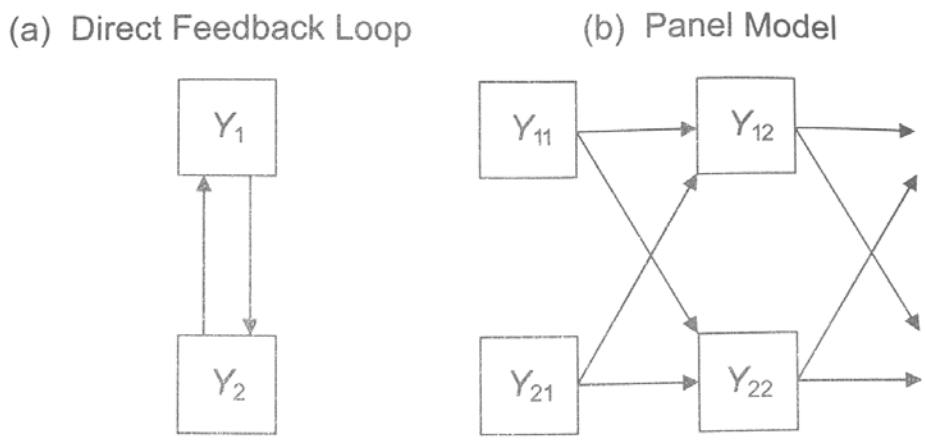
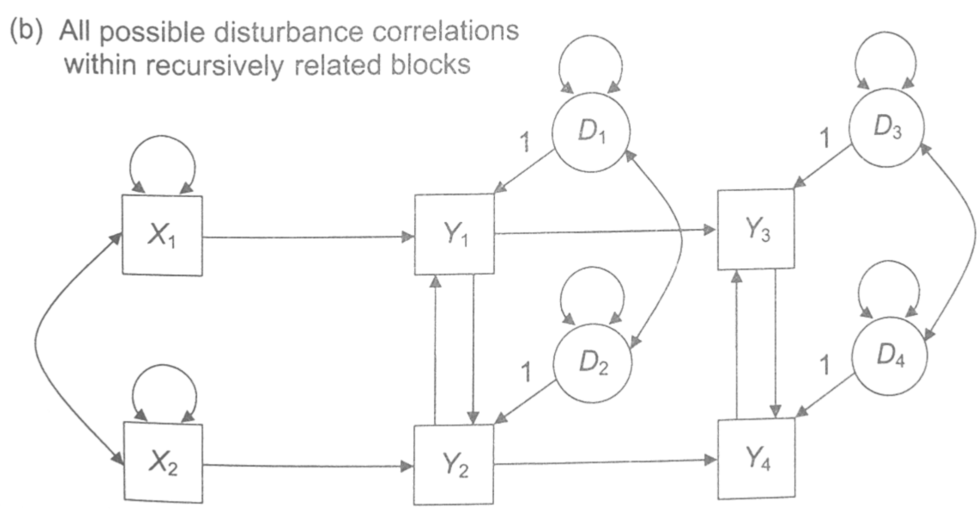
SC705: Advanced Statistics
Instructor: Natasha Sarkisian
Class notes: Nonrecursive models

These models may include feedback loops (A→ B → C → A) or reciprocal causal relationships, especially simultaneous (A→B and B→A). Note, that we can contrast simultaneous reciprocity with the lagged reciprocity that is used in panel models: A1→B2 and B1→A2. The latter is more realistic but only possible to model if we have longitudinal data. The former is appropriate if we assume that there is continuous bidirectional causation, and the causal effects are stable over time.



Finally, the term "block recursive" is typically applied when pairs of variables are linked by reciprocal causal relationships, but the links between the pairs are recursive.



Unlike recursive models, nonrecursive ones are not easy to estimate using regular regression techniques – note for instance that for a reciprocal causal model (A→B and B→A), the residuals of the models for A and B will not be independent. Special techniques involving multi-stage regression approaches have been developed to deal with such models, but SEM makes it easier to handle such models.

In regression approaches, the reciprocally related variables (A and B) are first each separately regressed on the full set of predictor variables that are hypothesized to have an effect on them. Second, predicted scores from such models are used in the final models alongside the rest of the predictors instead of the original variables to predict the other variable. Further, such models also have to allow the residuals of these final models to be correlated. In SEM models, that's not necessary – but we do allow for the residuals to be correlated if we think that there is a relationship between the unexplained portions of these variables.

Importantly, the set of predictors used to model A in the first step and the one included in the second step to model B should differ by at least one variable -- otherwise, it will be impossible to find a solution because the predicted value of A will be a perfect linear combination of other predictors. The variables excluded from the equation are called instrumental variables or instruments. Even though the estimation process in SEM is different and doesn't proceed in stages, the idea of an instrumental variable is still important, and we still need to locate such variables for each of the variables involved in a reciprocal relationship.

The most important difference in terms of working with nonrecursive models (as compared to recursive ones) is the issue of model identification. In contrast to recursive models, nonrecursive models may be not uniquely solvable even in the instances when the degrees of freedom suggest overidentification.

To ensure identification, certain conditions should be met. Specifically, for a model to be identified, we should have as many instrumental variables as there are variables in reciprocal relationships. And these instrumental variables need to be distributed in particular ways across equations. That is, some of the predictor variables should have no direct paths to certain endogenous variables. An instrument for a specific endogenous variable has to have direct paths to other endogenous variables but not to this specific one.

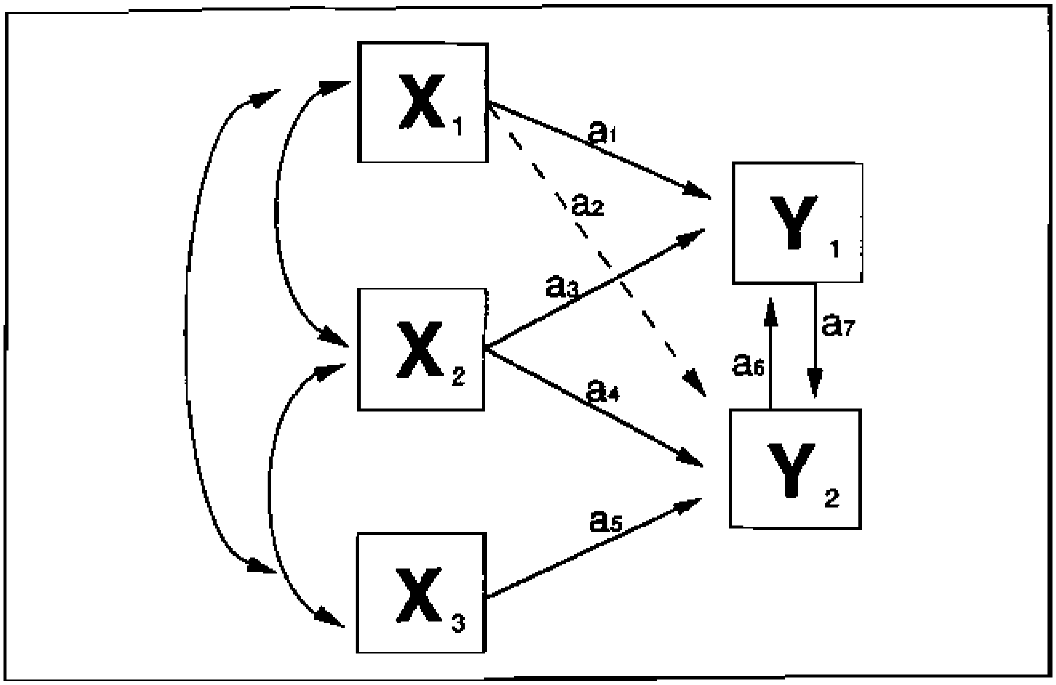


Figure 6.2. Nonrecursive Path Model to Illustrate Model Identification

These requirements are usually described as order condition and rank condition – order condition has to do with omitting one predictor from each equation involving a reciprocal relationship, and the rank condition has to do with the pattern of omitted variables across equations.

To test for the rank condition:
1. Form the matrix of all coefficients (exogenous and endogenous variables together in one matrix). E.g. for the picture on handout (with a2 omitted):

Y1=a1*X1 + a3X2 + 0*X3 + a6*Y2 + e1
Y2=a4*X2 + a5*X3 + a7*Y1 + e2

	X1	X2	X3	Y1	Y2
Y1	a1	a3	0	1	a6
Y2	0	a4	a5	a7	1

2. To test for identification of a particular equation, delete from the matrix (a) the row of that equation and (b) the columns that do not have a zero in the row of the equation of interest.
3. Find a nonzero determinant of rank N-1 from the remaining values. If you do, then the model is identified for that equation. Repeat for the next one.

Sometimes, if a nonrecursive model involves too many parameters or we don't have the necessary instruments, but we have prior information about the relative size of the relationships going in both directions, we can impose equality and proportionality constraints. E.g. if we think that the effects are equal in both directions, we can use EQ to set them to be the same, or we can set one to be twice the size of the other – that frees up one degree of freedom for the model.

Other issue involved in estimating nonrecursive models has to do with specifying starting values (ST command). Oftentimes, nonrecursive models don't converge, and we need to specify the starting values to help it. See Appendices 5-A and 7-A in Kline book for the guidelines in specifying those.

Example (p.255 in Kline):

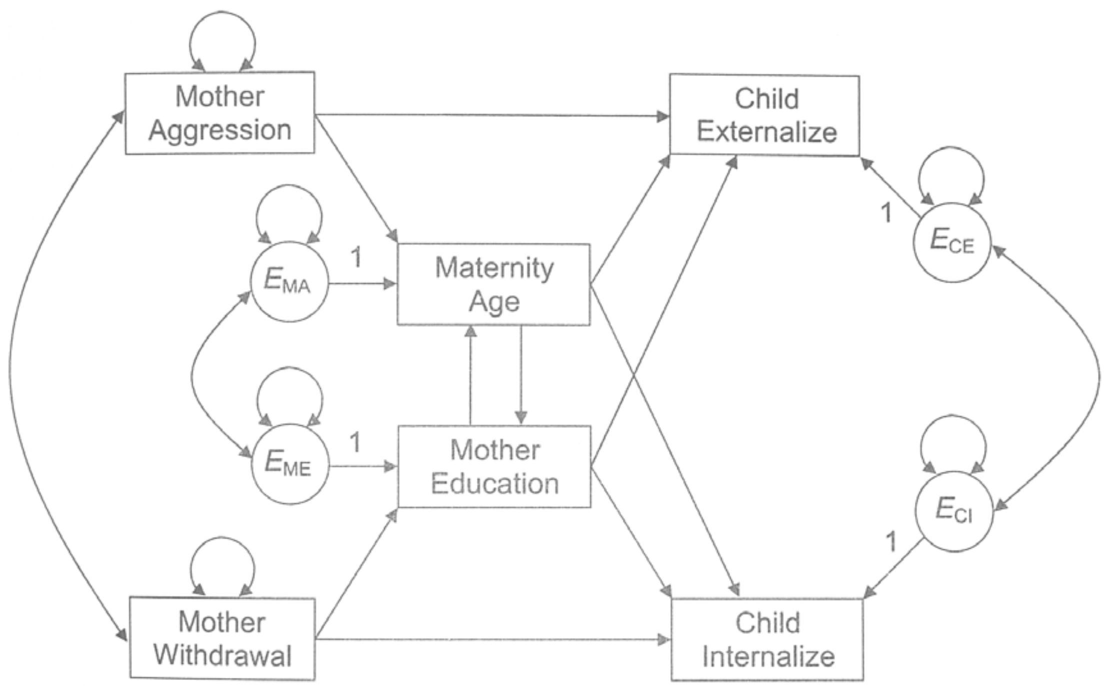


FIGURE 9.4. A nonrecursive path model of transgenerational adjustment problems.