Two cautions are urged. First, recall our discussion of partial correlation. If X were maternal encouragement and Y were SES, we would conclude that maternal encouragement has no direct effect on child achievement (W), only an indirect one through socioeconomic class. Conceptually, the opposite seems more reasonable. The second caution is that just because the observed correlations can be reproduced by the model at the right of Part *i* does not mean that this particular model is the only one that can reproduce the pattern of correlations. Indeed, one would have to test all 12 possible models before one could argue that one model is better than the others at reproducing the observed correlations. And if more than one model can reproduce the observed correlations, what criteria shall we use to compare models?

For example, consider now Part *ii* of Table 12. Here is an alternative model for this case. It suggests that Y produces both X (contemporaneously) and W (predictively). For example, SES might be related to maternal encouragement contemporaneously and also predict later child achievement, with earlier maternal encouragement not influencing later achievement at all. In fact, this model is equal to the first in reproducing the observed pattern of correlations despite the fact that the direction of causality was reversed in one of the two causal specifications. Therefore, it is sometimes possible to have directly opposite causal hypotheses in two models, both of which fit the data equally well.

The testing of hypotheses in path analysis is most directly performed with statistical technology developed for the analysis of covariance structures (see below). These techniques include methods for testing the adequacy of a simple model, the comparison of competing models, and the assessment of incremental fit. A relatively nontechnical discussion of these models and their applications has been presented by Bentler and Bonett (1980).

The advantage of path analysis is to eliminate possible models. But the number of possible models meriting assessment can be enormous, and selection of one model may not always be possible, as we have seen. Moreover, one must keep a careful eye on the assumptions and the cautions voiced above with respect to interpreting partial correlations. Therefore, while path analysis may be a useful technique, especially for achieving parsimonious models with high predictive efficiency, its ability to determine causality from observational data is probably more limited than many suppose.

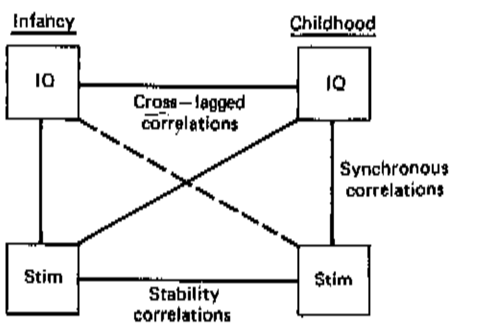
Cross-lagged Panel Analysis. What could be better for developmentalists concerned with individual differences than a procedure that could determine developmental causality from strictly observational longitudinal data and that requires nothing more complicated than the computation of a few simple correlation coefficients? No wonder cross-lagged panel analysis has been greeted so enthusiastically by some developmentalists. But the promise and simplicity of cross-lagged analysis must be purchased at the dear cost of a highly restrictive set of assumptions that are not often met by developmental data. And even when they are met, some specialists feel that conclusions derived from cross-lagged analysis may be inaccurate or misleading.

The logic and procedure of cross-lagged panel analysis is illustrated in Table 13. Suppose a mental test was administered during infancy and again during childhood to a sample of children, and their mothers were assessed for some aspect of intellectual stimulation or encouragement at these same points in time.

There are three types of correlations in such a "panel." First, child IQ can be correlated with mother's stimulation during infancy and again during childhood. Because these correlations are calculated on variables assessed at the same time, they are called *synchronous correlations*. Second, IQ can be correlated with itself across age and maternal stimulation can be correlated with itself across age. These longitudinal relationships between the same variables are called *stability correlations*. Third, the diagonals of the panel represent correlations across variables and across developmental time. They are the *cross-lagged correlations*.

Presumably, if the cross-lagged correlations are significantly different from one another, one has evidence suggesting causality. In the example, we would look for the correlation between maternal in-

Table 13. **General Scheme for Cross-Lagged Panel Analysis**



fant stimulation and childhood IQ to be statistically higher than the correlation between infant IQ and maternal childhood stimulation. Presumably, such a result would suggest that early stimulation by the mother produces a brighter child.

We offer several cautions. First, the statistical test comparing the two cross-lagged correlations must take into account the fact that the two *r*s are not independent. Researchers rarely consider this point. We have presented an appropriate test in Table 6 for the case of comparing two correlations involving four variables assessed on a single sample.

The next problem with cross-lagged panel analysis is that the desired interpretation depends on several assumptions that are unlikely to be met by the data and even more frequently disregarded completely by researchers. These assumptions include the following:

1. *The causal relationships for* X *and* Y *do not change over time*. A necessary, but not sufficient, sign that this assumption is met is that the synchronous correlations are equal. In our example, the correlation between maternal stimulation and child IQ must be the same during infancy as during childhood. But even if this assumption is supported by the data, Rogosa (1980) has shown that unequal cross-lagged correlations may nevertheless reflect a noncausal association.

2. *The stability correlations must be equal*. This implies two requirements. First, the two measures associated with Time 1 must be assessed at the same time, and the measures associated with Time 2 must also be measured at the same time. Second, the stability correlations must be equal. In our example, the stability for child IQ must be identical to the stability for maternal stimulation. In parent/child work, it might be expected that the stability of the parent variable will be greater than the stability of the child variable, thus violating this assumption.

3. We believe, in addition to these assumptions, interpretation is very difficult unless *all important variables have been measured and are included in the panel*. In the example in Table 13, many results consistent with these assumptions and a causal interpretation could be explained noncausally if IQ were genetic or if maternal stimulation was a correlate of maternal intelligence (and therefore of child intelligence).

Not only are these assumptions restrictive, but as Rogosa (1980) has pointed out, the interpretation of a cross-lagged panel analysis is not straightforward, *even when the assumptions are met*. The reason, in part, is that the relative sizes of the correlations just their significance or nonsignificance) are im tant. For example, moderate correlations betv maternal stimulation and later child IQ on the hand and maternal infant stimulation and mate childhood stimulation on the other might mean the mothers who stimulate early are not necess the same mothers who stimulate later. In additic the size of the correlations, it seems necessar have all of the relevant variables involved in system, and multiple variates and several ass ment occasions might be more informative.

Several statisticians and methodologists have cently suggested that cross-lagged analysis sir does not fulfill its promise as a method of disco ing causality (Rogosa, 1980; Wohlwill, 19 Rogosa (1980) is particularly blunt: "No justif tion was found for the use of CLC" (cross-lag correlations) and "CLC should be set aside as a d end" (p. 257). Other statisticians would sug using general simultaneous equation models and analysis of covariance structures instead of cr lagged procedures.

A more moderate view suggests that these m ods may help in eliminating some alternatives, as in the case of path analysis, especially if part partial correlations are calculated within the pa (i.e., the panel is treated somewhat like a path an sis). Even then, considerable interpretive cautio required, and one must always be on the lookout variables that are not assessed but that may unde the observed relationships.

Multiple Regression

Although multiple regression is commo known to developmental researchers, it is but technique in an extensive collection of regress methods that are less familiar, some of which h been discussed above. Indeed, some statistici would argue that regression techniques underlie most all statistical methods and that applied statist will someday be taught from this perspective.

We present here only a general introduction multiple regression. Readers are referred to seve comprehensive textbooks and articles on the subje Draper and Smith (1981) is a more-or-less classi text stressing the fundamentals of regression ana sis; Green (1978) and Kleinbaum and Kupper (197 are somewhat more advanced yet applied books t consider regression as part of general multivari methods; and Mosteller and Tukey (1977) prese regression analysis in the context of data analy and robust methods. Darlington (1968) offers a br and readable overview of multiple regression.