total effect*, which is the sum of the direct effect and all indirect effects of one variable on another. Ordinary least-squares regression can be used to estimate the parameters (effects) in *recursive* multi-equation models but not in *nonrecursive* models (i.e., models with feedback loops). Thus, learning to distinguish between recursive and nonrecursive models and understanding why OLS is invalid for estimating the parameters of nonrecursive models are also important objectives of this chapter. Before taking up the analysis of multi-equation models, however, it will be helpful to look more closely at some characteristics of single-equation models.

Single-Equation Models

Multiple regression analysis may be used to estimate the effects of several different independent variables on a single dependent variable. A single multiple regression equation represents the simplest type of causal model. It is often helpful to use causal diagrams, or path diagrams, to show the structure of causal models, especially when we are concerned with more complex models that will be introduced shortly. The path diagram for a causal model that would be estimated with a regression equation containing three independent variables is shown in Figure 8.1.

Some of the rules and principles that are used in path diagrams were presented in the section on causal diagrams in Chapter 1. You should review that section. In Figure 8.1, the single-headed arrows represent causal paths, or effects, between variables. The double-headed curved lines represent covariances or correlations between variables. The dependent variable Y is an en-

FIGURE 8.1 Path Diagram for a Single-Equation Causal Model



dogenous variable; the X's and ε *are exogenous* variables. The lower case letters a, b, and c on the causal paths represent the effect of each X on Y (the change in Y per unit increase in X). These letters are used instead of the usual β's because the subscript notation for the β's becomes somewhat unwieldy when we are dealing with more complex models. The causal effects of the X's are estimated with the following regression equation:

$$\hat{Y} = a + b_1X_1 + b_2X_2 + b_3X_3 \tag{8.1}$$

The intercept of the regression equation usually is not shown in the path diagram. The estimates of a, b, and c in the diagram are b_1, b_2, and b_3, respectively. Since the b's are unstandardized slopes, this indicates that the effects shown in the diagram are unstandardized effects. In the causal modeling literature, unstandardized effects are often called **structural coefficients.** Because all of the causal effects are estimated by Equation 8.1, the model shown in Figure 8.1 is called a *single-equation model.*

Notice that in Figure 8.1 there is no symbol on the path from the error term to Y. Since the error term represents the effects of unmeasured variables, we have to make an assumption about the scale of this hypothetical construct before we can specify the value of its effect on Y. The convention is to assume that ε has the same scale as Y; e.g., if Y is measured in dollars, then the scale for ε is also taken to be dollars. The consequence of this assumption is that the effect of a unit change in ε is equal to unity; e.g., if ε increases by one dollar, then Y increases by one dollar. Since the effect of ε is equal to unity, it can be omitted from the path diagram.

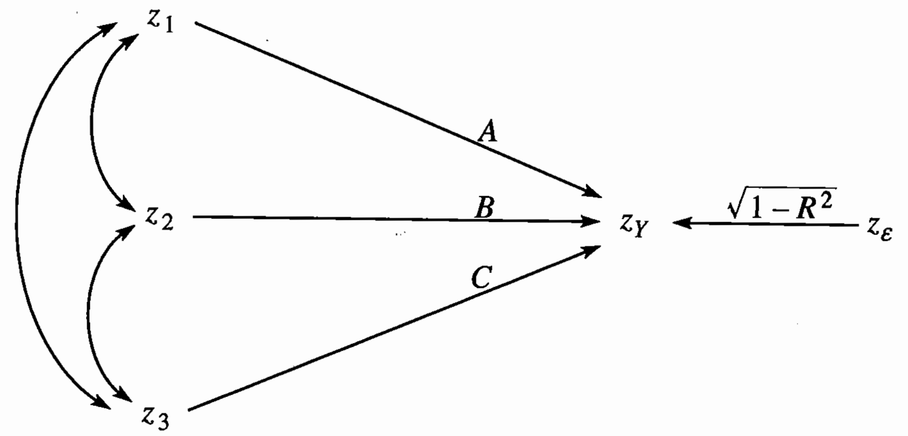
Causal models may also be specified in terms of standardized variables and coefficients (Figure 8.2). In this diagram, A, B, and C represent the causal effects of the standardized variables (the z_i). The standardized effects A, B, and C would be estimated with the standardized regression equation

$$\hat{z}_Y = B_1z_1 + B_2z_2 + B_3z_3 \tag{8.2}$$

The standardized effects for a causal model are often called **path coefficients** because the original principles of path analysis developed by biologist Sewell Wright (1921) were based entirely on standardized coefficients. Path coefficients are often symbolized by p_i. We will not use this notation, however. Instead, we will use our normal symbols B_i to represent the estimates of the standardized effects in causal models.

There is another difference between the diagrams for standardized and unstandardized models. In a standardized model we treat all of the variables as being standardized to have a unit variance and standard deviation (i.e., z scores). Thus, in Figure 8.2 it is also understood that the error term z_ε has a unit variance and standard deviation. The consequence of this specification is that the effect of ε on Y will not be equal to unity, as in the unstandardized model; an effect of unity for ε would mean that an increase in ε of one standard devia-

FIGURE 8.2 Path Diagram for a Standardized Causal Model



tion would cause a one-standard-deviation increase in Y, a perfect correlation. Instead of a unity effect, the effect of the standardized ε equals $\sqrt{1 - R^2}$, i.e., the square root of the residual proportion of variance; this effect equals the multiple correlation between the unmeasured variables and Y. Thus, if the path diagram represents a standardized model, the value of $\sqrt{1 - R^2}$ would be entered on the path from z_ε to z_Y, as shown in Figure 8.2.

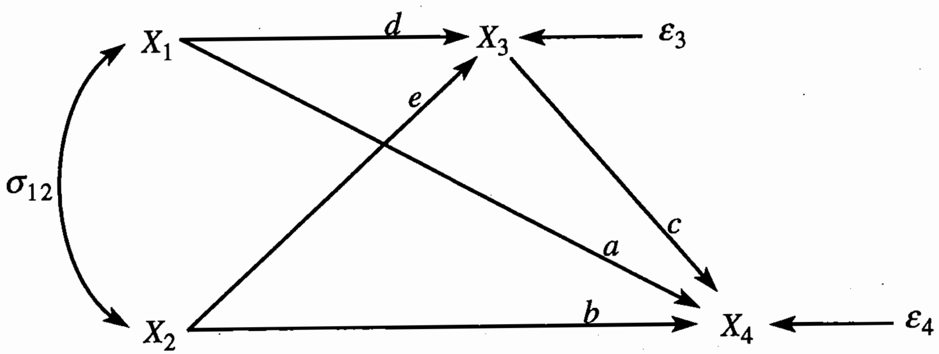
A Two-Equation Causal Model

A single-equation causal model does not include any causal effects between the independent variables; they may be correlated, but the model does not specify that these correlations are due to causal effects. If theory or known temporal sequence indicates that one or more of the X's may be dependent on one or more of the other X's, then the causal model in Figure 8.1 can be elaborated to include additional causal specifications. Let us assume that there is reason to believe that X_3 is affected by X_1 and X_2. Figure 8.3 shows this elaborated model.

Figure 8.3 contains two endogenous variables (X_3 and X_4) and four exogenous variables (X_1, X_2, ε_3, and ε_4). Because there are two endogenous variables in the model, we will no longer label one of them Y; thus, the original variable labeled Y in Figure 8.1 is now called X_4.[1] Endogenous variable X_4 is a dependent variable only; it is not specified as having an effect on any of the other variables in the model. Endogenous variable X_3 is a dependent variable rel-

1. In some notational systems, the endogenous variables are symbolized by Y's and the observed exogenous variables are represented by X's (e.g., Jöreskog and Sörbom 1989). We do not use this system because it would unduly complicate the notation in this chapter and Chapter 9.

FIGURE 8.3 Causal Model with Two Endogenous Variables



ative to X_1 and X_2, but it is an independent variable relative to X_4. Thus, this causal model reveals a new type of variable (i.e., X_3). X_3 is an *intervening* variable; it intervenes between X_1 and X_4 and also between X_2 and X_4.

How do we estimate the causal effects for the model shown in Figure 8.3? They can be estimated by running a regression for each endogenous variable (i.e., a variable that has arrows pointing at it). The variables that have arrows pointing at a particular endogenous variable will be the independent variables in the regression equation for that endogenous variable. In Figure 8.3, X_3 and X_4 will be dependent variables in two separate regression equations. The regression equations are

$$\hat{X}_4 = a_{4\cdot123} + b_{41\cdot23}X_1 + b_{42\cdot13}X_2 + b_{43\cdot12}X_3 \quad \textbf{(8.3)}$$

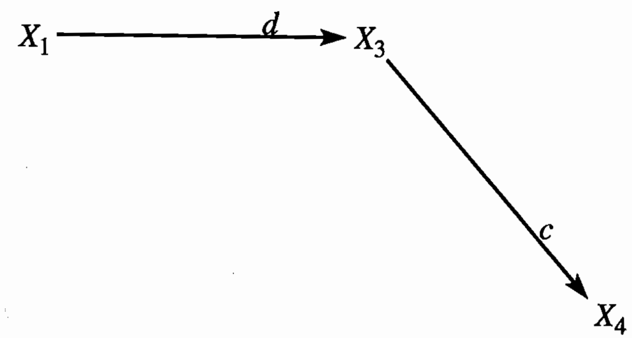
$$\hat{X}_3 = a_{3\cdot12} + b_{31\cdot2}X_1 + b_{32\cdot1}X_2 \quad \textbf{(8.4)}$$

Equation 8.3 provides the estimates of *a*, *b*, and *c* in Figure 8.3. This is the same as Equation 8.2, except that we are now calling the dependent variable X_4 instead of *Y*. This means that *a*, *b*, and *c* will have the same values in Figure 8.3 that they had in Figure 8.1. Equation 8.4 provides the estimates for *d* and *e* in Figure 8.3. These are the new parameters that were not estimated in the single-equation model. The model in Figure 8.3 is called a two-equation model because the diagram specifies two separate causal equations embedded in a single model. Finally, notice that the error terms ε_3 and ε_4 in Figure 8.3 are not correlated with X_1 and X_2 and also are not correlated with one another (there are no double-headed curved lines connecting the error terms with any variables in the model). This satisfies the regression assumption that unmeasured variables are not correlated with the independent variables in a regression equation (the absence of a correlation between ε_3 and ε_4 indicates that ε_4 is not

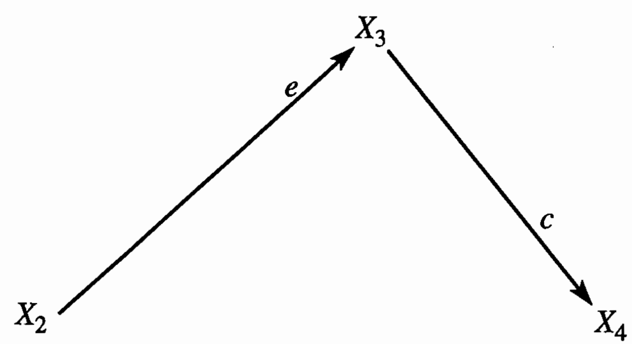
correlated with X_3). Thus, if this specification is correct, Equations 8.3 and 8.4 will provide unbiased estimates of the causal parameters in Figure 8.3.

Direct, Indirect, Total, and Spurious Effects (DITS)

What do we gain from introducing the two-equation model relative to what we would learn from the single-equation model? First, of course, we have learned about the effects of X_1 and X_2 on X_3. Just as importantly, however, we can now investigate **indirect effects**. Indirect effects are formed by compound causal paths or chains of paths in which one or more intervening variables mediate the effect of one variable on another. There are two indirect effects specified in Figure 8.3. One indirect effect is shown by the following compound path:



In this indirect effect, a change (or difference) in X_1 causes a change in X_3, and this change in X_3 then causes a change in X_4. *The notation for this indirect effect is* $\boldsymbol{I}_{431}$. The left-hand variable in the subscript is the final dependent variable in the compound path (X_4), the second is the intervening variable (X_3), and the third is the independent variable or source of the indirect effect (X_1). Thus, the order of the variables in the subscript moves backward along the path to the origin of the effect. The other indirect effect is represented as follows:



This indirect effect shows that a change in X_2 causes a change in X_3 that in turn causes a change in X_4. The notation is I_{432}.

Just as a direct effect (b or β) is defined as the change in the dependent variable per unit increase in the independent variable, *the value of an indirect effect is the change in the dependent variable at the end of a chain produced by a one-unit increase in the independent variable at the origin of the chain.* Thus, I_{431} equals ΔX_4 that is caused by ΔX_3 that is caused by $\Delta X_1 = 1$. To calculate an indirect effect, we first determine the change in the intervening variable produced by a one-unit increase in the source variable, and then we calculate the change in the final variable produced by this change in the intervening variable. Remembering that the change in a dependent variable equals the slope (effect) of the independent variable times the change in the independent variable, I_{431} and I_{432} are determined as follows:

I_{431}	I_{432}
$\Delta X_1 = 1$	$\Delta X_2 = 1$
$\Delta X_3 = d \cdot \Delta X_1 = d \cdot 1 = d$	$\Delta X_3 = e \cdot \Delta X_2 = e \cdot 1 = e$
$\Delta X_4 = c \cdot \Delta X_3 = c \cdot d$	$\Delta X_4 = c \cdot \Delta X_3 = c \cdot e$
$I_{431} = d \cdot c$	$I_{432} = e \cdot c$

The value of an indirect effect equals the product of the effects along the compound path or chain that links the source variable with the last variable in the chain. When there is one intervening variable, the indirect effect will equal the product of two coefficients, as shown above. If a compound path contains two intervening variables, the indirect effect will equal the product of three coefficients, and so on. The sign of the indirect effect will be positive if there are no negative coefficients or if there are an even number of negative coefficients. The sign of the indirect effect will be negative if there are an odd number of negative coefficients along the chain.

In addition to the indirect effects of X_1 and X_2 on X_4, the two exogenous variables also have **direct effects** on X_4, which equal a and b, respectively. *Direct effects are not mediated by any intervening variables.* We will use $\boldsymbol{D}_{ij}$ to indicate the direct effect of X_j on X_i. When an independent variable has both direct and indirect effects on a particular dependent variable, we can sum the two types of effects to determine the **total effect**, which will be designated by $\boldsymbol{T}_{ij}$. The total effects of X_1 and X_2 on X_4 are

$$T_{41} = D_{41} + I_{431} = a + dc$$

$$T_{42} = D_{42} + I_{432} = b + ec$$

If the indirect effect has the same sign as the direct effect, the total effect will be larger in absolute value than the direct effect. If the sign of the indirect effect is opposite the sign of the direct effect, the total effect will be either smaller in absolute value than the direct effect or it will be opposite in sign from the direct effect. If the intervening variable X_3 is redundant with the independent variable, the indirect effect will have the same sign as the direct effect; if suppression exists between the intervening variable and the independent variable, however, the indirect effect will be opposite in sign from the direct effect.

Not all of the causal relationships in a causal model will have both direct and indirect components. For the model in Figure 8.3, the following pairs of variables have only direct causal relationships, and thus the total effect equals the direct effect:

$$T_{43} = D_{43} = c$$

$$T_{31} = D_{31} = d$$

$$T_{32} = D_{32} = e$$

For any pair of independent and dependent variables in a causal model, the difference between the bivariate (zero-order) slope and the total effect equals the **spurious slope** or spurious "effect":

$$\text{Spurious Slope} = b_{ij} - T_{ij} = S_{ij}$$

The above equation is written in terms of the unstandardized regression slope; therefore, it pertains to a model containing structural coefficients. If we were working with a standardized model (i.e., path coefficients), we would use the standardized regression slope B_{ij} (or r_{ij}) to compute the spurious association.

The spurious component indicates the amount of the bivariate association (slope) between two variables that is not due to the effect of one variable on the other. Since indirect effects are included in the total effect, they are not counted as spurious association; indirect effects and direct effects are equally valid types of causal effects. Most commonly, the spurious component of association is thought of as having the same sign as the total effect. This would occur when the bivariate slope b_{ij} has the same sign and is larger in absolute value than the total effect. In this case, failure to control for other causes of X_i, which either are correlated with X_j or are causes of X_j, would lead to overestimates of the effects of X_j. However, if these other causes are suppressing the relationship between X_i and X_j, b_{ij} may be opposite in sign from the total effect or may have the same sign but be smaller in absolute value. In this case, the bivariate association either will be underestimating the strength of the causal effect or will be giving a wrong-signed estimate of the effect. Thus, spuriousness may involve overestimates, underestimates, or wrong-signed estimates of causal effects.

Direct (D_{ij}), indirect (I_{ij}), total (T_{ij}), and spurious (S_{ij}) effects will be referred to simply as DITS. The calculation of DITS in complex models can become tedious and error prone. We will see, however, how DITS calculations can be expedited through the use of simplified forms of causal models.

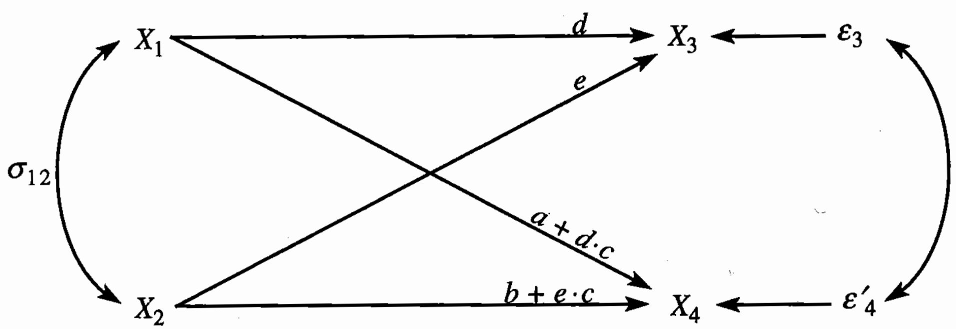
Reduced-Form Models and DITS

Causal models may be simplified by removing all of the paths leading from one endogenous variable to another. This produces a causal diagram that contains only paths from the exogenous variables to the endogenous variables. The result of this reduction is to eliminate all compound paths from the diagram. Such models are called **reduced-form models**. Reduced-form models represent only the effects of the exogenous variables. Figure 8.3 has only a single path connecting the endogenous variables, the path from X_3 to X_4. Removing this path gives Figure 8.4.

When the paths between endogenous variables are removed, the paths from the exogenous variables to the endogenous variables may no longer represent effects that are *independent* of other endogenous variables. In the full diagram of the causal system (Figure 8.3), the path from X_1 to X_4 represents an effect that is independent of X_3 because the model includes a path from X_3 to X_4. In the reduced-form model, however, there is no path from X_3 to X_4 to represent the effect of X_3; thus, the path from X_1 to X_4 does not represent an effect that is independent of X_3. Stated differently, the path from X_1 to X_4 no longer represents the *direct* effect of X_1 on X_4 because the indirect or compound path connecting X_1 to X_4 via X_3 is no longer represented.

In a reduced-form diagram, the causal path between an exogenous variable and an endogenous variable represents the direct effect plus the sum of any indirect effects that may exist between the pair of variables. In other words, *the paths in the reduced-form model represent the total effects of the exogenous variables*. In the case of Figure 8.4, since the path from X_3 to X_4 has been removed, the paths from X_1 and X_2 to X_4 stand for the total effects that were calculated from the parameters in Figure 8.3. These total effects are shown on the paths in Figure 8.4. Notice that the effects of X_1 and X_2 on X_3 in the reduced-form diagram are the same as their original direct effects; there were no intervening variables between the exogenous variables and X_3 to be eliminated in the reduced diagram.

FIGURE 8.4 Reduced-Form Model



Since the reduced-form diagram contains two endogenous variables, it is still a two-equation model. The effects in Figure 8.4 are estimated with the following regression equations:

$$\hat{X}_3 = a_{3\cdot12} + b_{31\cdot2}X_1 + b_{32\cdot1}X_2 \quad \textbf{(8.5)}$$

$$\hat{X}_4 = a_{4\cdot12} + b_{41\cdot2}X_1 + b_{42\cdot1}X_2 \quad \textbf{(8.6)}$$

There are only two X's in the equation for X_4 (8.6) because the path from X_3 to X_4 has been dropped from the diagram. In Equation 8.6 the slope $b_{41\cdot2}$ estimates the effect of X_1 on X_4 that is independent of X_2, since X_2 is in the equation and is thus held constant. This slope, however, does not estimate an effect that is independent of X_3 because X_3 is not included in the equation. The regression slopes in Equation 8.6 will equal the effects shown in Figure 8.4 (assuming no sampling error occurs).

$$b_{41\cdot2} = a + d\cdot c \qquad b_{42\cdot1} = b + e\cdot c$$

Again, since the equation for X_4 does not contain X_3, the above regression slopes will not be estimates of the direct effects of the exogenous variables. Instead, they will equal the sum of the direct effect plus the indirect effect that passes through X_3, i.e., the total effect.

It should be emphasized that the reduced-form model of Figure 8.4 does not represent an alternative causal model to that specified in Figure 8.3. It is instead a simplification of Figure 8.3 achieved by omitting the path from X_3 to X_4. The absence of a path from X_3 to X_4 does not mean that X_3 does not have an effect on X_4. It was omitted to simplify the causal analysis. A reduced-form model has meaning only when compared to the full model from which it is derived.

Notice in Figure 8.4 that the error term for X_4 is now written as ε'_4. The variance of ε'_4 will be larger than the variance of ε_4 in the three-variable equation because X_3 is not included in the equation. Thus, the variance in X_4 that is uniquely explained by X_3 is now contained in the error term ε'_4. Furthermore, since the unique variance in X_3 is the amount that is not related to X_1 and X_2, the error term for X_3 is the source of this unique variance. Therefore, ε_3 is the cause of the unique variance in X_3 that causes some of the variance in X_4 that is now summarized by ε'_4. As a consequence, ε_3 and ε'_4 will be correlated. This is shown by the doubled-headed curved line connecting ε_3 and ε'_4 in the reduced-form model in Figure 8.4. The error terms in Figure 8.4, however, are not of primary concern to us. We are principally interested in the total effects of the exogenous variables shown in the reduced-form diagram.

The model represented by Figure 8.4 is a valid model for the total effects of the exogenous variables on X_4. When these total effects are estimated with the above regression equation, however, the ability to distinguish between the direct and indirect effects is lost. However, if we estimated only the model shown in Figure 8.4 (that is, if we never estimated Figure 8.3, possibly because we did not have a measure of X_3), the estimates of the total effects of X_1 and X_2 would not be biased by failure to control for the intervening variable X_3. Discussions of regression assumptions (see Chapter 4) always emphasize that the failure to control for a variable that is a cause of Y and that is correlated with the independent variables included in the regression equation will lead to biased estimates of the effects of the variables included in the equation (e.g., $b_{41\cdot2}$ and $b_{42\cdot1}$ will be biased). This is because the effect of the omitted variable (e.g., X_3) will be part of the error term and thus will be correlated with the included X's. Consequently, part of this effect will be picked up by the b's for the included variables, and this will create biased estimates of their *direct effects*. If the omitted variable is an intervening variable, however, the amount of the bias will be equal to the indirect effects of the included X's, which are dc and ec for X_1 and X_2, respectively. We now see why we must qualify the consequences of violating the assumption that e is uncorrelated with the X's. If the omitted variables are intervening variables, their omission will not lead to biased estimates of the *total effects* of the included variables. If, on the other hand, the omitted variables are either causes of the included variables or are simply correlated with the included variables, the regression slopes will be biased.

There is a practical reason for introducing reduced-form models. Alwin and Hauser (1975) have shown how reduced-form equations can be used to expedite the calculations of indirect effects and total effects for more complex models. We will illustrate the basic principles of the Alwin-Hauser method with the relatively simple model in Figure 8.3. The method involves running a series of regression equations, including the bivariate equations, the reduced-form equation (8.6), and the full equation (8.3). The following SPSS commands may be used.

```
REGRESSION VARS = X1 X2 X3 X4/
  DEP = X4/ ENTER X1/
  DEP = X4/ ENTER X2/ ENTER X1/ ENTER X3
```

The above statements will give us two bivariate equations, one for X_4 and X_1 and one for X_4 and X_2. The commands will also give us the reduced-form equation for X_4 as predicted by X_1 and X_2. Finally, they will give us the full equation for X_4 regressed on X_1, X_2, and X_3.

$$\hat{X}_4 = a_{41} + b_{41}X_1$$

$$\hat{X}_4 = a_{42} + b_{42}X_2$$

$$\hat{X}_4 = a_{4\cdot12} + b_{41\cdot2}X_1 + b_{42\cdot1}X_2$$

$$\hat{X}_4 = a_{4\cdot123} + b_{41\cdot23}X_1 + b_{42\cdot13}X_2 + b_{43\cdot12}X_3$$

The regression coefficients from these equations can be used to get the total effects, indirect effects, and spurious association in terms of structural coefficients, as follows:

$$b_{41\cdot2} = a + d\cdot c = T_{41}$$
$$b_{42\cdot1} = b + e\cdot c = T_{42}$$
$$b_{41\cdot2} - b_{41\cdot23} = (a + d\cdot c) - a = d\cdot c = I_{431}$$
$$b_{42\cdot1} - b_{42\cdot13} = (b + e\cdot c) - b = e\cdot c = I_{432}$$
$$b_{41} - b_{41\cdot2} = b_{41} - T_{41} = S_{41}$$
$$b_{42} - b_{42\cdot1} = b_{42} - T_{42} = S_{42}$$

In this case, we have to make one subtraction for each exogenous variable to get its indirect effect, instead of multiplying the structural coefficients along the compound paths. We do not have to add indirect effects and direct effects together to get the total effect; it can be read off the printout for the reduced-form equation. Although there is only a small savings in computations for this simple model, the savings can be considerable for more complex models. Most importantly, perhaps, familiarity with reduced-form models sharpens our understanding of causal modeling.

A Three-Equation Causal Model

It is possible to elaborate further the causal model of Figure 8.3 by adding a causal path between the two exogenous variables. Let us assume that theory justifies specifying a causal path from X_1 to X_2. This gives Figure 8.5.

Figure 8.5 represents a three-equation model with three endogenous variables and only one exogenous variable, other than the three error terms. The three equations for estimating the parameters of Figure 8.5 are

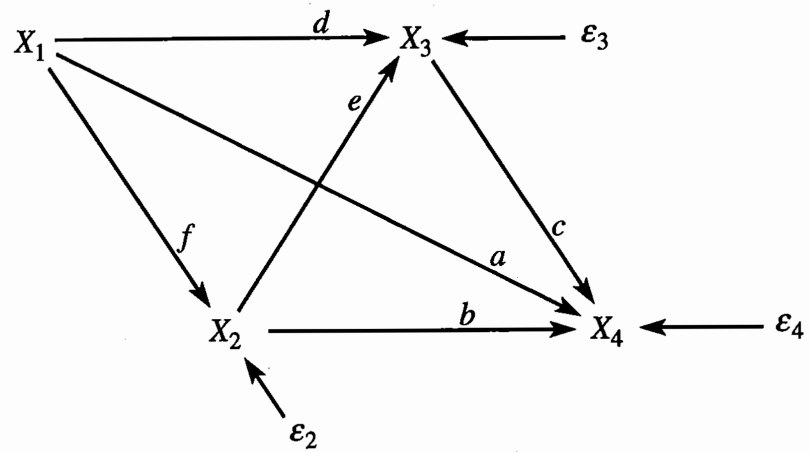
$$\hat{X}_4 = a_{4\cdot123} + b_{41\cdot23}X_1 + b_{42\cdot13}X_2 + b_{43\cdot12}X_3 \qquad \textbf{(8.7)}$$

$$\hat{X}_3 = a_{3\cdot12} + b_{31\cdot2}X_1 + b_{32\cdot1}X_2 \qquad \textbf{(8.8)}$$

$$\hat{X}_2 = a_{21} + b_{21}X_1 \qquad \textbf{(8.9)}$$

The parameters for X_4 and X_3 in Figure 8.5 are identical to those in Figure 8.3. Thus, Equations 8.7 and 8.8 for X_4 and X_3, respectively, are identical to Equations 8.3 and 8.4. The new parameter f and the new variable ε_2 in Figure 8.5 are estimated by using Equation 8.9.

FIGURE 8.5 Three-Equation Causal Model



DITS

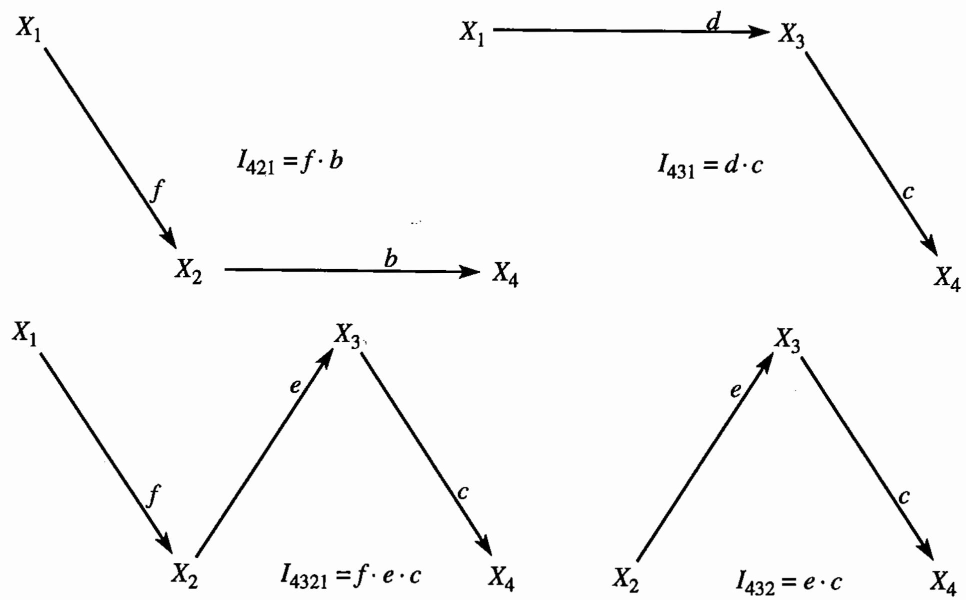
Although there is only one new structural coefficient in Figure 8.5 (i.e., f), the addition of X_2 as an endogenous variable that intervenes between X_1 and the final two endogenous variables (X_3 and X_4) creates an increase in the complexity of indirect effects. Figure 8.6 shows the four indirect effects that are now present in the model.

Indirect effects I_{431} and I_{432} are the same as in Figure 8.3. The new indirect effects are the ones that pass from X_1 through X_2, I_{421} and I_{4321}. The latter indirect effect involves a chain of three direct effects. The value of this longer indirect effect is determined in the same manner as for the shorter indirect effects.

I_{4321}
$\Delta X_1 = 1$
$\Delta X_2 = f \cdot \Delta X_1 = f \cdot 1 = f$
$\Delta X_3 = e \cdot \Delta X_2 = e \cdot f$
$\Delta X_4 = c \cdot \Delta X_3 = c \cdot (e \cdot f)$
$I_{4321} = f \cdot e \cdot c$

As before, indirect effects equal the product of all structural coefficients (or path coefficients in the case of a standardized model) along the compound path. When there is more than one indirect effect, the total effect of one variable on another will equal the direct effect plus the sum of indirect effects. Since Figure 8.5 does not add any new intervening variables between X_2 and the later endogenous variables (X_3 and X_4), the total effects of X_2 do not differ from those for Figure 8.3. The total effects of X_1, however, are now given by

FIGURE 8.6 Indirect Effects Present in Figure 8.5



$$T_{ij} = D_{ij} + \sum I_{i...j}$$

$$T_{21} = f$$

$$T_{31} = d + f \cdot e$$

$$T_{41} = a + d \cdot c + f \cdot b + f \cdot e \cdot c$$

The spurious association again equals the bivariate slope minus the total effect ($S_{ij} = b_{ij} - T_{ij}$). The spurious associations between X_2 and X_3, between X_2 and X_4, and between X_3 and X_4 are the same for Figure 8.5 as they were for Figure 8.3. The spurious associations of X_1 with the other variables, however, have changed because the total effects of X_1 have changed.

$$S_{21} = b_{21} - T_{21} = b_{21} - b_{21} = 0$$

$$S_{31} = b_{31} - T_{31} = 0$$

$$S_{41} = b_{41} - T_{41} = 0$$

Figure 8.5 indicates that the total effect of X_1 on X_2 is equal to its direct effect, which equals b_{21} according to Equation 8.9; therefore, all of the association between these variables is nonspurious. Furthermore, since no other variables in the model are correlated with X_1 or antecedent to X_1, all of the association between X_1 and X_3 and between X_1 and X_4 is due to the direct and indirect effects of X_1 on these variables. Therefore, there is no spurious association between X_1 and the other variables in Figure 8.5.

Reduced and Semi-Reduced Models

We can again simplify our model by eliminating all indirect effects from the diagram. The reduced form of Figure 8.5, containing only the paths from the exogenous variable to the endogenous variables, is shown in Figure 8.6. Each endogenous variable in Figure 8.7 is linked by a single causal path to the exogenous variable X_1. As in Figure 8.4, the coefficients represent the total effect of X_1 on each endogenous variable. These effects are estimated with the following bivariate regression equations:

$$\hat{X}_2 = a_{21} + b_{21}X_1 \quad \textbf{(8.10)}$$

$$\hat{X}_3 = a_{31} + b_{31}X_1 \quad \textbf{(8.11)}$$

$$\hat{X}_4 = a_{41} + b_{41}X_1 \quad \textbf{(8.12)}$$

Equation 8.10 is identical to the equation for X_2 in Figure 8.5. Thus, Equation 8.10 is not a reduced-form equation. Equations 8.11 and 8.12 differ from the equations for X_3 and X_4 in the full model because the intervening variables

FIGURE 8.7 Reduced Form of the Three-Equation Model

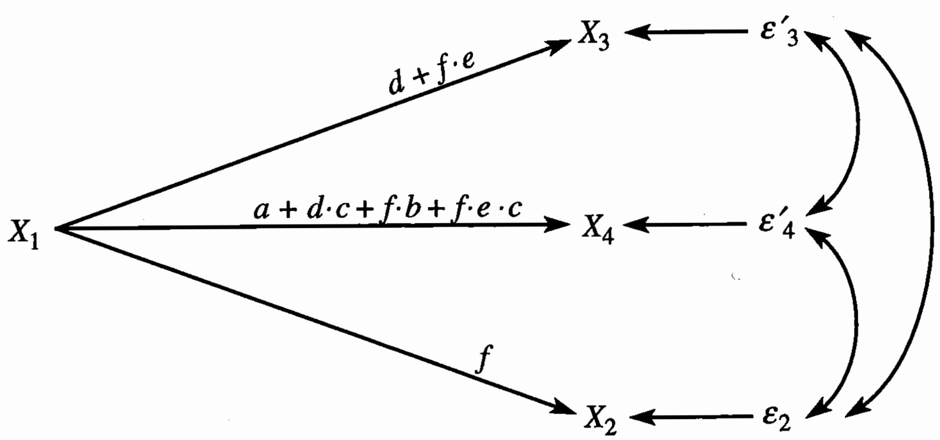
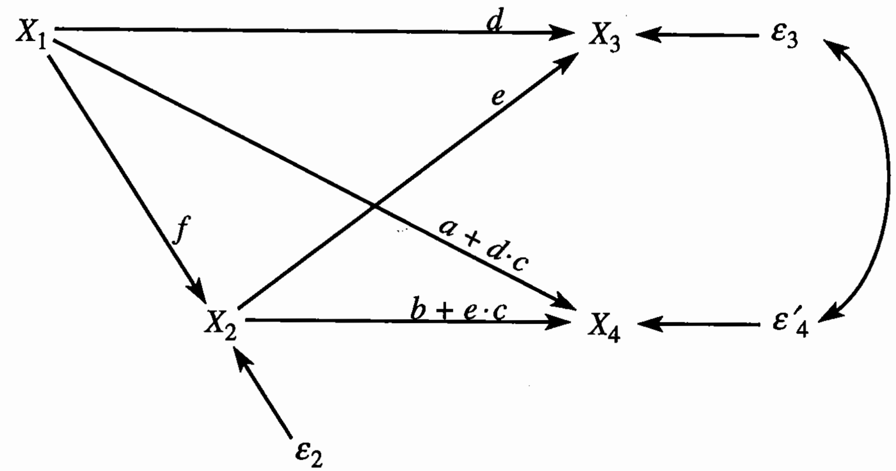


FIGURE 8.8 Semi-Reduced Form of the Three-Equation Model



have been removed from these equations. Thus, Equations 8.11 and 8.12 are reduced-form equations. The entire model is called a reduced model because some, but not all, of the equations are reduced. Since the total effects equal the bivariate slopes in this reduced model, we can again see that there is no spurious association between X_1 and the endogenous variables in this model.

The error terms for X_3 and X_4 in Figure 8.7 are different from those in Figure 8.5 because the variance in X_3 uniquely explained by X_2 and the variance in X_4 uniquely explained by X_2 and X_3 become part of the error terms in the reduced-form equations. Also, for reasons analogous to those given for reduced Figure 8.4, the error terms in Figure 8.7 are correlated.

It is also useful to consider the **semi-reduced** model shown in Figure 8.8. The semi-reduced model includes causal paths from the first endogenous variable, X_2, to the later endogenous variables, X_3 and X_4. As in the reduced model, however, the path from X_3 to X_4 is omitted. Thus, the semi-reduced model allows for the indirect effects that pass through the first endogenous variable, but it does not specify indirect effects through later endogenous variables. In semi-reduced models, each endogenous variable (except the first) is affected by all of the exogenous variables plus the first endogenous variable.

The equations for estimating the semi-reduced model are

$$\hat{X}_2 = a_{21} + b_{21}X_1 \quad \textbf{(8.13)}$$

$$\hat{X}_3 = a_{3\cdot12} + b_{31\cdot2}X_1 + b_{32\cdot1}X_2 \quad \textbf{(8.14)}$$

$$\hat{X}_4 = a_{4\cdot12} + b_{41\cdot2}X_1 + b_{42\cdot1}X_2 \quad \textbf{(8.15)}$$

Equations 8.13 and 8.14 are the same as the equations for X_3 and X_4 in the full model (Figure 8.5). These equations, therefore, are neither reduced nor semi-reduced. Only the equation for X_4 (Equation 8.15) is different in the semi-reduced model. Equation 8.15 does not contain X_3, the second endogenous variable, but it does contain X_2, the first endogenous variable. Equation 8.15 is thus a semi-reduced equation. The entire model is called a semi-reduced model, even though not all equations in the model are semi-reduced. Since X_3 is not contained in Equation 8.15, the slope coefficients in the semi-reduced equation equal the direct effects of each variable plus the indirect effects that pass through X_3.

Indirect Effects. By using the Alwin-Hauser method, we can use the differences between the slopes in the reduced, semi-reduced, and full models to compute various indirect effects.

$$\text{Reduced Model } b_{ij} - \text{Semi-Reduced Model } b_{ij} = \sum I_{i\ldots j} \text{ Through First Endogenous } X$$

$$b_{41} - b_{41\cdot2} = (a + d\cdot c + f\cdot b + f\cdot e\cdot c) - (a + d\cdot c)$$

$$= f\cdot b + f\cdot e\cdot c = I_{421} + I_{4321}$$

$$b_{31} - b_{31\cdot2} = (d + f\cdot e) - d = f\cdot e = I_{321}$$

$$\text{Semi-Reduced } b_{ij} - \text{Full } b_{ij} = \sum I_{i\ldots j} \text{ Through Second Endogenous } X$$

$$b_{41\cdot2} - b_{41\cdot23} = (a + d\cdot c) - a = d\cdot c = I_{431}$$

$$b_{42\cdot1} - b_{42\cdot13} = (b + e\cdot c) - b = e\cdot c = I_{432}$$

The sum of the indirect effects that pass through the first endogenous variable, in this case X_2, equals the difference between the slope in the reduced-form equation for a particular variable and the slope in the semi-reduced equation for that variable. In our model, X_1 is the only exogenous variable, and thus it is the only variable for which we can calculate the difference between its reduced and semi-reduced coefficients. With respect to the effect of X_1 on X_4, we see that the difference between its reduced and semi-reduced b's equals the sum of two indirect effects, I_{421} and I_{4321}. The first is a two-path chain and the second is a three-path chain; both, however, initially pass through X_2. The difference between the reduced and semi-reduced slopes does not allow us to differentiate between these two indirect effects (I_{421} and I_{4321}); it simply provides a summary of all indirect effects passing through X_2. We would have to multiply the coefficients along the paths to calculate each distinct indirect effect.

We can also use the reduced and semi-reduced models to calculate the indirect effect of X_1 on X_3 (I_{321}) as shown in the formulas above. In this case, there is only one indirect effect that passes through X_2. Finally, the difference between the b's in the semi-reduced and full equations for X_4 can be used to calculate the indirect effects of X_1 and X_2 on X_4 that pass through X_3 (I_{431} and I_{432}), as shown in the above formulas.

Different Models Produce Different DITS

In order to study some of the rules of causal analysis, we have examined the properties of three different types of causal models that can be specified for four variables: the single-equation model shown by Figure 8.1, the two-equation model shown by Figure 8.3, and the three-equation model shown by Figure 8.5. Table 8.1 provides a summary of the various total effects that were derived for each model.

Clearly, the model that we choose can potentially make a big difference in the magnitude of the total effects that we might find for each of the independent variables, with the exception of X_3. If we read across any row in the table except the last one, there is at least one difference between models with respect to the total effect of the independent variable in that row on the dependent variable in that row. We say that there is a potential difference between models, because it is certainly possible that one of the letters representing a difference in effects between models might turn out to be zero, or nearly zero, when we empirically estimate it.

If the model that we choose to estimate can make a big difference in the size of the effects that we estimate, how do we choose between models? It is not valid to try them all and choose the one whose results we like best. Nor is it possible to determine empirically which model fits the data best. The differences between the models is a matter of the number of causal equations that we can theoretically justify. With nonexperimental data there is no way to test whether the causal order that we might specify is valid. All that we can do is use the best theory that is available and the best information about the temporal sequence of the variables that is available to specify certain equations. Once we have specified a model, we can empirically estimate the size of each

TABLE 8.1 Summary of Total Effects

Dependent	*Independent*	*Figure 8.1*	*Figure 8.3*	*Figure 8.5*
X_2	X_1	0	0	f
X_3	X_1	0	d	$d + f\cdot e$
	X_2	0	e	e
X_4	X_1	a	$a + d\cdot c$	$a + d\cdot c + f\cdot b + f\cdot e\cdot c$
	X_2	b	$b + e\cdot c$	$b + e\cdot c$
	X_3	c	c	c

structural or path coefficient. These estimates are valuable for the information they provide about the relative sizes of the different direct and indirect effects specified by the model. But the validity of these estimates is dependent on the validity of the causal order that we have specified.

A Causal Analysis of SES and Self-Esteem

The principles of causal analysis that have been presented will be illustrated with data from the 1986 Akron Area Survey, a telephone survey of residents of Summit County, Ohio. This example uses the same sample of cases ($n = 513$) that were used in the anomia example in Chapter 4. Instead of anomia, however, this example uses *self-esteem* (Rosenberg 1965). The respondents' scores on a self-esteem index (ESTEEM) equal the sum of their coded responses to the four questions shown in Figure 8.9. Note that the scoring has been reversed on the last two questions because the wording of these two items expresses low self-esteem (Figure 8.9). The range of values on ESTEEM is 4 to 16. The other three variables (which were also used in the anomia example) are years of education (1–20), family income (*less than $5,000* = 1; *$5,000–9,999* = 2; *$10,000–14,999* = 3; *$15,000–19,999* = 4; *$20,000–24,999* = 5; *$25,000–34,999* = 6; *$35,000–49,999* = 7; *$50,000 or more* = 8), and a measure of the respondents' subjective

FIGURE 8.9 Self-Esteem and Subjective Income Questions

Self-Esteem

I feel that I am a person of worth, at least on an equal basis with others. (*strongly agree* = 4; *somewhat agree* = 3; *somewhat disagree* = 2; *strongly disagree* = 1)

I am able to do things as well as most other people. (*strongly agree* = 4; *somewhat agree* = 3; *somewhat disagree* = 2; *strongly disagree* = 1)

I wish I could have more respect for myself. (*strongly agree* = 1; *somewhat agree* = 2; *somewhat disagree* = 3; *strongly disagree* = 4)

I certainly feel useless at times. (*strongly agree* = 1; *somewhat agree* = 2; *somewhat disagree* = 3; *strongly disagree* = 4)

Subjective Income

How seriously do you feel a personal shortage of money these days—a great deal, quite a bit, some, or little or none? (*great deal* = 1; *quite a bit* = 2; *some* = 3; *little or none* = 4)

assessment of the level of their incomes (SHORTINC in Figure 8.9). A high score on SHORTINC indicates a positive assessment of income.

As we did in the previous sections, a two-equation model will first be specified and analyzed according to the principles that have been discussed. Then a three-equation model will be specified to illustrate the additional causal information that can be extracted from such a model. Although other social scientists might argue that causal orderings should be specified that are different from the ones we will analyze, including models with reciprocal causation (see the section on nonrecursive models), we will examine the causal information that can be extracted from the models under the assumption that they are validly specified.

A Two-Equation Model

The two-equation model is shown in Figure 8.10. The following SPSS commands can be used to compute all of the regression equations necessary for estimating the coefficients for the causal model and for computing the various DITS.

```
REGRESSION DESCRIPTIVES = DEFAULTS COV XPROD/
  VARIABLES = ESTEEM SHORTINC INCOME EDUC/
  DEP = SHORTINC/ ENTER INCOME/
  DEP = SHORTINC/ ENTER EDUC/ENTER INCOME/
  DEP = ESTEEM/ ENTER INCOME/
  DEP = ESTEEM/ ENTER SHORTINC/
  DEP = ESTEEM/ ENTER EDUC/ ENTER INCOME/ ENTER SHORTINC
```

The coefficients on the paths in Figure 8.10 come from Equation 5 for ESTEEM and from Equation 3 for SHORTINC in Table 8.2. These are structural

FIGURE 8.10 A Two-Equation Model for Self-Esteem

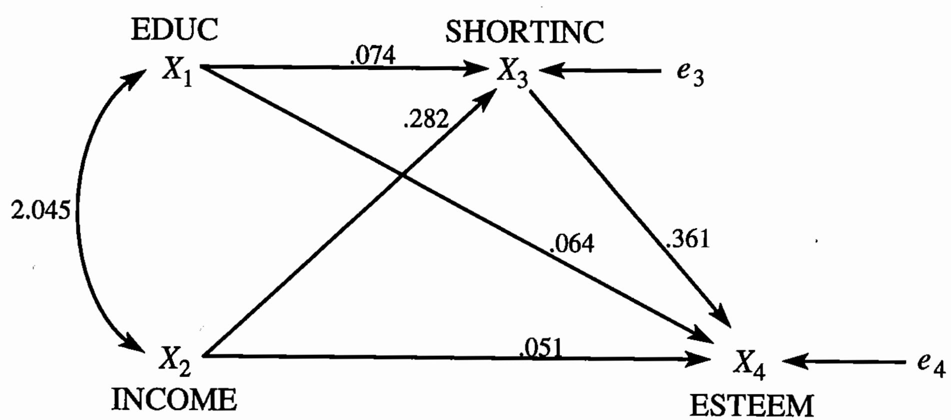


TABLE 8.2 Unstandardized Regression Equations for INCOME, SHORTINC, and ESTEEM (Standardized Coefficients)

	ESTEEM Equations				
	1	2	3	4	5
EDUC	.133*	—	—	.091*	.064
	(.165)			(.113)	(.080)
INCOME	—	.196*	—	.152*	.051
		(.184)		(.143)	(.048)
SHORTINC	—	—	.440*	—	.361*
			(.267)		(.219)
Constant	11.590	12.352	11.985	11.358	11.107
R^2	.027	.034	.071	.045	.081
$\hat{s}_\varepsilon$	2.169	2.162	2.119	2.151	2.113

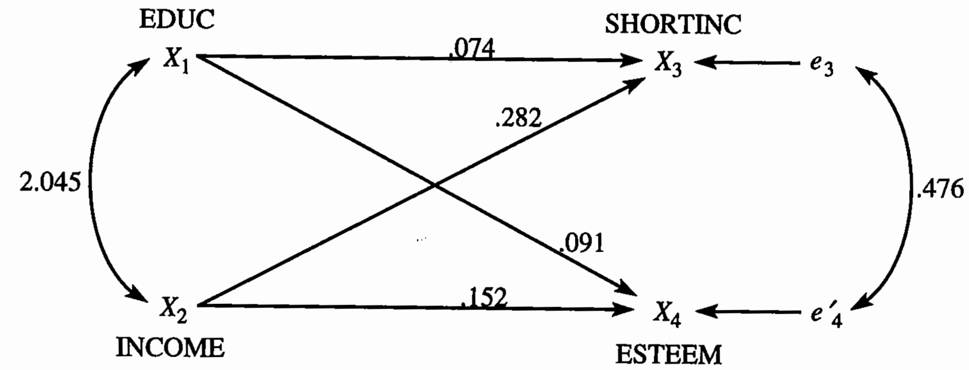
	SHORTINC Equations			*INCOME*
	1	2	3	1
EDUC	.151*	—	.074*	.274*
	(.309)		(.151)	(.364)
INCOME	—	.317*	.282*	—
		(.490)	(.435)	
Constant	1.124	1.501	.696	1.520
R^2	.096	.240	.260	.132
$\hat{s}_\varepsilon$	1.270	1.164	1.150	1.920

*$p \leq .05$

(unstandardized) coefficients. The diagram also shows a covariance of 2.045 between education and income (not given in Table 8.2). SHORTINC is the only variable having a significant ($p < .05$) effect on ESTEEM (Table 8.2). The positive coefficient indicates that the more a person feels he or she has enough income, holding constant education and actual income, the higher will be his or her self-esteem. It is interesting that subjective income is more important than objective income for self-esteem, whereas just the opposite was true for anomia (see Table 4.3). Both education and income, however, have significant positive effects on SHORTINC. We will use all coefficients in Figure 8.10, whether significant or not, for computing direct, indirect, total, and spurious effects (DITS).

Figure 8.11 shows the reduced form of the two-equation self-esteem model. This form shows only the effects of the exogenous variables (education and income) on self-esteem (there is no path from SHORTINC to ESTEEM). The structural coefficients for ESTEEM come from Equation 4 in Table 8.2. They represent the total effects of education and income on self-esteem, and both are signifi-

FIGURE 8.11 Reduced Form of the Two-Equation Model for Self-Esteem



cant. Figure 8.11 also shows a covariance of .476 between the error terms for SHORTINC and ESTEEM. This is the covariance between the regression residuals for SHORTINC and ESTEEM (not shown in Table 8.2).[2]

The direct, indirect, total, and spurious effects (DITS) and their computations are shown in Table 8.3. Two equivalent methods are shown: the *path-diagram* method multiplies the coefficients on compound paths in Figure 8.10 to get indirect effects and sums the direct and indirect effects to get the total effect; the Alwin-Hauser *hierarchical-equations* methods determines the indirect and total effects from the full equation coefficients and reduced-form equation coefficients in Table 8.2.

With respect to the effects of education on self-esteem, the indirect effect is smaller than the direct effect. However, the indirect effect is large enough so that when it is added to the direct effect, the total effect is significant (see reduced equation in Table 8.2). For income, however, its indirect effect on self-esteem is larger than its direct effect. In both these cases a failure to take into account indirect effects would lead the researcher to conclude that educational

2. The following SPSS commands can be used to compute this covariance:
```
REGRESSION DESCRIPTIVES/
  VARIABLES = ESTEEM SHORTINC INCOME EDUC/
  DEP = SHORTINC/ ENTER EDUC INCOME/
  SAVE = RESID (SHORTRES)
REGRESSION DESCRIPTIVES/
  VARIABLES = ESTEEM SHORTINC INCOME EDUC/
  DEP = ESTEEM/ ENTER EDUC INCOME/
  SAVE = RESID (ESTEMRES)
REGRESSION DESCRIPTIVES = DEFAULTS XPROD COV/
  VARIABLES = ESTEEM SHORTINC INCOME EDUC SHORTRES ESTEMRES/
  DEP = ESTEMRES/ ENTER SHORTRES
```

TABLE 8.3 Direct, Indirect, Total, and Spurious Effects (DITS) for the Two-Equation Self-Esteem Model

Dep. Var.	*Indep. Var.*	*DITS Formulas*	*Effects*
ESTEEM (X_4)	EDUC (X_1)	Path-Diagram Method	
		$D_{41} = b_{41.23} =$	.064
		$I_{431} = b_{31.2}b_{43.12} = (.074)(.361)$	.027
		$T_{41} = D_{41} + I_{431} =$	.091
		$S_{41} = b_{41} - T_{41} = .133 - .091 =$	.041
		Hierarchical-Equations Method	
		$D_{41} = b_{41.23} =$	.064
		$I_{431} = b_{41.2} - b_{41.23} = .091 - .064 =$	.027
		$T_{41} = b_{41.2} =$	.091
		$S_{41} = b_{41} - b_{41.2} = .133 - .091 =$	.041
	INCOME (X_2)	Path-Diagram Method	
		$D_{42} = b_{42.13} =$	.051
		$I_{432} = b_{32.1}\, b_{43.12} = (.282)(.361)$	.102
		$T_{42} = D_{42} + I_{432} =$	.153
		$S_{42} = b_{42} - T_{42} = .196 - .153 =$	.043
		Hierarchical-Equations Method	
		$D_{42} = b_{42.13} =$	.051
		$I_{432} = b_{42.1} - b_{42.13} = .152 - .051 =$	.101
		$T_{42} = b_{42.1} =$	.152
		$S_{42} = b_{42} - b_{42.1} = .196 - .152 =$	.044
	SHORTINC (X_3)	Path-Diagram Method	
		$D_{43} = b_{43.12} =$	.361
		$I_{4.3} =$ none	—
		$T_{43} = D_{43} + I_{4\ 3} =$	.361
		$S_{43} = b_{43} - T_{43} = .440 - .361 =$	.079
SHORTINC (X_3)	EDUC (X_1)	Path-Diagram Method	
		$D_{31} = b_{31.2} =$	.074
		$I_{3.1} =$ none	—
		$T_{31} = D_{31} + I_{321} =$	.074
		$S_{31} = b_{31} - T_{31} = .151 - .074 =$	.077
	INCOME (X_2)	Path-Diagram Method	
		$D_{32} = b_{32.1} =$	.282
		$I_{3.2} =$ none	—
		$T_{32} = D_{32} + I_{3..2} =$	.282
		$S_{32} = b_{32} - T_{32} = .317 - .282 =$	.035

attainment and objective income play no causal role in determining self-esteem. Nevertheless, the direct effect of subjective income is greater than the total effects of either education or income itself. For education, income, and subjective income, the spurious component of DITS is less than half as large as each variable's total effect. Also, the total effect of each variable is less than its bivariate slope and of the same sign, an indicator of redundancy among these variables.

A Three-Equation Model

The three-equation model is shown in Figure 8.12. The only difference between it and the two-equation model is that education is now specified as a cause of income. The structural coefficient for this new path is estimated by the regression equation in the last column of Table 8.2.

Figure 8.13 shows the reduced form of the model. It contains causal paths for only the single exogenous variable, education. The coefficients for each of these three paths, which represent the total effects of education, are estimated with the bivariate regression equations shown in Table 8.2. The covariances between the error terms for income, subjective income, and self-esteem are estimated by saving the residuals from the three bivariate regression equations and computing the covariances between these residuals (footnote 1 gives the SPSS commands for correlating the residuals of the two endogenous variables in the reduced form of the two-equation model). Figure 8.14 shows the semi-reduced model. In addition to causal paths from the exogenous variable education, the semi-reduced model shows the total effects of the first endogenous

FIGURE 8.12 A Three-Equation Model for Self-Esteem

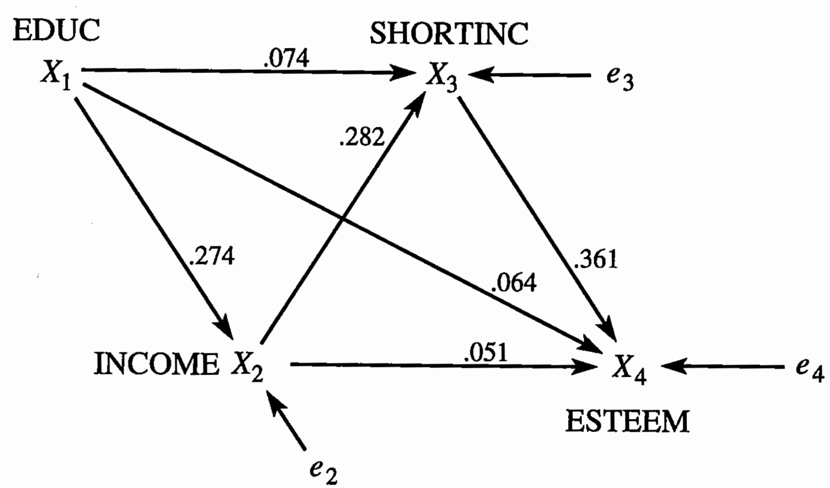


FIGURE 8.13 Reduced Form of the Three-Equation Model for Self-Esteem

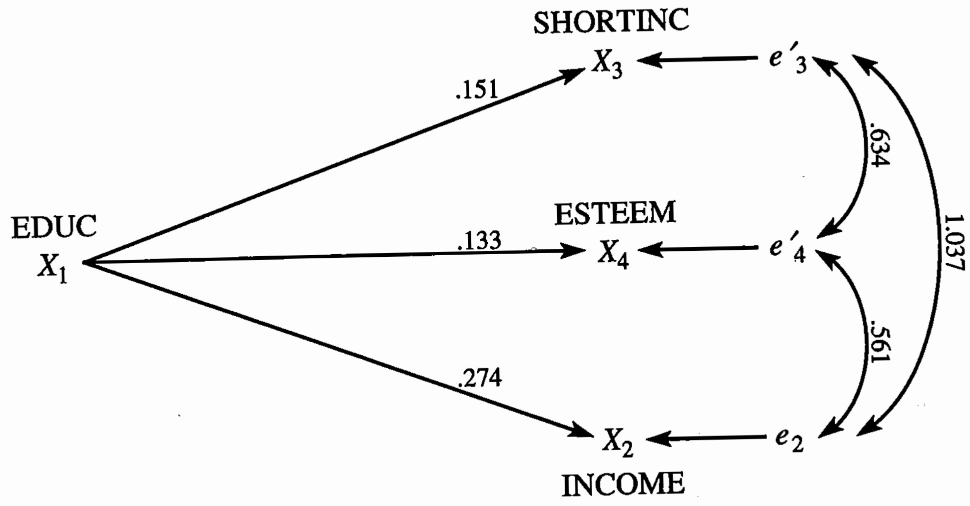


FIGURE 8.14 Semi-Reduced Form of the Three-Equation Model for Self-Esteem

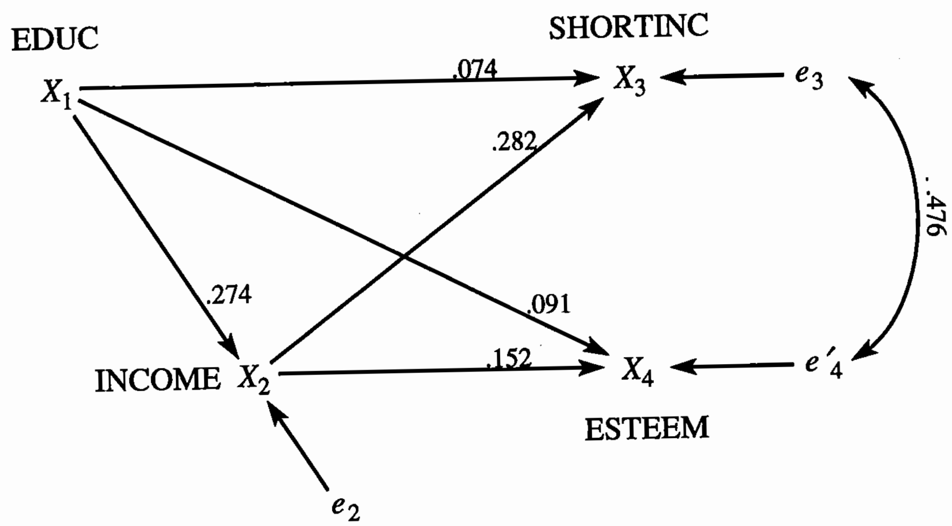


TABLE 8.4 Direct, Indirect, Total, and Spurious Effects (DITS) of Education in the Three-Equation Self-Esteem Model

Dep. Var.	*Indep. Var.*	*DITS Formulas*		*Effects*
ESTEEM (X_4)	EDUC (X_1)	Path-Diagram Method		
		$D_{41} = b_{41.23} =$		.064
		$I_{421} = b_{21}b_{42.13} = (.274)(.051) =$	.014	
		$I_{4321} = b_{21}b_{32.1}b_{43.12} = (.274)(.282)(.361) =$	.028	
		$I_{431} = b_{31.2}b_{43.12} = (.074)(.361) =$	.027	
		$\sum I_{4...1} =$		.069
		$T_{41} = D_{41} + \sum I_{4...1} =$		.133
		$S_{41} = b_{41} - T_{41} = .133 - .133 =$		.000
		Hierarchical-Equations Method		
		$D_{41} = b_{41.23} =$		.064
		$I_{421} + I_{4321} = b_{41} - b_{41.2} = .133 - .091 =$	.042	
		$I_{431} = b_{41.2} - b_{41.23} = .091 - .064 =$	.027	
		$\sum I_{4...1} = b_{41} - b_{41.23} = .133 - .064 =$		.069
		$T_{41} = b_{41} =$		.133
		$S_{41} = b_{41} - b_{41} =$		.000
SHORTINC (X_3)	EDUC (X_1)	Path-Diagram Method		
		$D_{31} = b_{31.2} =$		.074
		$I_{321} = b_{21}b_{32.1} = (.274)(.282) =$		.077
		$T_{31} = D_{31} + I_{321} =$		.151
		$S_{31} = b_{31} - T_{31} = .151 - .151 =$		.000
		Hierarchical-Equations Method		
		$D_{31} = b_{31.2} =$		.074
		$I_{321} = b_{31} - b_{31.2} = .151 - .074 =$		.077
		$T_{31} = b_{31} =$		.151
		$S_{31} = b_{31} - b_{31} = .151 - .151 =$		.000
INCOME (X_2)	EDUC (X_1)	Path-Diagram Method		
		$D_{21} = b_{21} =$		.274
		$I_{2..1} =$ none		—
		$T_{21} = D_{21} + I_{2..1} =$		.274
		$S_{21} = b_{21} - T_{21} = .274 - .274 =$		.000

variable, income, on the other two endogenous variables. Notice that except for the path from education to income, the semi-reduced form in Figure 8.14 is the same as the reduced form in Figure 8.11, including the covariance between the error terms.

The calculation of all new direct, indirect, total, and spurious effects for the three-equation model are shown in Table 8.4.

The three-equation model differs from the two-equation model only with respect to the total effects and indirect effects of education. By making income dependent upon education, several new indirect effects of education appeared. Table 8.4 gives all of the DITS effects for education. There are now two additional indirect effects of education on ESTEEM, both of which pass through INCOME; education increases income, which in turn has a slight effect on self-esteem (I_{421}), and education increases income, which increases subjective income, which increases self-esteem (I_{4321}). Notice that with the Alwin-Hauser method the difference between the reduced form and the semi-reduced form slope of self-esteem on education equals the sum of these two indirect effects. In many models, the hierarchical changes in slopes equal the sum of several indirect effects. Thus, although the Alwin-Hauser method is quick and accurate, if you want to know the values of each indirect effect, you may have to compute them by multiplying coefficients along the compound paths that define each indirect effect. But if you are mainly interested in total effects and total indirect effects, the Alwin-Hauser method is ideal.

The two new indirect effects of education on self-esteem increase the total indirect effect enough that it is now slightly greater than the direct effect of education. There is also an indirect effect of education on SHORTINC that is as large as its direct effect. Thus, not only does the three-equation model show the effect of education on income, it consequently opens up new indirect paths to subjective income and self-esteem. As a consequence, education assumes a more powerful explanatory role in the elaborated model. And since education is now the single exogenous variable in the model, none of its bivariate slope is spurious.

Nonrecursive Models

The models that we have examined are called **recursive** models. Recursive models do not have any causal loops; the causal flow is all in one direction. In a recursive model, the path effects leaving any particular variable will never return to that variable. As a consequence of the absence of any causal loops, *the parameters of recursive models can be estimated with ordinary least-squares regression.*

Nonrecursive models, however, are characterized by the presence of causal loops. Figure 8.15 provides an example. Figure 8.15 is the same as Figure 8.12 except that it includes a path from X_4 to X_3. This creates a causal loop between X_4 and X_3. A change in X_4 will cause a change in X_3, which will in turn feed back and cause a change in X_4. The return effect on X_4 would start another cycle around the loop, and so on. Although there are mathematical rules that can be used under certain circumstances to determine the total effect of the loop, they will not concern us here. The loop effect, however, is like an indirect effect, one that returns to cause a change in the original source of the effect. The loop effect could start with X_3 as well as with X_4.

FIGURE 8.15 A Nonrecursive Model

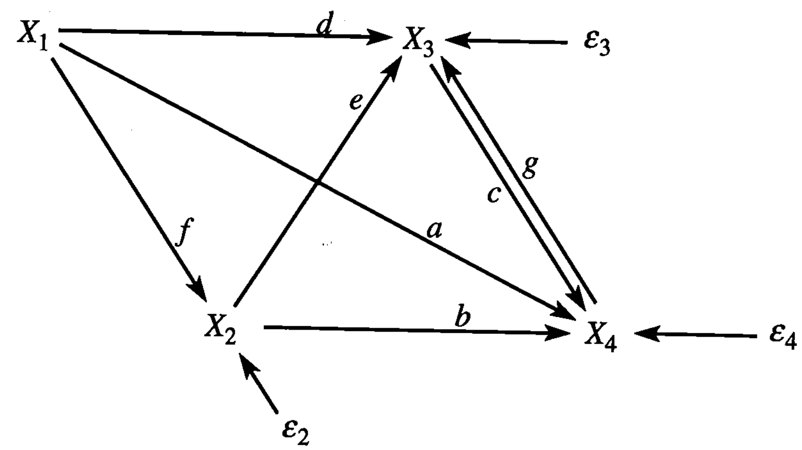


Figure 8.15 represents a three-equation model, one equation for each endogenous variable. Since the model is nonrecursive, not all of the parameters can be estimated with ordinary least-squares regression (for reasons to be demonstrated below). Therefore, we will not write these equations with regression notation, but instead, we will use the letters in the diagram as the coefficients for the variables in the equations. In order to avoid confusion between the effect a and the a that we have used as a constant in the regression equations, we will simply omit the intercept from the following equations (alternatively, we may assume that the X's are deviation scores, in which case the intercept will equal zero).

$$X_2 = fX_1 + \varepsilon_2 \quad \textbf{(8.16)}$$

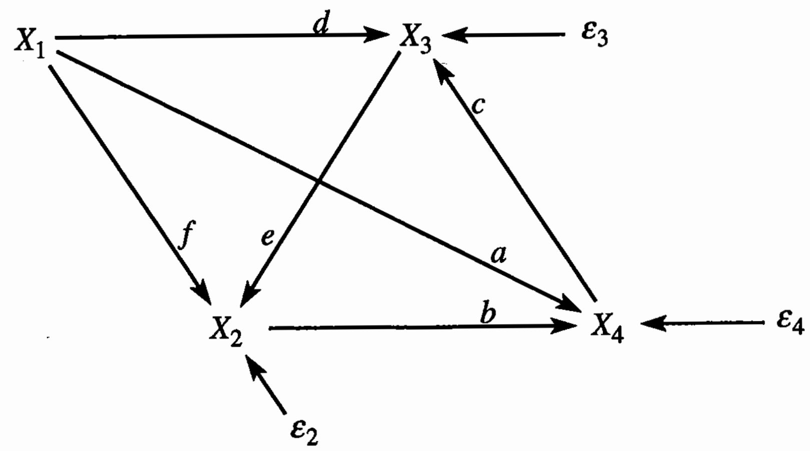
$$X_3 = dX_1 + eX_2 + gX_4 + \varepsilon_3 \quad \textbf{(8.17)}$$

$$X_4 = aX_1 + bX_2 + cX_3 + \varepsilon_4 \quad \textbf{(8.18)}$$

Equation 8.16 is actually a recursive equation because X_2 is not in a loop with any other variables. Thus, not all equations in a nonrecursive model are nonrecursive equations. Because Equation 8.16 is recursive, the parameter f can be estimated with an ordinary least-squares regression equation. Notice that the equation for X_3 (Equation 8.17) now contains X_4 because Figure 8.15 shows a path from X_4 to X_3. The equation for X_4 in turn contains X_3. Thus, each variable is included in the equation for the other.

Figure 8.15 shows an indirect path running from the error term for X_4 to X_3, that is, $\varepsilon_4 \rightarrow X_4 \rightarrow X_3$. This indirect path means that ε_4 and X_3 will covary or be correlated. Looking at Equation 8.18, this means that one of the independent

FIGURE 8.16 Another Nonrecursive Model



variables in the equation for X_4, namely X_3, is correlated with the error term for X_4. A basic assumption for using ordinary least-squares regression is that the error term must be uncorrelated with the independent variables. Since this assumption is clearly violated due to the loop in the model, least-squares regression should not be used to estimate the parameters in the equation for X_4. If we were to use the regression equation

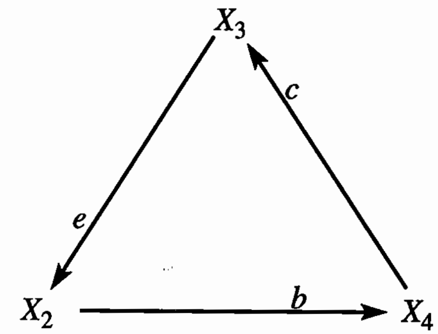
$$\hat{X}_4 = b_{41\cdot23}X_1 + b_{42\cdot13}X_2 + b_{43\cdot12}X_3$$

to estimate the effects on X_4, $b_{43\cdot12}$ would be a biased estimate of c in Equation 8.18 because X_3 is correlated with the error term for X_4.

Figure 8.15 also shows the indirect path $\varepsilon_3 \rightarrow X_3 \rightarrow X_4$. This path will cause X_4 to be correlated with ε_3. Therefore, X_4 will be correlated with the error term for X_3 in Equation 8.17. If least-squares regression were used to estimate the effect of X_4 on X_3, $b_{34\cdot12}$ would be a biased estimate of parameter g.

To summarize, Figure 8.15 shows that the error term for each of the variables in the loop will be correlated with the other variable in the loop. This is the characteristic of nonrecursive models or equations that causes ordinary least-squares regression to give biased estimates of the effects of the variables in the loop.

Figure 8.16 shows another nonrecursive model. At a glance, this diagram looks like Figure 8.12. However, the direction of the paths between X_3 and X_4 and between X_3 and X_2 have been reversed in Figure 8.16. This creates a loop between X_3, X_4, and X_2.



There are three equations in Figure 8.16, each containing two independent variables:

$$X_3 = dX_1 + cX_4 + \varepsilon_3$$

$$X_2 = fX_1 + eX_3 + \varepsilon_2$$

$$X_4 = aX_1 + bX_2 + \varepsilon_4$$

Because of the loop, each of the variables in the loop will be correlated with the error term in the equation in which they are one of the independent variables. This means that the least-squares regression equations will provide biased estimates of the effects of the variables in the loop, as shown:

$\varepsilon_4 \rightarrow X_4 \rightarrow X_3 \rightarrow X_2$, $\therefore r(\varepsilon_4 X_2) \neq 0$, $\therefore b_{42\cdot1}$ is a biased estimate of b

$\varepsilon_3 \rightarrow X_3 \rightarrow X_2 \rightarrow X_4$, $\therefore r(\varepsilon_3 X_4) \neq 0$, $\therefore b_{34\cdot1}$ is a biased estimate of c

$\varepsilon_2 \rightarrow X_2 \rightarrow X_4 \rightarrow X_3$, $\therefore r(\varepsilon_2 X_3) \neq 0$, $\therefore b_{23\cdot1}$ is a biased estimate of e

Bias is always a matter of degree. If one or more paths in a loop are weak compared to the others, the bias resulting from using least-squares regression may not be severe. Furthermore, there is undoubtedly always some bias in the regression estimates of recursive models because it is often the case that not all relevant independent variables are included in the regression equation, creating at least some covariance between the included variables and the error term.

Still, it is hard to justify using ordinary least-squares regression when we believe that causal loops are present in our models. There are other techniques that may be used to attempt to estimate the parameters of nonrecursive models, such as indirect least-squares and two-stage least-squares (see Duncan 1975; Heise 1975; Wonnacott and Wonnacott 1979). These techniques, however, do not provide easy solutions to the difficult problems of estimating nonrecursive models. In order to use these techniques, additional variables, called *instru-*

mental variables, must be found and added to the models. The instrumental variables must have strong statistical properties, and we must make strong theoretical assumptions about the absence of certain causal relationships between these variables and the variables in the loops. Further discussion, however, of two-stage least-squares and related techniques is beyond the scope of this book. For the present, we must be aware of the potential for causal loops in our models and recognize the inappropriateness of using ordinary least-squares regression when we believe that nonrecursiveness is present in our models.

A Model for Anomia and Self-Esteem

In this chapter we used a model for self-esteem that included education, income, and subjective income. In Chapter 4, anomia was used as a dependent variable in an equation that also used education, income, and subjective income as independent variables. Would it therefore be possible to add anomia as an endogenous variable to the three-equation model for self-esteem, thus converting it to a four-equation model? If so, would anomia be specified as a cause of self-esteem or vice versa? Some would feel it would be difficult to choose between these two alternatives and might want to specify a nonrecursive model in which anomia and self-esteem have reciprocal effects on each other. Such a model is shown in Figure 8.17.

As indicated in the previous discussion, estimating the reciprocal relationship between anomia and self-esteem would be difficult and beyond the scope of this book. However, there is a way out of this dilemma that allows us to salvage much of the model. We can choose not to attempt to estimate the reciprocal relationship but instead to estimate the remainder of the model with ordinary least-squares. If we eliminate the paths between anomia and self-esteem, we will be left with four equations, one each for INCOME, SHORTINC, ESTEEM, and ANOMIA, each of which can be estimated with ordinary least-squares. In fact, all of these regression equations have already been computed. The equation for ANOMIA was reported as Equation 1 in Table 4.3. The equations for INCOME, SHORTINC, and ESTEEM were given in Table 8.2. If we redraw Figure 8.17 and enter the coefficients from these tables, we get Figure 8.18.

Figure 8.18 is actually a semi-reduced form of Figure 8.17 produced by eliminating the causal paths between the last two endogenous variables, ESTEEM and ANOMIA. As such, the structural coefficients on the arrows leading to ESTEEM and ANOMIA are equal to a direct effect plus an indirect effect that passes through the arrows between ANOMIA and ESTEEM that are included in the full model (Figure 8.17). For example, −.213 on the path from SHORTINC to ANOMIA equals the direct effect of SHORTINC on ANOMIA (D_{53}) plus the

FIGURE 8.17 A Nonrecursive Model for Self-Esteem and Anomia

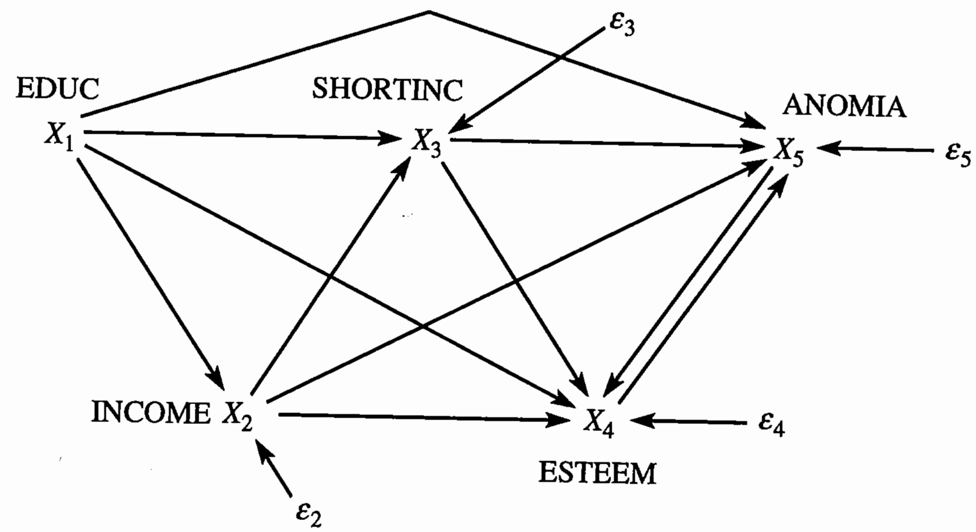
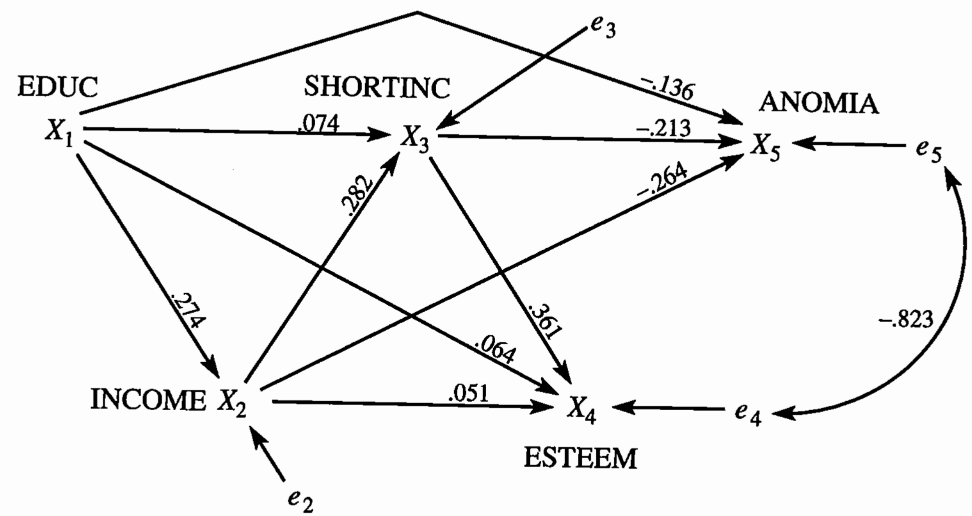


FIGURE 8.18 Semi-Reduced Model for Anomia and Self-Esteem



indirect effect that runs from SHORTINC to ESTEEM to ANOMIA (I_{543}).[3] Thus, from this semi-reduced model we can compute valid estimates of the total effects of education, income, and subjective income on anomia and self-esteem. Because we have chosen not to try to estimate the effects of anomia and self-esteem on one another, however, we cannot estimate the direct effects of EDUC, INCOME, and SHORTINC on ANOMIA and ESTEEM. Finally, the path between the error terms for ANOMIA and ESTEEM contains the covariance between these two errors. This covariance is produced by the causal effects of ANOMIA and ESTEEM on one another. The covariance may also be due to common causes outside the system that are independent of the education, income, and the subjective income variables.

Summary

In this chapter we have used multi-equation causal models to compute direct, indirect, total, and spurious (DITS) effects of one variable on another. The rules for constructing path diagrams to represent causal systems were described. If the model is recursive (i.e., no feedback loops), the structural coefficients or path coefficients for each path can be estimated by running an OLS regression equation for each endogenous variable. These coefficients represent the direct effects of one variable on another. Indirect effects between two variables (i.e., those that are mediated by at least one intervening variable) can then be calculated by multiplying the coefficients along the compound path that connects the two variables. In relatively complex models there may be several indirect paths between a pair of variables. The sum of the indirect effects for all of these paths gives the total indirect effect. When the total indirect effect is added to the direct effect, we get what is called the total effect.

A second method of calculating indirect and total effects (the Alwin-Hauser method) involves the use of reduced-form equations and semi-reduced equations. Reduced-form equations are created by omitting all of the endogenous variables that are included among the independent variables (i.e., a reduced-form equation contains only exogenous variables as independent variables). The slopes from a reduced-form equation represent estimates of the total effects of the exogenous variables. The difference between a variable's slope in the reduced-form equation and its slope in the full equation represents the total indirect effect of the variable. Adding the first intervening endogenous variable to the reduced-form equation produces a semi-reduced equation. The reduced-form slope minus the semi-reduced slope of an exogenous variable represents the sum of all indirect effects that pass through the first endogenous variable, and the slope for the first endogenous variable in the semi-reduced equation represents its total effect. The logic of this hierarchical approach can be used to calculate additional indirect and total effects by adding successive endogenous variables to the semi-reduced equation.

When feedback loops exist (nonrecursive models), ordinary least-squares cannot, in general, be used to estimate the structural or path coefficients, although some paths in a nonrecursive model may be recursive and are thus estimable with OLS regression. The reason that OLS regression gives biased estimates of nonrecursive paths is that if an independent variable is involved in a feedback loop that also involves the dependent variable, the independent variable will be correlated with the error term for the dependent variable, which is a violation of one of the regression assumptions.

This chapter considered the calculation of indirect effects for systems of equations involving only linear and additive effects. Methods for computing indirect effects in nonlinear and nonadditive models are described and illustrated in Stolzenberg (1979). Tests of statistical significance for indirect effects also were not covered here. Such tests are not readily available from regression programs. However, methods for estimating the standard errors of indirect effects are given by Sobel (1982).

3. Special algebraic rules that we have not covered are needed to compute indirect effects that pass through variables that are in a loop.

References

Alwin, Duane F., and Robert M. Hauser. 1975. "The Decomposition of Effects in Path Analysis." *American Sociological Review* 40:37–47.

Duncan, Otis Dudley. 1975. *Introduction to Structural Equation Models*. New York: Academic Press.

Heise, David. 1975. *Causal Analysis*. New York: Wiley.

Jöreskog, Karl G., and Dag Sörbom. 1989. *LISREL 7 User's Reference Guide*. Mooresville, Indiana: Scientific Software, Inc.

Rosenberg, Morris. 1965. *Society and the Adolescent Self-Image*. Princeton, New Jersey: Princeton University Press.

Sobel, Michael. 1982. "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models." Pp. 290–312 in *Sociological Methodology 1982*, edited by S. Leinhardt. San Francisco: Jossey-Bass.

Stolzenberg, Ross M. 1979. "The Measurement and Decomposition of Causal Effects in Nonlinear and Nonadditive Models." Pp. 459–488 in *Sociological Methodology 1980*, edited by Karl Schuessler. San Francisco: Jossey-Bass.

Wonnacott, Ronald J., and Thomas H. Wonnacott. 1979. *Econometrics*. New York: Wiley.

Wright, Sewell. 1921. "Correlation and Causation." *Journal of Agricultural Research* 20:557–585.