# Advanced Topics

- Multi-group analyses
- Non-recursive models

# I. Multi-group Analyses

- Can compare models across groups
  - » men, women
  - » experimental vs. control
  - » different cultural or sub-cultural groups
  - » no mathematical limit on number of groups, but easier if it is relatively small number (2-4)
- Question:  Do model parameters vary across group?
- i.e., does group membership moderate any of the relationships in the model?
  - » efficient way of testing for interactions

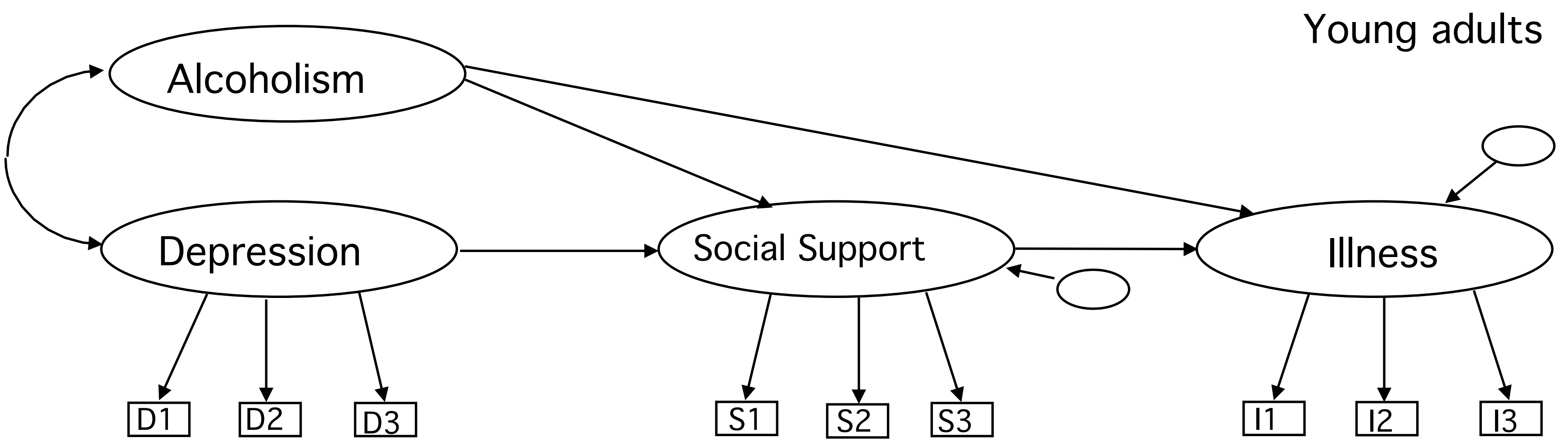
# Eye-balling Group Differences

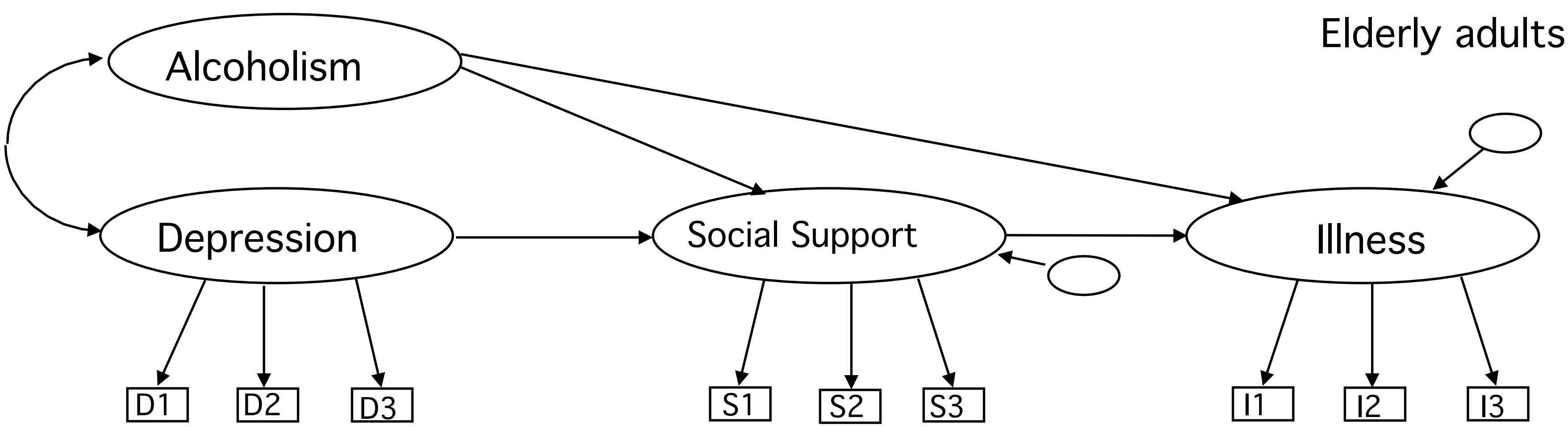
- Imprecise method of looking for differences between groups
  - » run the model (regression, EFA, CFA, path analysis, hybrid model) separately for the two groups
  - » if paths look different (especially if a path is significant in one group but not in the other), conclude that they are different
- Problem: how big of a difference in the sample statistics do we need to conclude that there's a difference in the populations?
  - » need a method for inferential statistical conclusions

# Logic of Inferential SEM Group Comparisons

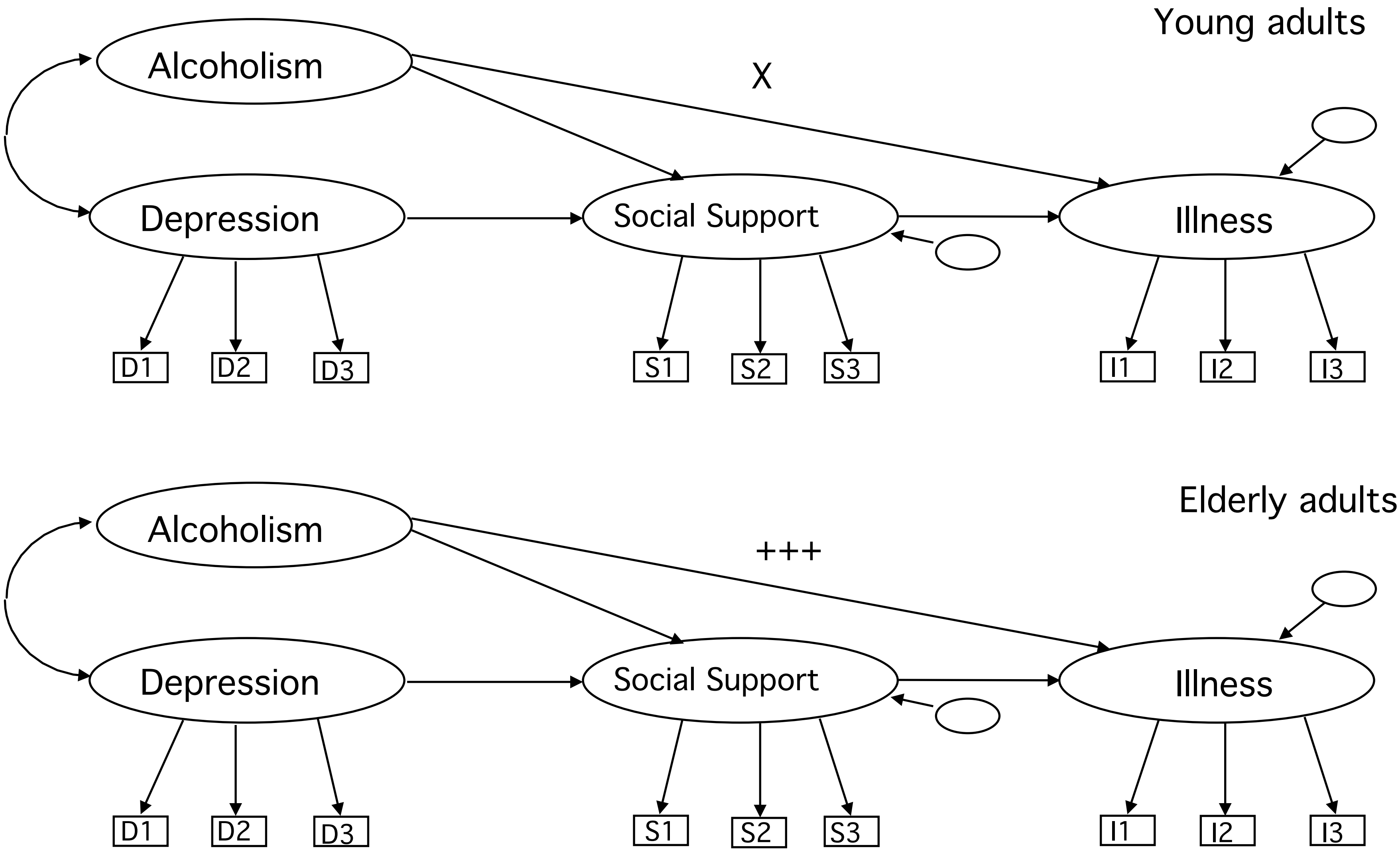
- SEM allows you to estimate parameters separately for the two (or three or four ....) groups, within one run
- You can allow the parameter to vary in size between the groups (the parameter is estimated separately for each group) [Model 1] OR
- You can constrain the solution such that the value of the parameter must be the same across the groups [Model 2]
- Model 2 is nested within Model 1, so a chi-square difference test can be performed
- If fit is sig. better in Model 1, there's a significant group difference

# Multi-Sample Example

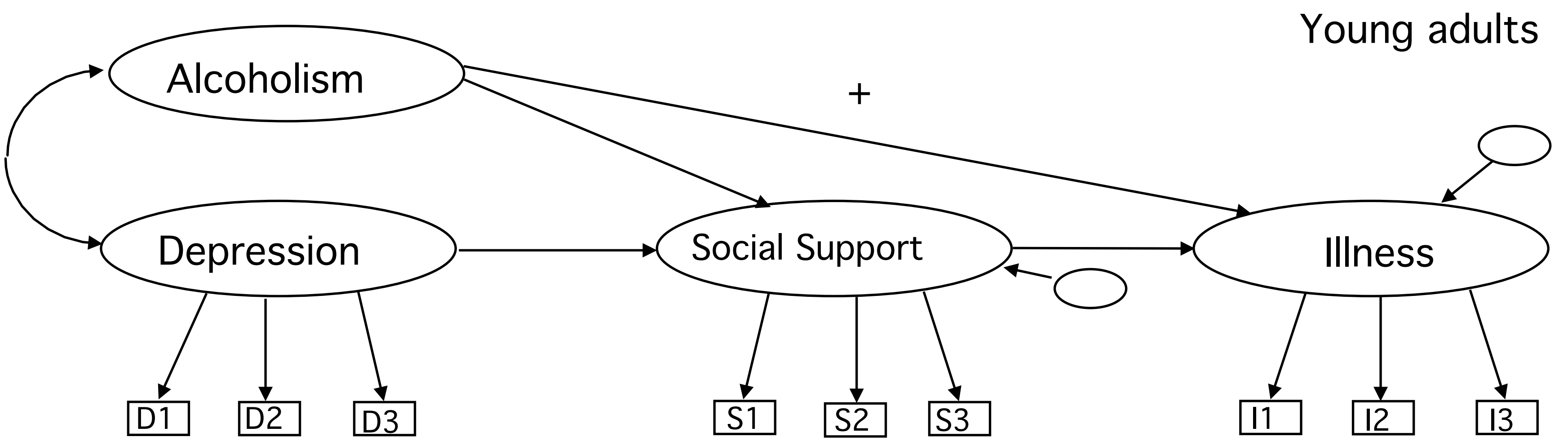


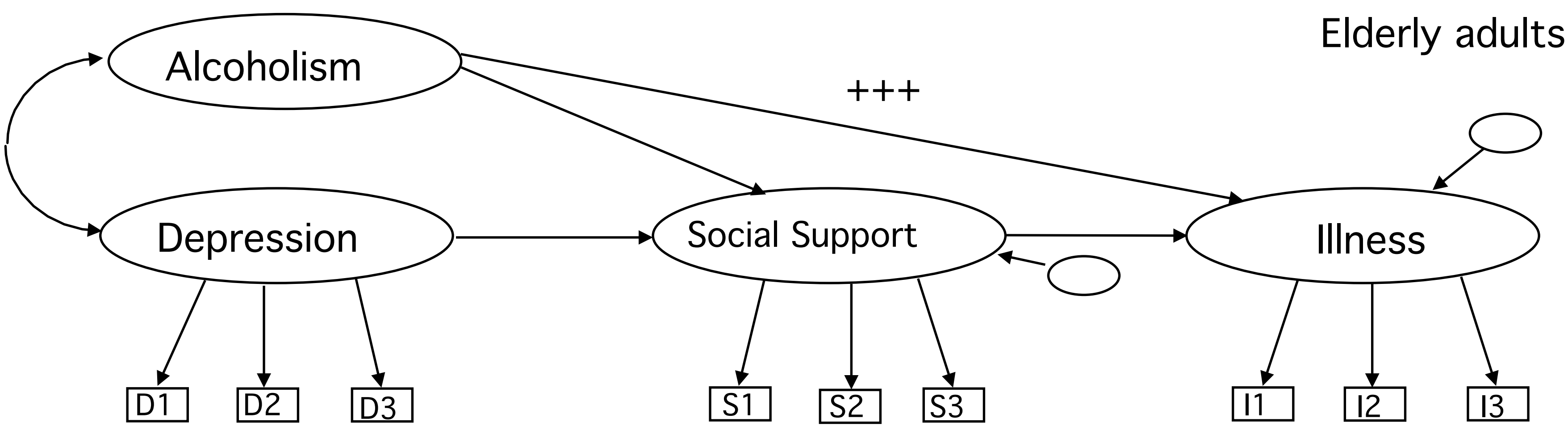



# Multi-Sample Example



# Multi-Sample Example

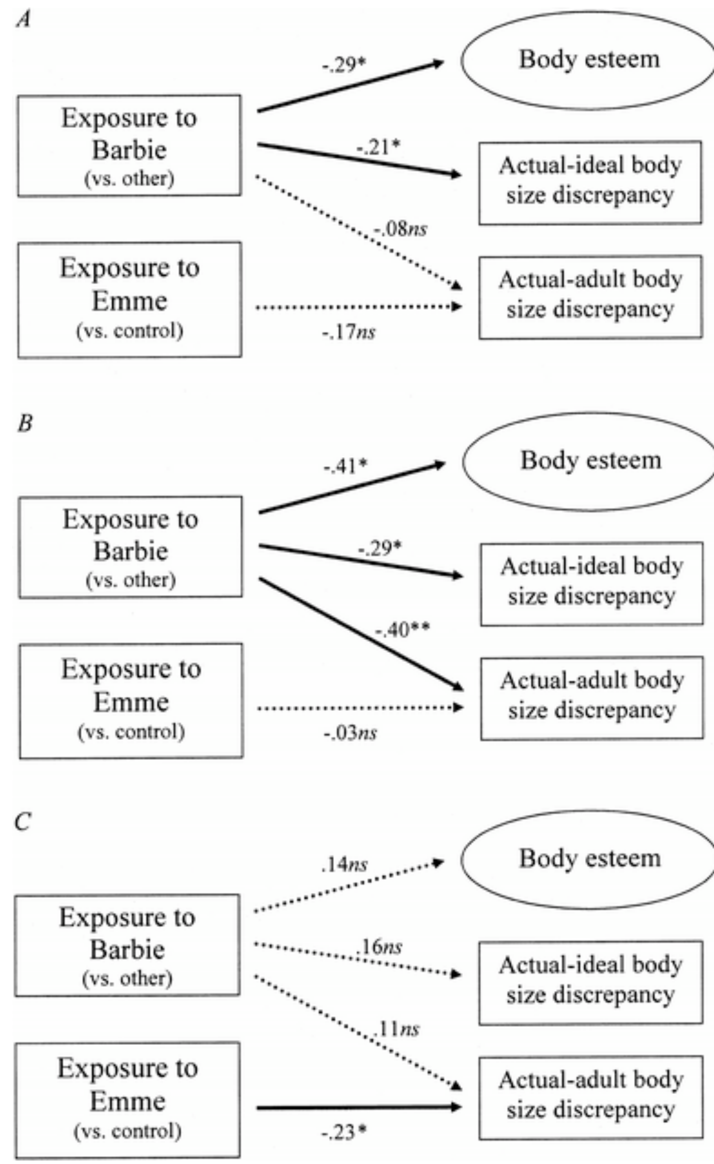


# Dittmar, Halliwell, & Ive (2006)

Experiment
Three different age groups
3 DVs
Latent variable for one DV (body esteem)

Figure 2. Structural equation models of exposure to different dolls on girls' body esteem and desired body shape, separately by year group. For visual clarity, factor loadings for body esteem items and error terms are not shown. A: Year 1 (5½ to 6½ years old). B: Year 2 (6½ to 7½ years old). C: Year 3 (7½ to 8½ years old). Dashed lines indicate nonsignificant paths. * p < .05, one-tailed. ** p < .01, one-tailed



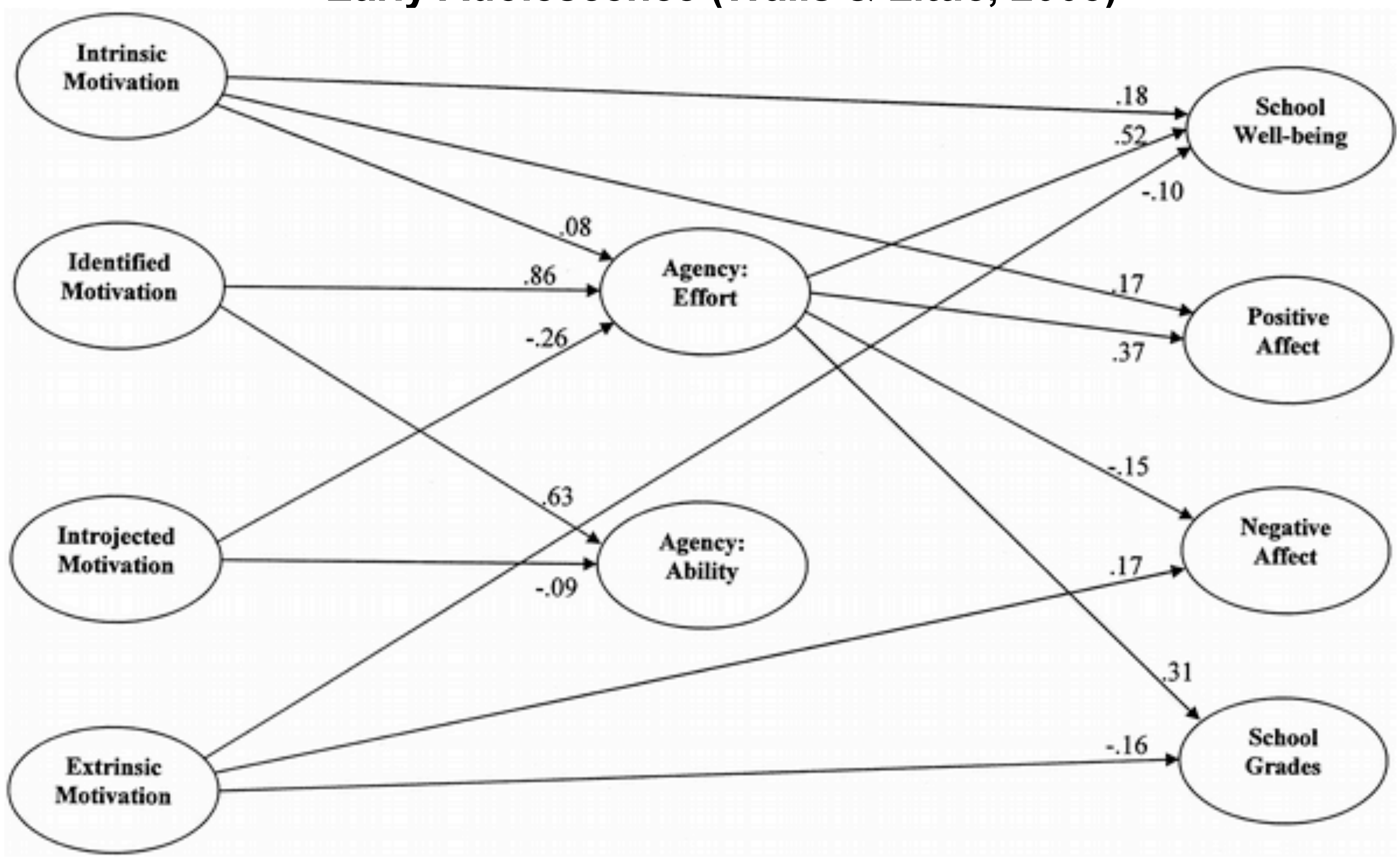
# Strategies for Constraining

- Constrain everything to equality and free up constraints only as suggested by LaGrange modification indices (model building/adding parameters)
- Allow everything to be freely estimated across groups, and constrain paths only as suggested by Wald modification indices (model trimming/reducing parameters)
- Test specific equality constraints, as predicted by theory or suggested by previous research

# Strategies for Constraining

- Take a hierarchical approach to setting and releasing constraints
  - » conceptually similar to an omnibus test in an ANOVA, with follow-up tests only if the omnibus test is significant
- Example: Allow all paths to be estimated freely in each group (Model 1); constrain all paths to equality (Model 2)
  - » If there is no sig difference in the fit of Models 2 and 1, there are no interactions with group. You are done.
  - » If there is a sig difference, take a blockwise approach to finding out where the interaction effects are. For example . . .

## Relations Among Personal Agency, Motivation, and School Adjustment in Early Adolescence (Walls & Little, 2005)




Block 1: 10 paths from exogenous vars
Block 2: 4 paths from mediator vars

OR

Block 1: 5 paths into mediators
Block 2: 9 paths into outcome vars

Block 1: 5 direct path
Block 2: 9 indirect paths

# EQS Syntax

- Separate syntax for each group, appended into one file (with a few caveats)
  - Create diagram
  - Build EQS to create syntax
  - Copy, paste, and edit syntax for each additional group
  - Other edits, as required

```
/TITLE
...        } Input file for group 1
/END
/TITLE
...        } Input file for group 2
/END
...
...
/TITLE
...        } Input file for group m
/END
```

# Syntax Modification for Multi Sample

- GROUPS in the first group
  - » Indicate # of groups in /SPECIFICATION section of the first group
- /CONSTRAINTS in the last group
  - » Cross-group equality constraints must be specified in the last group
- /LMTEST in the last group is recommended
  - » in order to test the appropriateness of the cross-group equality constraints

```
/TITLE
2 GROUP EXAMPLE FROM WERTS ET AL 1976 - GROUP 1
1 FACTOR MODEL WITH UNEQUAL FACTOR CORRELATIONS
/SPECIFICATIONS
CASES = 865; VARIABLES = 4; GROUPS = 2;  ! TWO GROUPS ARE SPECIFIED
/MODEL
 (V1,V2) ON F1 = 5*;              ! TWO FACTOR MODEL, WITH FREE FACTOR LOADINGS
 (V3,V4) ON F2 = 5*;
 VAR F1,F2;                       ! FACTOR VARIANCES FIXED FOR IDENTIFICATION
 COV FF = .5*; VAR E1-E4 = 50*;
/MATRIX
63.382
70.984 110.237
41.710  52.747 60.584
30.218  37.489 36.392 32.295
/END
16 CUMULATED RECORDS OF INPUT MODEL FILE WERE READ (GROUP   1)
/TITLE
2 GROUP EXAMPLE FROM WERTS ET AL 1976  - GROUP 2
/SPECIFICATIONS
CASES = 900; VARIABLES = 4;  ! DIFFERENT SAMPLE SIZE
/MODEL
 (V1,V2) ON F1 = 5*;             ! SAME MODEL AS IN GROUP 1.  REMEMBER THAT
 (V3,V4) ON F2 = 5*;             ! START VALUES MUST BE THE SAME IF CROSS-
 VAR F1,F2;                      ! GROUP EQUALITIES ARE IMPOSED.
 COV FF = .5*; VAR E1-E4 = 50*;
/MATRIX
67.898
72.301 107.330
40.549  55.347 63.203
28.976  38.896 39.261 35.403
/CONSTRAINTS                    ! BETWEEN GROUP CONSTRAINTS GIVEN HERE
(1,F2,F1)=(2,F2,F1);            ! FACTOR CORRELATION  EQUAL ACROSS GROUPS.
SET = GVF;                      ! EACH FACTOR LOADING EQUAL ACROSS GROUPS.
/LMTEST                         ! SEE LM TEST, BELOW, FOR LIST OF CONSTRAINTS.
/END
```

# Output for Multi Sample

- There will be a separate output section for each group
  - » with equations, standardized solutions, etc.
  - » no fit statistics
- Fit statistics are provided at the end
  - » this is fit for the entire, multi-sample, model

# Fit Summary

```
ALL EQUALITY CONSTRAINTS WERE CORRECTLY IMPOSED ! IMPORTANT MESSAGE
GOODNESS OF FIT SUMMARY FOR METHOD = ML
INDEPENDENCE MODEL CHI-SQUARE =        5472.640 ON    12 DEGREES OF FREEDOM
CHI-SQUARE =       10.870 BASED ON     7 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS     0.14440 ! GOOD MODEL FIT
BENTLER-BONETT NORMED    FIT INDEX=       0.998  ! ALL FIT INDICES INDICATE
BENTLER-BONETT NONNORMED FIT INDEX=       0.999  ! GOOD MODEL FIT.
COMPARATIVE FIT INDEX (CFI)      =       0.999
```

# II. Recursive & Nonrecursive Models

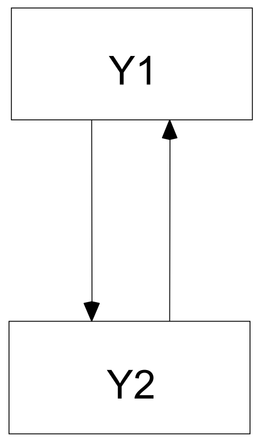
- Recursive models
  - » Unidirectional causal effects
  - » No disturbance correlations between endogenous variables having direct effects between them
  - » All recursive models are identified
  - » May not adequately represent "real world" (Kline, p 237) causal processes which involve feedback loops of some kind

# Recursive & Nonrecursive Models

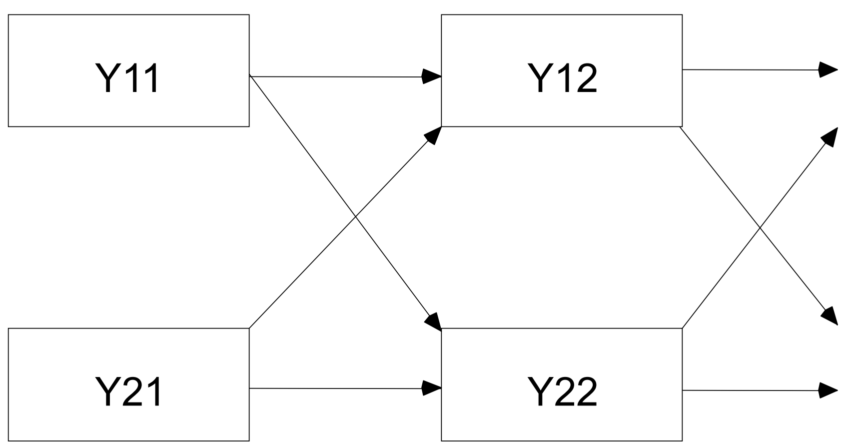
- Nonrecursive models
  - » More difficult to analyze
  - » Not always identified
  - » Very important to determine whether a model is identified before collecting data. May need more variables, can't collect data later.
  - » More observations than freely estimated parameters
  - » All latent variables must be scaled

# Reciprocal Causal Effects

From Kline, 2nd ed, fig 9.1 p 238




Direct Feedback Loop
Cross-sectional Design
Nonrecursive




Panel Model
Longitudinal Design
Often recursive

# Equilibrium

- Assumption when using cross-sectional data to study reciprocal effects that whatever the presumed feedback loop is, that the observed data represents a steady state – the reciprocal process is over.
- This means that the dynamic system has completed its cycle, that the inputs which produced the current steady state are the same over time. One ought not to study this process when it is just beginning.

# Identification of Nonrecursive Models

- $df_M \geq 0$ (# of free parameters does not exceed # of observations).
- Every latent variable has a scale.
- Conditions for identification depend on pattern of disturbance correlations
- Models with unanalyzed associations (all possible disturbance correlations)
  - » Necessary: Order condition
  - » Sufficient: Rank condition
- Order and rank may be too conservative if model has no disturbance correlations or fewer than all possible. Could be identified even if order and/or rank not met

# Order Condition

- A counting rule applied to each endogenous variable in a nonrecursive model
  - » With all possible disturbance correlations, or
  - » Block recursive
- If the order condition is not met for a particular endogenous variable, then the equation for that variable is underidentified.
- "the number of excluded variables for each endogenous variable equals or exceeds the total number of endogenous variables minus 1." (Kline, p242)

# The order condition

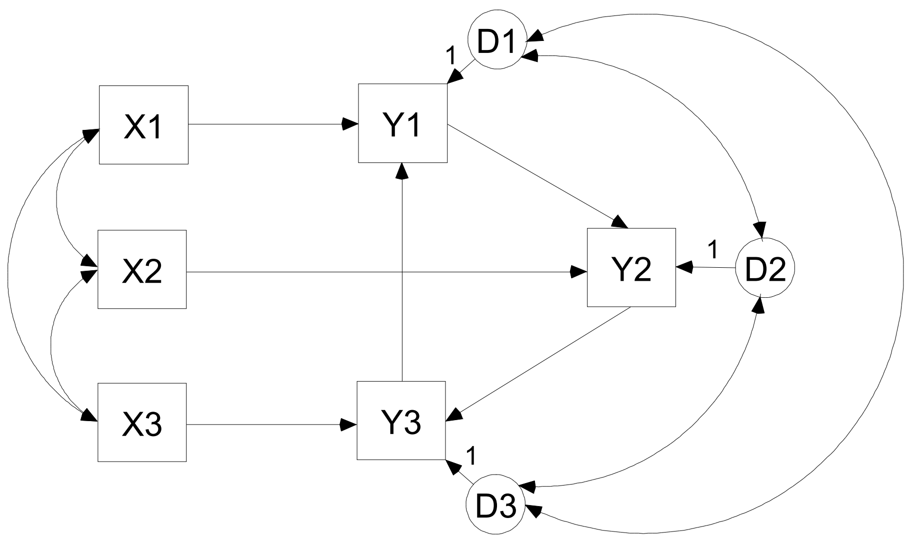



Figure 9.2(a) from Kline 2nd ed, p 241
All possible disturbance correlations

There are six variables in this model: three exogenous, three endogenous.

3 – 1 = 2: two of the 6 variables in the model must be excluded from the equation of each endogenous variable.

$Y_1 = B_{11}X_1 + B_{13}Y_3 + 0X_2 + 0X_3 + 0Y_2 + e$ [3 excluded]

$Y_2 = B_{22}X_2 + B_{21}Y1 + 0X_1\ 0X_3 + 0Y_2 + e$ [3 excluded]

$Y_3 = B_{33}X_3 + B_{32}Y_2 + 0X_1 + 0X_2 + 0Y_1 + e$ [3 excluded]

The order condition is met.

# The order condition

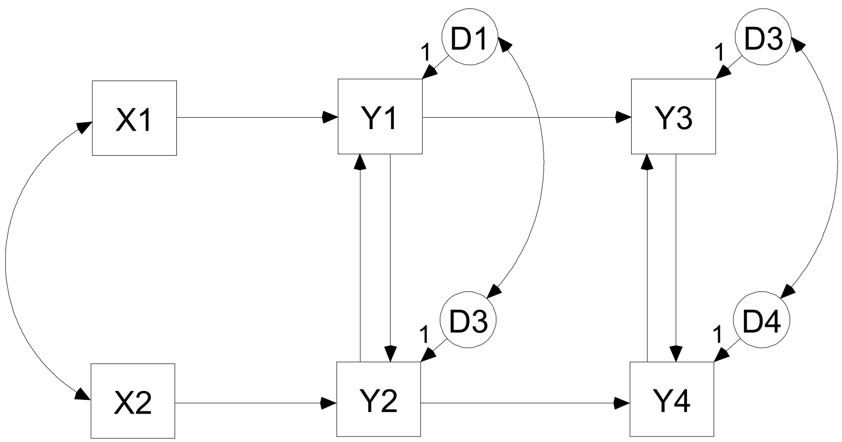



Figure 9.2(b) from Kline 2nd ed, p 241
All possible disturbance correlations within recursively related blocks

In this block recursive model, there are two recursively related blocks with two endogenous variables in each block. Count each block separately.

2 – 1 = 1: one of the 4 vars in each block must be excluded from the equation of each endogenous var.

$Y_1 = B_{11}X_1 + B_{12}Y_2 + 0X_2 + e$

$Y_2 = B_{22}X_2 + B_{21}Y1 + 0X_1 + e$

$Y_3 = B_{31}Y_1 + B_{34}Y_4 + 0Y_2 + e$

$Y4 = B_{42}Y_2 + B_{43}Y_3 + 0Y_1 + e$

1 variable excluded from each equation; the order condition is met.

# The Rank Condition

- The order condition is necessary, but not sufficient for a nonrecursive model to be identified.
- The rank condition is a ‘sufficient’ condition, so if it is satisfied, then the nonrecursive model will be identified.
- The rank condition means that each endogenous variable in a feedback loop must have a unique pattern of direct effects on it from variables outside the loop. (Kline)
- Complex matrix description requires familiarity with linear algebra.
- Berry (1984) provides an algorithm for determining rank that is easier to use.
- Kline provides a set of steps

# Checking the Rank Condition

- **Step 1** - Form a *system matrix* (y rows, x and y columns) for the observed variables.
- a "1" indicates that the column variable has a direct effect on the row variable.
- a "1" also appears in each of the elements where the row and column variable are the same.
- all remaining elements are 0, representing excluded direct effects.

$Y_1 = B_{11}X_1 + B_{13}Y_3 + 0X_2 + 0X_3 + 0Y_2 + e$

$Y_2 = B_{22}X_2 + B_{21}Y1 + 0X_1\ 0X_3 + 0Y_2 + e$

$Y_3 = B_{33}X_3 + B_{32}Y_2 + 0X_1 + 0X_2 + 0Y_1 + e$

| | $X_1$ | $X_2$ | $X_3$ | $Y_1$ | $Y_2$ | $Y_3$ |
|---|---|---|---|---|---|---|
| $Y_1$ | 1 | 0 | 0 | 1 | 0 | 1 |
| $Y_2$ | 0 | 1 | 0 | 1 | 1 | 0 |
| $Y_3$ | 0 | 0 | 1 | 0 | 1 | 1 |

# Checking the Rank Condition

- **Step 2**: ‘Read’ the matrix
- Each row represents the equation for each endogenous variable
- Repeat for each equation:

1. Cross out all the entries in the row of interest.
2. Cross out the elements in any column with a “1” in the row of interest
3. Form a new matrix with the entries that remain
4. Delete all zero rows
5. Delete duplicate rows
6. Delete rows that can be reproduced by adding other rows together.
7. # of remaining rows = rank

| | $X_1$ | $X_2$ | $X_3$ | $Y_1$ | $Y_2$ | $Y_3$ |
|---|---|---|---|---|---|---|
| $Y_1$ | 1 | 0 | 0 | 1 | 0 | 1 |
| $Y_2$ | 0 | 1 | 0 | 1 | 1 | 0 |
| $Y_3$ | 0 | 0 | 1 | 0 | 1 | 1 |

| | | |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 1 | 1 |

| | | |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 1 | 1 |

| | | |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 1 | 1 |

Rank = 2 for each equation. Rank must be > or = # of endogenous variables minus 1: 3 – 1. Therefore, rank condition satisfied. Model identified.