to the causal ordering of the variables, which was required to be specified at the outset. It is, therefore, an exceedingly misleading statistic. (One does not often see the mistake in quite this crude a form. But naïvely regressing causes on effects is far from being unknown in the literature.)

Another way to raise $R^2_{3(21)}$ would be to introduce another variable, say x_3', that is essentially an alternative measure of x_3, though giving slightly different results. The regression of x_3 on x_3', x_2, and x_1 is then guaranteed to yield a high value of R^2.

Indeed, the best-known examples of very high correlations are those selected to convey the notion of "spurious correlation," "nonsense correlation in time series," or other kinds of artifact. This shows us that high values of R^2, in themselves, are not sufficient to evaluate a model as successful.

Before worrying too much about his R^2, therefore, the investigator does well to reconsider the entire specification of the model. If that specification cannot be faulted on other grounds, the R^2 as such is not sufficient reason to call it into question.

Exercise. *To conclude, for the time being, your study of fully recursive models, review the material on estimation and testing in Chapter 3 and restate the essential points so that they apply to the recursive model as expressed without standardization of variables (page 52).*

FURTHER READING

An example of a recursive sociological model presented in terms of both standardized and nonstandardized coefficients appears in Duncan (1969). Note that the more interesting conclusions were developed on the basis of the latter. On the questionable value of commonly used measures of "relative importance" or "unique contribution" of the several variables in an equation, see Ward (1969), Cain and Watts (1970), and Duncan (1970).

5

A Just-Identified Nonrecursive Model

The model considered throughout this chapter is

$$x_3 = b_{31}x_1 + b_{34}x_4 + u$$

$$x_4 = b_{42}x_2 + b_{43}x_3 + v$$

For convenience, $E(x_j) = 0$, $j = 1, \ldots, 4$, and $E(u) = E(v) = 0$. However, we do *not* put the variables in standard form. Variables x_1 and x_2 are *exogenous*; their variances and their covariance are not explained within the model. Variables x_3 and x_4 are *jointly dependent* or *endogenous*; the purpose of the model is to explain the behavior of these variables. Variables u and v are, respectively, the *disturbances* in the x_3-equation and the x_4-equation. Their presence accounts for the fact that x_3 and x_4 are not fully explained by their explicit determining factors. The model will be operational only if we can assume that disturbances are uncorrelated with exogenous variables; hence the specification $E(x_1u) = E(x_1v) = E(x_2u) = E(x_2v) = 0$. *This is a serious assumption.* The research worker must carefully consider what circumstances would violate it and whether his theoretical understanding of the situation under study permits him to rule out such violations.

In contrast to the case of a fully recursive model, in the nonrecursive model the specification of zero covariances between disturbances and exogenous variables does not lead to either zero covariance between the two disturbances or zero covariance between the disturbance and each explanatory variable in an equation. To see this, let us multiply through each equation of the model by every variable in the model, and take expectations. To express the result of this operation in a convenient form, we will adopt the notation $E(x_j^2) = \sigma_{jj}$ where *sigma* with the repeated subscript refers to the (population) variance of x_j, and $E(x_h x_j) = \sigma_{hj}$ where (if $h \neq j$) *sigma* with two different subscripts refers to the (population) covariance of x_h and x_j. From the x_3-equation in the model we obtain, after multiplying through by x_1 and x_2:

$$E(x_1 x_3) = b_{31} E(x_1^2) + b_{34} E(x_1 x_4) + E(x_1 u)$$
$$E(x_2 x_3) = b_{31} E(x_1 x_2) + b_{34} E(x_2 x_4) + E(x_2 u)$$

or, since $E(x_1 u) = E(x_2 u) = 0$,

$$\sigma_{13} = b_{31}\sigma_{11} + b_{34}\sigma_{14}$$
$$\sigma_{23} = b_{31}\sigma_{12} + b_{34}\sigma_{24} \qquad \text{Set } (i)$$

However, in multiplying through the x_3-equation by endogenous variables and disturbances, not all covariances involving the disturbance drop out; we find:

$$\sigma_{33} = b_{31}\sigma_{13} + b_{34}\sigma_{34} + \sigma_{3u}$$
$$\sigma_{34} = b_{31}\sigma_{14} + b_{34}\sigma_{44} + \sigma_{4u}$$
$$\sigma_{3u} = b_{34}\sigma_{4u} + \sigma_{uu}$$
$$\sigma_{3v} = b_{34}\sigma_{4v} + \sigma_{uv} \qquad \text{Set } (ii)$$

Similarly, in multiplying through the x_4-equation by exogenous variables, we obtain:

$$\sigma_{14} = b_{42}\sigma_{12} + b_{43}\sigma_{13}$$
$$\sigma_{24} = b_{42}\sigma_{22} + b_{43}\sigma_{23} \qquad \text{Set } (iii)$$

But, in multiplying it through by endogenous variables and disturbances, we find:

$$\sigma_{34} = b_{42}\sigma_{23} + b_{43}\sigma_{33} + \sigma_{3v}$$
$$\sigma_{44} = b_{42}\sigma_{24} + b_{43}\sigma_{34} + \sigma_{4v}$$
$$\sigma_{4u} = b_{43}\sigma_{3u} + \sigma_{uv}$$
$$\sigma_{4v} = b_{43}\sigma_{3v} + \sigma_{vv} \qquad \text{Set } (iv)$$

Equations in Sets (*i*), (*ii*), (*iii*), and (*iv*) are population moment equations, analogous to but different in significant ways from those pertaining to recursive models. Note the lack of symmetry in the pattern of σ's on the right-hand side in Sets (*i*) and (*iii*), in contrast with the symmetric pattern of the normal equations on page 53. Note also—or carry out an *Exercise* to show this—that we cannot replace σ_{3u} by σ_{uu} or σ_{4v} by σ_{vv}, so that the simplification mentioned on page 53 for the recursive model is not available here. The population moment equations serve to express the relationships holding among the structural coefficients of the model (the b's) and the variances and covariances in the population under study. A number of important properties of the model are disclosed in studying these sets of equations.

Note, first, that if the b's in Set (*i*) are regarded as unknown and the σ's as known, it is possible to solve these two equations for the b's:

$$b_{31} = \frac{\sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23}}{\sigma_{11}\sigma_{24} - \sigma_{12}\sigma_{14}}$$
$$b_{34} = \frac{\sigma_{11}\sigma_{23} - \sigma_{12}\sigma_{13}}{\sigma_{11}\sigma_{24} - \sigma_{12}\sigma_{14}} \qquad \text{Set } (v)$$

In practice, of course, we would not know the population variances and covariances. However, if we replace the σ's by the corresponding sample moments, $m_{jj} = \sum (x_j^2)$ and $m_{hj} = \sum (x_h x_j)$, where the summation is over all sample observations on variables x_h and x_j, we obtain

$$\hat{b}_{31} = \frac{m_{13}m_{24} - m_{14}m_{23}}{m_{11}m_{24} - m_{12}m_{14}}$$
$$\hat{b}_{34} = \frac{m_{11}m_{23} - m_{12}m_{13}}{m_{11}m_{24} - m_{12}m_{14}} \qquad \text{Set } (vi)$$

This method of obtaining the estimates, $\hat{b}_{31}$ and $\hat{b}_{34}$, of the structural coefficients is termed instrumental variables. Here it is equivalent to the method of indirect least squares, for reasons that will become clear later. The method works here because the number of equations in Set (*i*), obtained by multiplying through the x_3-equation of the model by all exogenous variables, is just the same as the number of unknown structural coefficients. This fact is implied when we describe the model as "just identified" or "exactly identified" with respect to the x_3-equation. If Set (*i*) included more equations than structural coefficients, we would describe the x_3-equation as "overidentified"; if there were fewer moment equations in Set (*i*) than structural coefficients, the x_3-equation would be "underidentified" or "unidentified." In the case of overidentification, we replace instrumental variables (IV) or indirect least squares (ILS) by special methods of estimation. In the case of underidentification, estimation of structural coefficients is not possible.

Turning to the moment equations in Set (*iii*), we find that the x_4-equation of the model is likewise exactly identified. The solution for the *b*'s is

$$b_{42} = \frac{\sigma_{14}\sigma_{23} - \sigma_{13}\sigma_{24}}{\sigma_{12}\sigma_{23} - \sigma_{13}\sigma_{22}}$$

$$b_{43} = \frac{\sigma_{12}\sigma_{24} - \sigma_{14}\sigma_{22}}{\sigma_{12}\sigma_{23} - \sigma_{13}\sigma_{22}} \qquad \text{Set } (vii)$$

IV estimates, $\hat{b}_{42}$ and $\hat{b}_{43}$, may be obtained as before by replacing the σ's with sample moments, m_{jj} and m_{hj}:

$$\hat{b}_{42} = \frac{m_{14}m_{23} - m_{13}m_{24}}{m_{12}m_{23} - m_{13}m_{22}}$$

$$\hat{b}_{43} = \frac{m_{12}m_{24} - m_{14}m_{22}}{m_{12}m_{23} - m_{13}m_{22}} \qquad \text{Set } (viii)$$

Further study of this model is facilitated by solving for its *reduced form*, in which each endogenous variable is represented as a function of exogenous variables and disturbances only. Each equation is substituted into the other one. That is, for x_4 in the x_3-equation we substitute the right-hand side of the x_4-equation; and for x_3 in the x_4-equation we substitute the right-hand side of the x_3-equation. After collecting terms, this algebra yields the two reduced-form equations:

$$x_3 = \frac{1}{1 - b_{34}b_{43}}(b_{31}x_1 + b_{34}b_{42}x_2 + u + b_{34}v)$$

$$x_4 = \frac{1}{1 - b_{34}b_{43}}(b_{43}b_{31}x_1 + b_{42}x_2 + b_{43}u + v)$$

Let us adopt the notation,

$$a_{31} = \frac{b_{31}}{(1 - b_{34}b_{43})}$$

$$a_{32} = \frac{b_{34}b_{42}}{(1 - b_{34}b_{43})}$$

$$a_{41} = \frac{b_{43}b_{31}}{(1 - b_{34}b_{43})}$$

$$a_{42} = \frac{b_{42}}{(1 - b_{34}b_{43})}$$

If we now multiply through each reduced-form equation by the exogenous variables, we obtain

$$\sigma_{13} = a_{31}\sigma_{11} + a_{32}\sigma_{12}$$

$$\sigma_{23} = a_{31}\sigma_{12} + a_{32}\sigma_{22} \qquad \text{Set } (ix)$$

and

$$\sigma_{14} = a_{41}\sigma_{11} + a_{42}\sigma_{12}$$

$$\sigma_{24} = a_{41}\sigma_{12} + a_{42}\sigma_{22} \qquad \text{Set } (x)$$

since terms like $b_{43}E(x_1 u)$ and $E(x_2 v)$ drop out.

It appears that the reduced-form parameters (the *a*'s) are exact nonlinear functions of the structural coefficients (the *b*'s), and vice versa. Indeed, the four expressions defining the *a*'s could just as well be regarded as four equations in the unknown *b*'s, so that if the *a*'s were known we could solve for the b's by the following routine:

$$b_{34} = \frac{a_{32}}{a_{42}}$$

$$b_{43} = \frac{a_{41}}{a_{31}} \qquad \text{Set } (xi)$$

whence

$$1 - b_{34}b_{43} = 1 - \frac{a_{32}a_{41}}{a_{31}a_{42}}$$

and

$$b_{31} = a_{31}\left(1 - \frac{a_{32}a_{41}}{a_{31}a_{42}}\right)$$

$$b_{42} = a_{42}\left(1 - \frac{a_{32}a_{41}}{a_{31}a_{42}}\right) \qquad \text{Set } (xii)$$

Of course, the a's are not known, since they are functions of population variances and covariances, as shown by Sets (ix) and (x). But we could obtain estimates of the reduced-form parameters by replacing the σ's in Sets (ix) and (x) by corresponding sample moments; thus:

$$\hat{a}_{31} = \frac{m_{13}m_{22} - m_{12}m_{23}}{m_{11}m_{22} - m_{12}^2}$$

$$\hat{a}_{32} = \frac{m_{11}m_{23} - m_{12}m_{13}}{m_{11}m_{22} - m_{12}^2}$$

and

$$\hat{a}_{41} = \frac{m_{14}m_{22} - m_{12}m_{24}}{m_{11}m_{22} - m_{12}^2}$$

$$\hat{a}_{42} = \frac{m_{11}m_{24} - m_{12}m_{14}}{m_{11}m_{22} - m_{12}^2}$$

We see that these are precisely the same as the estimates we would obtain for the coefficients of the ordinary least squares (OLS) regressions of (respectively) x_3 on x_2 and x_1, and x_4 on x_2 and x_1, that is, of each endogenous variable on all exogenous variables.

Suppose we now replace the a's in Sets (xi) and (xii) by the corresponding OLS estimates, the $\hat{a}$'s. We will obtain estimates of the b's that are the very IV estimates presented earlier as Sets (vi) and $(viii)$; that is,

$$\hat{b}_{34} = \frac{\hat{a}_{32}}{\hat{a}_{42}}$$

$$\hat{b}_{43} = \frac{\hat{a}_{41}}{\hat{a}_{31}}$$

$$\hat{b}_{31} = \hat{a}_{31}\left(1 - \frac{\hat{a}_{32}\hat{a}_{41}}{\hat{a}_{31}\hat{a}_{42}}\right)$$

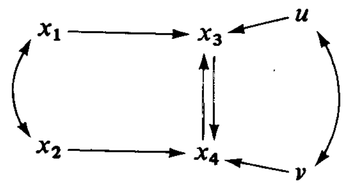
$$\hat{b}_{42} = \hat{a}_{42}\left(1 - \frac{\hat{a}_{32}\hat{a}_{41}}{\hat{a}_{31}\hat{a}_{42}}\right)$$

are the IV estimates of the b's. This fact may not be immediately obvious, but it is easily proved by algebraic substitutions.

Exercise. *Carry out this algebra.*

The fact that the $\hat{b}$'s are obtained from the $\hat{a}$'s, which in turn are OLS estimates of reduced-form regression coefficients, justifies the name indirect least squares.

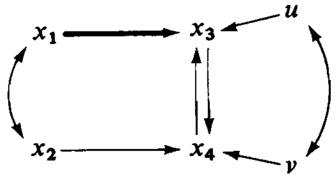
In addition to its possible use for purposes of estimation, the reduced form of the model is instructive in the way it displays the mechanisms through which the exogenous variables influence the endogenous variables. In this connection, study of the path diagram of the model is also instructive.



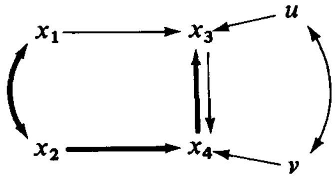
The covariance of x_1 and x_3 is obtained by multiplying through the reduced-form x_3-equation by x_1:

$$\sigma_{13} = \frac{b_{31}}{1 - b_{34}b_{43}}\sigma_{11} + \frac{b_{34}b_{42}}{1 - b_{34}b_{43}}\sigma_{12}$$

Note that there is a direct effect of x_1 on x_3:



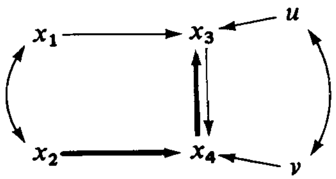
But another part of σ_{13} arises from the correlation (covariance) of x_1 with another cause (namely, x_2) of x_3, even though the latter works only indirectly:



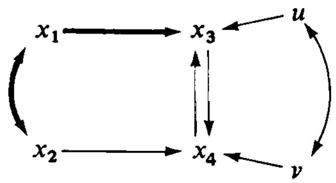
Note that we are reading the path diagram according to Sewall Wright's general principle, but are not using it as an algorithm for computing covariances. In particular, when actually calculating σ_{13} we must inflate both the direct path and the component due to a correlated cause by the factor $1/(1 - b_{34}b_{43})$. This is the "multiplier effect" in the model due to the "simultaneity" or "reciprocal causation" of the two endogenous variables. The covariance of x_2 and x_3 is

$$\sigma_{23} = \frac{b_{31}}{1 - b_{34}b_{43}}\sigma_{12} + \frac{b_{34}b_{42}}{1 - b_{34}b_{43}}\sigma_{22}$$

We see that it is produced by an indirect effect of x_2 on x_3:



and by a component due to the correlation of x_2 with x_1:



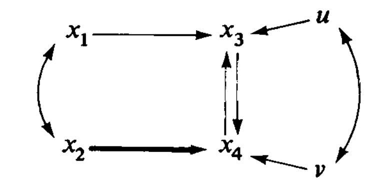
Again the multiplier effect comes into play for both components.

The same kinds of components can be found for the covariances of the two exogenous variables with x_4, shown below as they are obtained from the reduced-form x_4-equation:

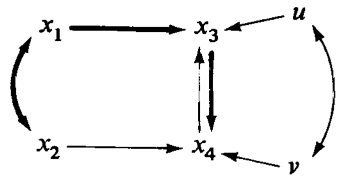
$$\sigma_{24} = \frac{b_{42}}{1 - b_{34}b_{43}}\sigma_{22} + \frac{b_{43}b_{31}}{1 - b_{34}b_{43}}\sigma_{12}$$

$$\sigma_{14} = \frac{b_{42}}{1 - b_{34}b_{43}}\sigma_{12} + \frac{b_{43}b_{31}}{1 - b_{34}b_{43}}\sigma_{11}$$

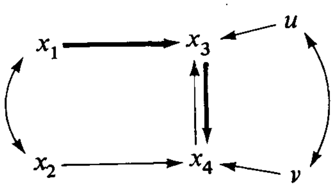
Thus, σ_{24} arises from the direct effect:



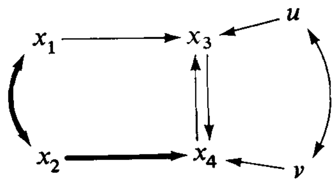
and the contribution due to a correlated cause, operating indirectly:



while σ_{14} arises from an indirect effect



and the contribution of a correlated cause:



Again, all these components are seen to include the multiplier effect of the factor $1/(1 - b_{34}b_{43})$.

It is important to understand precisely why special methods, instead of the conventional statistical procedure of OLS, are required for estimating structural coefficients in nonrecursive models. In the x_3-equation (to take it as a typical example), the two explanatory variables on the right-hand side are x_1 and x_4. If one were to estimate the coefficients of the x_3-equation by OLS, therefore, he would obtain as the estimate of b_{31}

$$\frac{m_{13}m_{44} - m_{14}m_{34}}{m_{11}m_{44} - m_{14}^2}$$

and as the estimate of b_{34}

$$\frac{m_{11}m_{34} - m_{13}m_{14}}{m_{11}m_{44} - m_{14}^2}$$

As we have seen, in Sets (*i*) and (*ii*), the covariances of x_3 with these explanatory variables are

$$\sigma_{13} = b_{31}\sigma_{11} + b_{34}\sigma_{14}$$

$$\sigma_{34} = b_{31}\sigma_{14} + b_{34}\sigma_{44} + \sigma_{4u}$$

We may solve for the b's in terms of the σ's:

$$b_{31} = \frac{\sigma_{13}\sigma_{44} - \sigma_{14}\sigma_{34} + \sigma_{14}\sigma_{4u}}{\sigma_{11}\sigma_{44} - \sigma_{14}^2}$$

$$b_{34} = \frac{\sigma_{11}\sigma_{34} - \sigma_{13}\sigma_{14} - \sigma_{11}\sigma_{4u}}{\sigma_{11}\sigma_{44} - \sigma_{14}^2}$$

The implication is clear if we note that, apart from $\sigma_{14}\sigma_{4u}$ in the numerator, the expression for b_{31} is the population counterpart of the OLS estimator of b_{31}; and apart from $-\sigma_{11}\sigma_{4u}$ in the numerator, the expression for b_{34} is the population counterpart of the OLS estimator of b_{34}. Thus, even if our sample were infinitely large, so that we could form OLS estimators from population variances and covariances (instead of sample moments), the OLS estimates would be biased. Indeed, the OLS procedure would not estimate b_{31} but rather

$$b_{31} - \frac{\sigma_{14}\sigma_{4u}}{\sigma_{11}\sigma_{44} - \sigma_{14}^2}$$

and it would not estimate b_{34} but rather

$$b_{34} + \frac{\sigma_{11}\sigma_{4u}}{\sigma_{11}\sigma_{44} - \sigma_{14}^2}$$

Similarly we can show that OLS applied to the x_4-equation would estimate not b_{42} but rather

$$b_{42} - \frac{\sigma_{23}\sigma_{3v}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2}$$

and, instead of b_{43}, it would estimate

$$b_{43} + \frac{\sigma_{22}\sigma_{3v}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2}$$

The basic reason for the failure of OLS, then, is that not all the explanatory variables in the equation are uncorrelated with the disturbance. And this is inescapably so given the jointly dependent (simultaneous, reciprocally influencing) relationships of the endogenous variables of a nonrecursive model. For if $x_3 \to x_4$, then $x_u \to x_3$ implies that $x_u \to x_3 \to x_4$ will contribute a nonzero component to σ_{4u}. Similarly,

$x_v \to x_4 \to x_3$ will contribute a nonzero component to σ_{3v}. It could happen that one or the other of these is cancelled out by σ_{uv}, for in Sets (*ii*) and (*iv*) we find

$$\sigma_{3v} = b_{34}\sigma_{4v} + \sigma_{uv}$$

and

$$\sigma_{4u} = b_{43}\sigma_{3u} + \sigma_{uv}$$

Hence, if $\sigma_{uv} = -b_{34}\sigma_{4v}$, then $\sigma_{3v} = 0$; or if $\sigma_{uv} = -b_{43}\sigma_{3u}$, then $\sigma_{4u} = 0$. But for either of these to hold would be merely a coincidence, and for both to hold would be a rare coincidence indeed.

Exercise. *Working from Sets (i), (ii), (iii), and (iv) in the spirit of Chapter 4, pages 53–55, show that the variances and covariances of the observable variables in the model studied in this chapter may be expressed as functions of (1) the variances and covariances of the exogenous variables, (2) a (nonlinear) combination of structural coefficients, and (3) variances and covariances of the disturbances. Verify Table 5.1.*

Table 5.1 Sources of Observable Variances and Covariances

Variance or covariance	*Is a function of*									
	σ_{11}	σ_{12}	σ_{22}	b_{31}	b_{42}	b_{34}	b_{43}	σ_{uu}	σ_{uv}	σ_{vv}
* σ_{11}	×	...	...	...	...	...	...	...	...	...
* σ_{12}	...	×	...	...	...	...	...	...	...	...
* σ_{22}	...	...	×	...	...	...	...	...	...	...
σ_{13}	×	×	...	×	×	×	×	...	...	...
σ_{14}	×	×	...	×	×	×	×	...	...	...
σ_{23}	...	×	×	×	×	×	×	...	...	...
σ_{24}	...	×	×	×	×	×	×	...	...	...
σ_{33}	×	×	×	×	×	×	×	×	×	×
σ_{34}	×	×	×	×	×	×	×	×	×	×
σ_{44}	×	×	×	×	×	×	×	×	×	×

* Exogenous

Discuss implications of the possibility that some, if not all, of the parameters across the top of the table are invariant across populations.

FURTHER READING

Chapters 5, 6, and 7 are essentially the elementary portions of the standard econometric presentation of simultaneous-equation models. Several advanced econometrics texts are listed among the references at the end of this book. More accessible presentations are available in Wonnacott and Wonnacott (1970, Part I), and Wallis (1973). The latter does, however, make use of matrices.