Brief Overview of LISREL & Related Programs & Techniques (Optional)

Richard Williams, University of Notre Dame, https://www3.nd.edu/~rwilliam/
Last revised April 6, 2015

STRUCTURAL AND MEASUREMENT MODELS: LISREL. We have focused on structural models. Such models assume that all variables are measured without error. Of course, this assumption is often not reasonable. As we saw earlier in the course,

- Random measurement error in the dependent variable does not bias regression coefficients. However, it does result in larger standard errors.

- Random measurement error in the independent variables results in biased estimates. In the case of a bivariate regression, estimates will be biased toward zero. With more IVs, the bias can be upwards or downwards.

- Systematic error, of course, can produce either an upward or downward bias.

Factor analysis is one way of dealing with measurement error. With factor analysis, a large number of items are reduced to a smaller number of factors, or "latent variables". For example, 7 personality measures might be reduced into a single "locus of control" scale. This scale would be more reliable than any of the individual measures that constructed it.

Factor analysis can be either

- exploratory — the computer determines what the underlying factors are

- confirmatory — the researcher specifies what factor structure she thinks underlies the measures, and then tests whether the data are consistent with her hypotheses.

Programs such as LISREL make it possible to combine structural equation modeling and confirmatory factor analysis. (I understand programs like AMOS and M-Plus and the gllamm addon routine to Stata can do these sorts of things too but I have never used them. These programs may be easier to use and/or cheaper than LISREL, so you may want to check them out if you want to do heavy-duty work in this area.) Some traits of LISREL:

- There is both a measurement model and a structural model.

  - The measurement model indicates how observed indicators are linked to underlying latent variables. (e.g. X1 and X2 may be indicators of Locus of control; X3 and X4 may be indicators of Socio-economic status).

  - The structural model indicates how the latent variables are linked to each other.

  - By controlling for measurement error, a correctly specified LISREL model makes it possible to obtain unbiased estimates of structural coefficients. (Of course, getting the model correctly specified is the trick.)

- LISREL can handle a wide array of problems and models. These include
    - Models with measurement error
    - Nonrecursive models
    - Manova-type problems
    - Multiple group comparisons (e.g. you can have separate models for blacks & whites)
    - Tests of constraints (e.g. two or more coefficients equal each other, a subset of coefficients equals zero, parameters are equal across populations)
    - Confirmatory factor analysis models
    - Ordinal regression
    - Hierarchical Linear Models

I'll give just a few examples, not all of which I will talk about in class. A free trial edition and a limited but free student edition of LISREL and a LISREL tutorial can currently be found at

http://www.ssicentral.com/lisrel/downloads.html

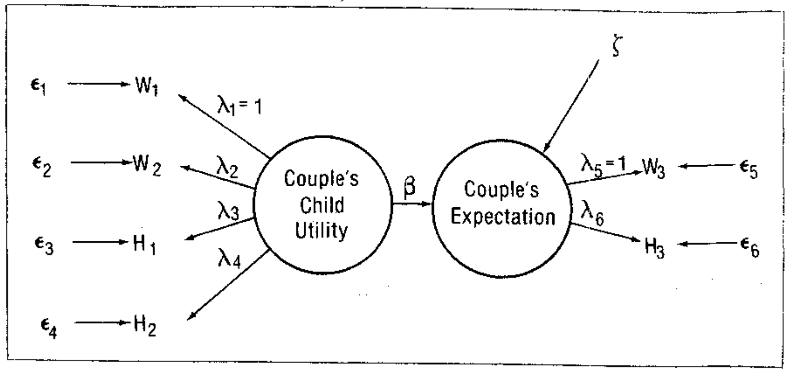
LISREL EXAMPLE 1: Measurement and Structural Models Combined. In their classic 1982 paper, "Beyond Wives Family Sociology: A Method for Analyzing Couple Data," Thomson and Williams estimate both measurement and structural parameters in a series of models of couple childbearing expectations. In their data, husbands and their wives were presented with several possible consequences of having another child within 20 months.

- Products of their subjective probability of each consequence (0 = no chance to 10 = certain) and their evaluations of the consequence (-3 = extremely bad thru +3 = extremely good) were constructed to form "subjective expected utilities" of another child. The subjective expected utilities of "a fulfilled family life" (W1 and H1) and "watching another child grow and develop" (W2 and H2) were used as multiple indicators of child utility.

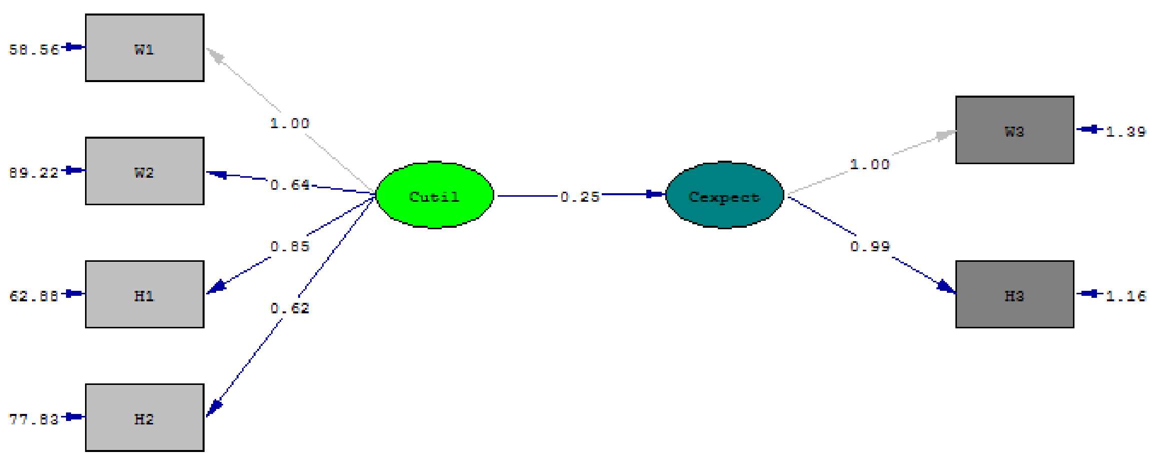
- Also, respondents were asked to estimate the likelihood that the couple would have another child within 20 months (1 = extremely unlikely thru 7 = extremely likely.) Responses of both partners (W3 and H3) were used as multiple indicators of couple childbearing expectations.

Thomson and Williams began by estimating a "couple" model, in which the wife's and husband's responses about the utility of another child are all imperfectly measured indicators of a single latent variable, the couple's child utility. Here is their original diagram for this model:

FIGURE 1. COUPLE'S UTILITY OF ANOTHER CHILD AND COUPLE'S CHILDBEARING EXPECTION (VARIABLE LABELS DEFINED IN TABLE 1)



As part of its output, LISREL can generate a (somewhat ugly and not always accurate) drawing of the path diagram for you. I would probably want to draw the diagram myself for a paper submission but the LISREL diagram is good for making sure you've specified the model you thought you did. The path diagram LISREL produces is



Minimum Fit Function Chi-Square=58.74, df=8, P-value=0.00000

The following LISREL program estimates this model. LISREL syntax used to be a nightmare but in recent years it has gotten more user-friendly.

LISREL Program	Interpretation

`Beyond Wive's Family Sociology - Model 1`
`Observed Variables W1 W2 H1 H2 W3 H3`
`Sample size 340`
`Correlation Matrix`
` 1`
` 0.470   1`
` 0.460   0.270   1`
` 0.312   0.223   0.495   1`
` 0.628   0.421   0.498   0.381   1`
` 0.596   0.347   0.586   0.422   0.816   1`
`Means`
` 11.36   22.34   9.75   18.50   3.64   3.66`
`Standard Deviations`
` 11.45   10.89  10.73   10.30   2.66   2.60`
	We start by reading in the data.

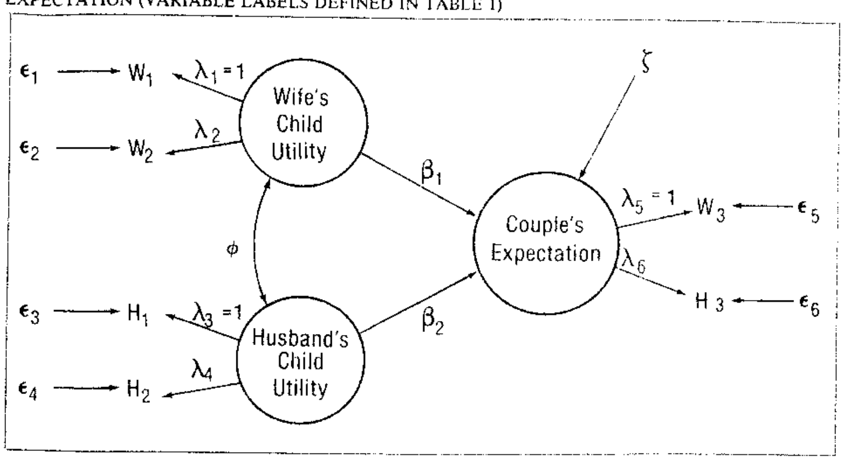
`Latent variables Cutil Cexpect`	This tells LISREL we have 2 latent variables

`W1 = 1*Cutil`
`W2 = Cutil`
`H1 = Cutil`
`H2 = Cutil`
 	
`W3 = 1*Cexpect`
`H3 = Cexpect`
	This is the measurement part of the model. It shows how the 6 observed variables are related to the 2 underlying latent variables.

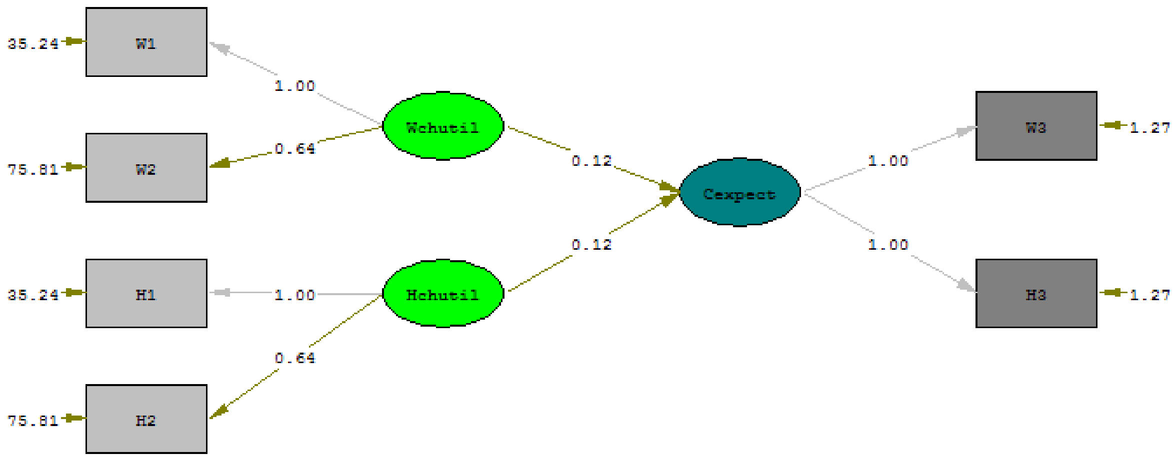
`Cexpect = Cutil`	This is the structural model
`Path Diagram` `End of Problem`	This causes LISREL to include a path diagram as part of its output.

Thomson and Williams argued that the fit of this model was unacceptable and that rather than having a single couple utility variable, there should be two separate variables, one for husbands and one for wives:

FIGURE 2. WIFE'S AND HUSBAND'S UTILITY OF ANOTHER CHILD AND COUPLE'S CHILDBEARING EXPECTATION (VARIABLE LABELS DEFINED IN TABLE 1)



Also, in their final model (which for some reason they hid in the discussion instead of presenting in the tables) all corresponding parameters between wives and husbands were constrained to be equal. The final LISREL-produced diagram looks like this:



Minimum Fit Function Chi-Square=27.46, df=13, P-value=0.011

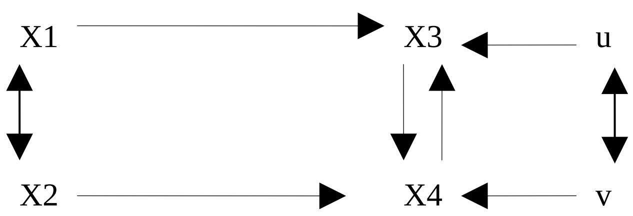
The LISREL syntax that produces this is

LISREL Program	Interpretation
Beyond Wive's Family Sociology - Model 2F Observed Variables W1 W2 H1 H2 W3 H3 Sample size 340 Correlation Matrix 1 0.470 1 0.460 0.270 1 0.312 0.223 0.495 1 0.628 0.421 0.498 0.381 1 0.596 0.347 0.586 0.422 0.816 1 Means 11.36 22.34 9.75 18.50 3.64 3.66 Standard Deviations 11.45 10.89 10.73 10.30 2.66 2.60	Read in the data.
Latent variables Wchutil Hchutil Cexpect	Tells LISREL there are 3 latent variables.
W1 = 1*Wchutil W2 = Wchutil H1 = 1*Hchutil H2 = Hchutil W3 = 1*Cexpect H3 = Cexpect	Measurement model – shows how the 6 observed variables are related to the 3 latent variables
Cexpect = Wchutil Hchutil Let Wchutil and Hchutil covary	The Structural Model. The 2 exogenous latent variables have a non-zero covariance.

`Equal Error Variances: W1 H1`
`Equal Error Variances: W2 H2`
`Set the Path from Wchutil to W2 Equal to the`
`    Path from Hchutil to H2`
`Set the variance of Wchutil equal to the`
`    variance of Hchutil`
 	
`Equal Error Variances: W3 H3`
`Set the path from Cexpect to H3 equal to 1`
`Set the path from Wchutil to Cexpect equal to`
`   the path from Hchutil to Cexpect`
	Imposes equality constraints between corresponding husband and wife measurement and structural parameters.

`Path Diagram` `End of Problem`	Produce Path diagram, end program.

LISREL EXAMPLE 2: Nonrecursive Models – The LISREL Alternative to 2SLS. Recall the nonrecursive model we previously estimated with 2sls:



We only have single indicators of each X, so LISREL's measurement model is not used here. Here is a LISREL program that will estimate this model.

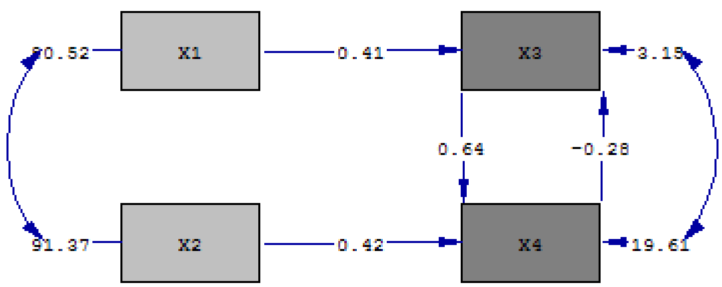
LISREL Program	Interpretation

`Nonrecursive Model Example`
 	
`Observed Variables: X1 X2 X3 X4`
`Correlation Matrix`
` 1`
` -.154   1`
` .812    -.354   1`
` .290     .608   -.127   1`
`Standard Deviations`
` 8.97349 9.55875 3.98072 5.06170`
`Sample size = 500`
	Read in data

`Equation: X3 = X1 X4`
`Equation: X4 = X2 X3`
`Let the errors between X3 X4 Correlate`
	Structural model. The residuals of X3 and X4 are allowed to be correlated.

`Path Diagram`
`Method of Estimation: Maximum`
`   Likelihood`
`End of Problem`
	Produce path diagram. Use ML to estimate the model (could have used Two Stage Least Squares if we preferred).

Here is the LISREL diagram:



The results are almost the same as we got with 2sls, but that won't always be the case.

Also, recall our previous example of a nonrecursive model of peer influence from Duncan-Haller-Portes:

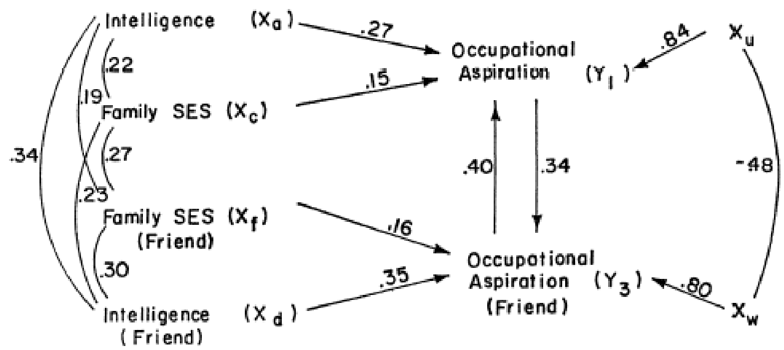


FIG. 2.—Model II

To estimate this in LISREL,

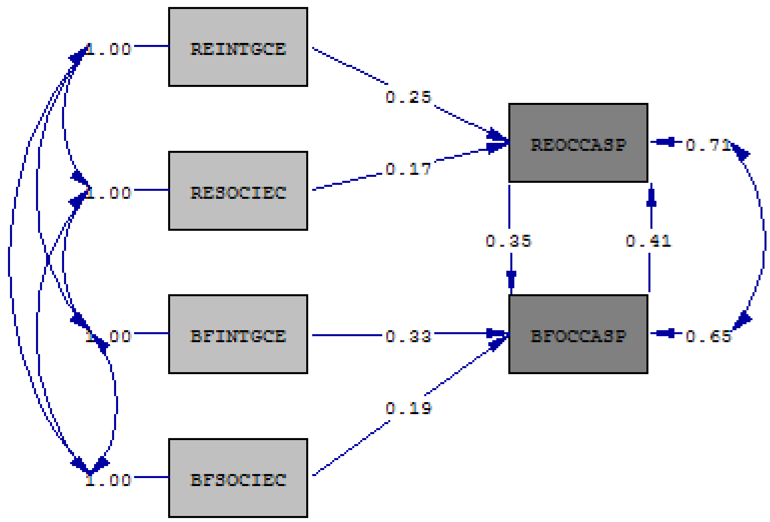
LISREL Program	Interpretation

`Peer Influences on Ambition - Duncan Haller`
`   Portes Replication`
`---------------------------`
`Observed Variables from File EX8.LAB`
`Correlation Matrix from File EX8.COR`
`Reorder Variables: REPARASP REINTGCE RESOCIEC`
`    BFSOCIEC BFINTGCE BFPARASP`
`    REOCCASP 'RE EDASP' 'BF EDASP' BFOCCASP`
`Sample Size 329`
	Read in data

`Relationships`
`   REOCCASP = REINTGCE RESOCIEC BFOCCASP`
`   BFOCCASP = BFINTGCE BFSOCIEC REOCCASP`
`Let the errors of REOCCASP and BFOCCASP covary`
	Specify Structural relationships

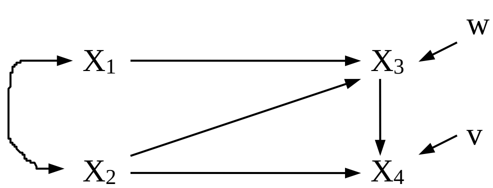
`Path Diagram` `End of Problem`	Include diagram in output

The LISREL produced diagram is



Results differ slightly because Duncan-Haller-Portes used 2SLS and LISREL is using Maximum Likelihood. You can also tell LISREL to use 2SLS instead, in which case the results are identical.

Example 3: Using LISREL to Decompose Correlations. In HW 7, you were given the following model:



Here is how LISREL can estimate this model:

LISREL Program	Interpretation

`Path Model from Homework 7`
`Observed variables X1 X2 X3 X4`
`Correlation Matrix`
`   1.00`
`    .60     1.00`
`   0.54     0.58   1.00`
`    .57     0.79   0.79  1.00`
`Means`
`      0          0        0      0`
`Standard Deviations`
`      1          1        1      1`
`Sample Size = 1000`
	Read in data

`X3 = X1 X2` `X4 = X2 X3`	Structural Model

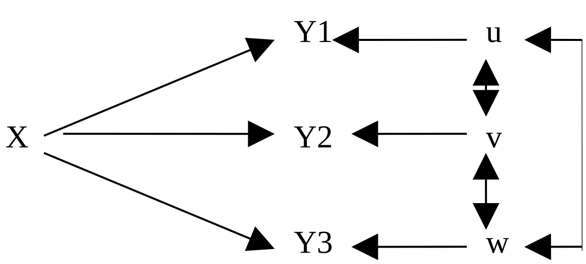
`LISREL OUTPUT EF`
 	
`Path Diagram`
`End of Problem`
	Request Effects and Path Diagram as part of the output

Note the LISREL OUTPUT Card. The EF parameter asks LISREL to give the total direct and indirect effects of each variable. Part of the Output is

```
Total and Indirect Effects

        Total Effects of X on Y

                X1          X2
            --------    --------
      X3       0.30        0.40
             (0.03)      (0.03)
               9.73       12.98

      X4       0.15        0.70
             (0.02)      (0.02)
               9.20       32.70


        Indirect Effects of X on Y

                X1          X2
            --------    --------
      X3        - -         - -

      X4       0.15        0.20
             (0.02)      (0.02)
               9.20       11.78


        Total Effects of Y on Y

                X3          X4
            --------    --------
      X3        - -         - -

      X4       0.50         - -
             (0.02)
              28.06
```

Hence, LISREL can do some of the decomposition of effects that you originally did by hand. In complicated models, such decompositions are difficult to compute manually. Knowing the total effect of a variable can be useful, since it tells you how much a 1 unit change in an IV will change the expected value of a DV.

Example 4: Using LISREL for Manova. We previously worked this problem using SPSS Manova:



Here is how LISREL can estimate this model:

<table>
<tr><th>LISREL Program</th><th>Interpretation</th></tr>
<tr><td><pre>Manova Using LISREL Example

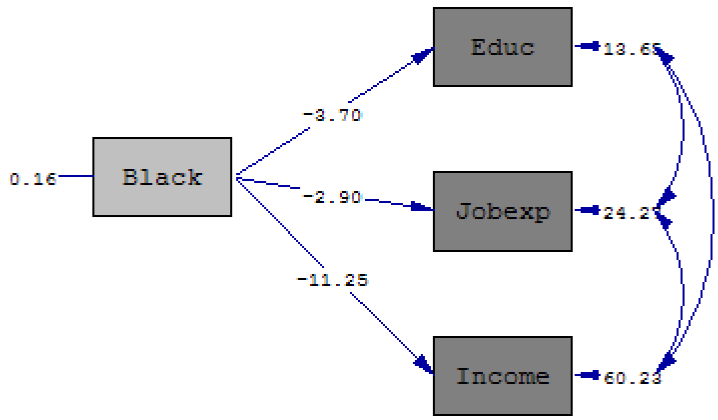
Observed Variables Educ Jobexp Income Black
Correlation Matrix
   1.0000000
   -.1274659   1.0000000
    .8119504   .2895017   1.0000000
   -.3721649   -.2294014   -.5019793
1.0000000
Standard Deviations
   3.9807150   5.0617034   8.9734914   .4004006
Means
   13.16   13.52   27.79   .20
Sample size 500</pre></td><td>Read in data</td></tr>
<tr><td><pre>Educ Jobexp Income = Black</pre></td><td>Structural Model. This is a shorthand way of saying that all the DVs should be regressed on all the IVs. If there were more IVs, we would just add them on the right.

To do a global test of whether Black has any effects, change this to

<code>Educ Jobexp Income = 0*Black</code>

When we do this, we get a chi-square statistic of 129.6 with 3 d.f. This is highly significant, suggesting that Black affects at least one of the DVs.</td></tr>
<tr><td><pre>Let the Error Covariances of Educ - Income be
      free</pre></td><td>This is a shortcut way of saying that all the endogenous residuals should be allowed to freely covary. The shortcut works because Educ, Jobexp and Income were consecutively ordered in the input data</td></tr>
<tr><td><pre>Path Diagram
Method of Estimation: Maximum Likelihood
End of Problem</pre></td><td>Request Path Diagram as part of the output</td></tr>
</table>

The LISREL diagram is



The output includes

```
     Educ =  - 3.70*Black, Errorvar.= 13.65, Rý = 0.14
              (0.41)                  (0.87)
              -8.95                    15.78

   Jobexp =  - 2.90*Black, Errorvar.= 24.27, Rý = 0.053
              (0.55)                  (1.54)
              -5.26                    15.78

   Income =  - 11.25*Black, Errorvar.= 60.23, Rý = 0.25
              (0.87)                   (3.82)
              -12.95                    15.78
```

The t-values (-8.95, -5.26, -12.95) are the square roots of the Univariate F tests that Manova reported (80.06600, 27.66301, 167.76045), i.e. Manova earlier reported

```
EFFECT .. BLACK (Cont.)
 Univariate F-tests with (1,498) D. F.

 Variable    Hypoth. SS   Error SS Hypoth. MS   Error MS           F  Sig. of F

 EDUC        1095.20000 6812.00000 1095.20000   13.67871    80.06600       .000
 JOBEXP       672.80000 12112.0000  672.80000   24.32129    27.66301       .000
 INCOME      10125.0000 30056.2500 10125.0000   60.35392   167.76045       .000
```

Example 5: Using LISREL for Group Comparisons. We previously worked this problem using SPSS (homeworks 5 & 6). Separate regressions are run for 225 males and 275 females, yielding the following:

<pre>-> * Group comparisons -- males only.
-> REGRESSION /VARIABLES EDUC JOBEXP female INCOME
->   /select = female eq 0
->   /DEPENDENT INCOME
->   /ENTER EDUC JOBEXP.
 	
Multiple R   .80941
R Square   .65514
Adjusted R Square   .65204
Standard Error   6.77361
 	
Analysis of Variance
                    DF   Sum of Squares   Mean Square
Regression   2   19350.45823   9675.22912
Residual   222   10185.76399   45.88182
 	
F =   210.87283   Signif F = .0000
 	
------------------ Variables in the Equation ------------------
 	
Variable   B   SE B   Beta   T Sig T
 	
EDUC   .819538   .107082   .319493   7.653 .0000
JOBEXP   1.384972   .089525   .645812   15.470 .0000
(Constant)   -.929413   1.497770   -.621 .5355</pre>
	<pre>-> * Group comparisons -- females only.
	-> REGRESSION /VARIABLES EDUC JOBEXP female INCOME
	-> /select = female eq 1
	-> /DEPENDENT INCOME
	-> /ENTER EDUC JOBEXP.
	
	Multiple R .68469
	R Square .46881
	Adjusted R Square .46490
	Standard Error 4.68853
	
	Analysis of Variance
	                    DF Sum of Squares Mean Square
	Regression 2 5276.94290 2638.47145
	Residual 272 5979.19347 21.98233
	
	F = 120.02693 Signif F = .0000
	
	------------------ Variables in the Equation ------------------
	
	Variable B SE B Beta T Sig T
	
	EDUC 1.525582 .100410 .684139 15.194 .0000
	JOBEXP -.004920 .077359 -.002864 -.064 .9493
	(Constant) 5.470545 1.589722 3.441 .0007</pre>

Here is how a LISREL problem can be set up to do multiple-group comparisons:

LISREL Program	Interpretation
Group 1: Males Observed Variables Educ Jobexp Income Correlation Matrix 1 .330 1 .532 .751 1 Standard Deviations 4.47657 5.35450 11.4829 Means 11.2222 14.1111 27.811 Sample Size 225	Read in the data for Group 1
Income = Educ Jobexp	Define the structural model for group 1
Group 2: Females Correlation Matrix 1 -.192 1 .685 -.134 1 Standard Deviations 2.87427 3.73073 6.4094 Means 10.6364 12.3636 21.636 Sample Size 275	Read in the data for group 2
Let the path from Educ to Income be free Let the path from Jobexp to Income be free Let the error variance of Income be free	By default, LISREL constrains parameters (except exogenous variances and covariances) to be equal across groups. This will let structural parameters be free to vary.
Path Diagram End of Problem	Create path diagram, end run.

Here are excerpts from the LISREL Output:

Group 1: Males	Group 2: Females
Structural Equations Income = 0.82*Educ + 1.38*Jobexp, Errorvar.= 45.54, R² = 0.65 (0.11) (0.089) (4.31) 7.65 15.49 10.56	Structural Equations Income = 1.53*Educ - 0.0044*Jobexp, Errorvar.= 21.80, R² = 0.47 (0.10) (0.077) (1.87) 15.23 -0.057 11.68

Note that LISREL's numbers are the same as we got from SPSS, EXCEPT LISREL does not include an intercept term. This is because we've never actually told the program to use the information about the means. It is fairly common with LISREL models just to ignore the means and treat all variables as being centered about their group mean so as to have a mean of zero. But, if you want to use the means in the analysis, you can; it just gets a little more complicated.

Also, recall that we got these results when we pooled the male-female samples and included a dummy variable for gender:

```
Multiple R           .73111
R Square             .53452          R Square Change    .53452
Adjusted R Square    .53170          F Change        189.85393
Standard Error      6.53532          Signif F Change    .0000

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           3         24326.24810       8108.74937
Residual           496         21184.38940         42.71046

F =     189.85393       Signif F =  .0000

------------------ Variables in the Equation ------------------

Variable              B        SE B       Beta         T  Sig T

EDUC           1.281368     .080380    .495122    15.941  .0000
JOBEXP          .773848     .065286    .373709    11.853  .0000
FEMALE        -4.071767     .599007   -.212324    -6.798  .0000
(Constant)     2.511457    1.269321                1.979  .0484
```

Here is what we get in LISREL when we constrain parameters to be equal across groups (i.e. we get rid of the "let the path/error variance" be free statements:

Group 1: Males	Group 2: Females
LISREL Estimates (Maximum Likelihood) Structural Equations Income = 1.28*Educ + 0.77*Jobexp, Errorvar.= 42.56, R² = 0.61 (0.080) (0.065) (2.70) 15.93 11.85 15.75	LISREL Estimates (Maximum Likelihood) Structural Equations Income = 1.28*Educ + 0.77*Jobexp, Errorvar.= 42.56, R² = 0.29 (0.080) (0.065) (2.70) 15.93 11.85 15.75

LISREL gives the same estimates as SPSS, i.e. running multi-group models in LISREL with parameters constrained to be equal gives the same result as running a regression with dummy variables for group membership but no interactions. Also, LISREL reports

```
                 Global Goodness of Fit Statistics

                        Degrees of Freedom = 3
          Minimum Fit Function Chi-Square = 168.09 (P = 0.0)
  Normal Theory Weighted Least Squares Chi-Square = 138.45 (P = 0.0)
          Estimated Non-centrality Parameter (NCP) = 135.45
      90 Percent Confidence Interval for NCP = (100.59 ; 177.73)
```

The large chi-square value tells us that one or more parameters is different across the 2 populations. This is the same conclusion we reached before using incremental F tests.

LISREL provides a fairly powerful, and convenient, means for examining group differences. Once you have the basic model set up, you can free or constrain parameters as you wish. LISREL provides diagnostic information, called Modification Indices, that can help you to identify equality constraints that are especially suspect. In this case, LISREL Says

```
        The Modification Indices Suggest to Add the
  Path to  from      Decrease in Chi-Square    New Estimate
 Income    Income            28.5                 -0.29 IN GROUP 2
 Income    Educ              34.1                  1.70 IN GROUP 2
 Income    Jobexp           178.2                 -0.02 IN GROUP 2
```

In other words, LISREL is saying that our most problematic constraint is in saying that the effect of Job experience is the same for both men and women. This is consistent with what our earlier analysis showed. We could let that path be free to vary across groups and then re-assess whether further changes were needed.

In effect, then, LISREL can help you to identify which interaction effects should be in your models and which ones should not. Just be careful though, because relying on the Modification Indices is very similar to using Stepwise regression to build your models. But, sometimes you'll find that a relatively simple and reasonable change will make for a big improvement in your model.

Note, too, that Modification Indices can be used to diagnose any problem in model specification; they aren't limited just to multiple group comparisons.