Multiple-Group Analysis with Missing Data

Sörbom (1981) reanalyzed some data from the Head Start summer program previously analyzed by Magidson (1977). Sörbom used data on 303 white children consisting of a Head Start sample (N = 148) and a matched Control sample (N = 155). The children were matched on gender and kindergarten attendance, but no attempt had been made to match on social status variables.

The variables used in Sörbom's re-analysis were:

X_1 = Mother's education (MOTHEDUC)
X_2 = Father's education (FATHEDUC)
X_3 = Father's occupation (FATHOCCU)
X_4 = Family income (FAMILINC)
Y_1 = Score on the Metropolitan Readiness Test (MRT)
Y_2 = Score on the Illinois Test of Psycholinguistic Abilities (ITPA)

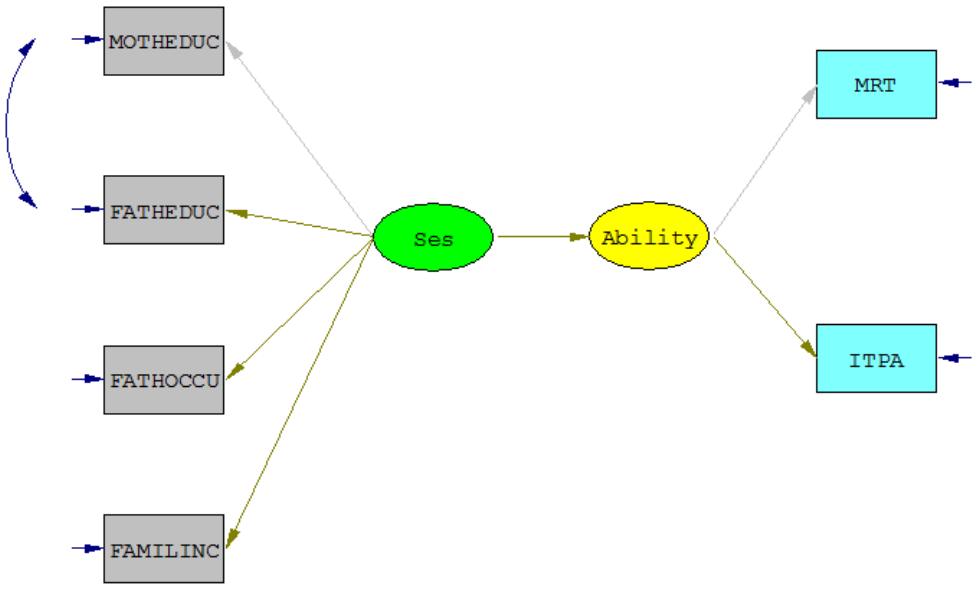
The following issues were examined:

- Test whether X_1, X_2, X_3 and X_4 can be regarded as indicators of a single construct Ses (socioeconomic status) for both groups. Is the measurement model the same for both groups? Is there a difference in the mean of Ses between groups?
- Assuming that Y_1 and Y_2 can be used as indicators of another construct Ability (cognitive ability), test whether the same measurement model applies to both groups. Test the hypothesis of no difference in the mean of Ability between groups.
- Estimate the structural equation:

$$Ability = \alpha + \gamma Ses + z$$

- Is γ the same for the two groups? Test the hypothesis $\alpha = 0$ and interpret the results.

A conceptual path diagram for the model fitted to the data is shown below.



The fitted covariance matrices obtained from **ex16d.spl** (in the **SIMPLIS examples** folder) were used to simulate a control group data set (sample size 550 and percentage missing 15%) and to simulate a Head Start data set (sample size of 600 and percentage missing 10%).

Note:

To invoke the FIML procedure for the analysis of missing data, the following three statements must be given in the SIMPLIS syntax in the order shown below.

```
Missing Value Code <value>
Sample Size: <nsize>
Raw data from file <filename>
```

The complete SIMPLIS syntax file is shown below.

```
FIML: Example 1: SIMPLIS syntax
Group = Control
Observed Variables: MOTHEDUC FATHEDUC FATHOCCU FAMILINC MRT ITPA
Missing Value Code –9
Sample Size: 550
Raw Data from File CONTROL.DAT
Latent Variables: Ses Ability
Relationships:
  MOTHEDUC          = CONST + 1*Ses
  FATHEDUC - FAMILINC = CONST +   Ses
  MRT               = CONST + 1*Ability
  ITPA              = CONST +   Ability
  Ability = Ses
Let the Errors of MOTHEDUC and FATHEDUC correlate
Group = Head Start
Missing Value Code –9
Sample Size: 600
Raw Data from File EXPERIM.DAT

Relationships:
  Ses    = CONST
```

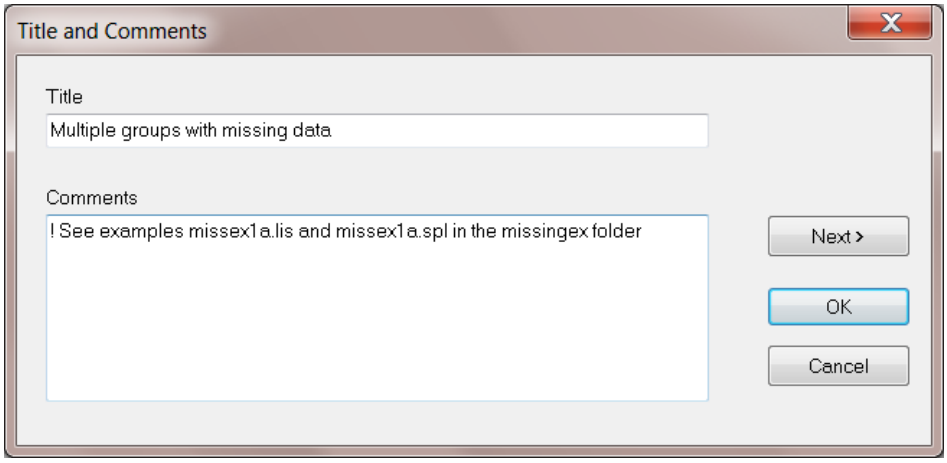
```
   Ability = CONST + Ses
Set the Error Variances of MOTHEDUC - ITPA free
Set the Variance of Ses free
Set the Error Variance of Ability free
Let the Errors of MOTHEDUC and FATHEDUC correlate
LISREL Output: ND=3
Path Diagram
End of Problem
```

If the data are stored in a LISREL system file (***.lsf**), the Observed Variables, Missing Value Code, Sample Size and Raw data from file control.dat can be replaced with the statement

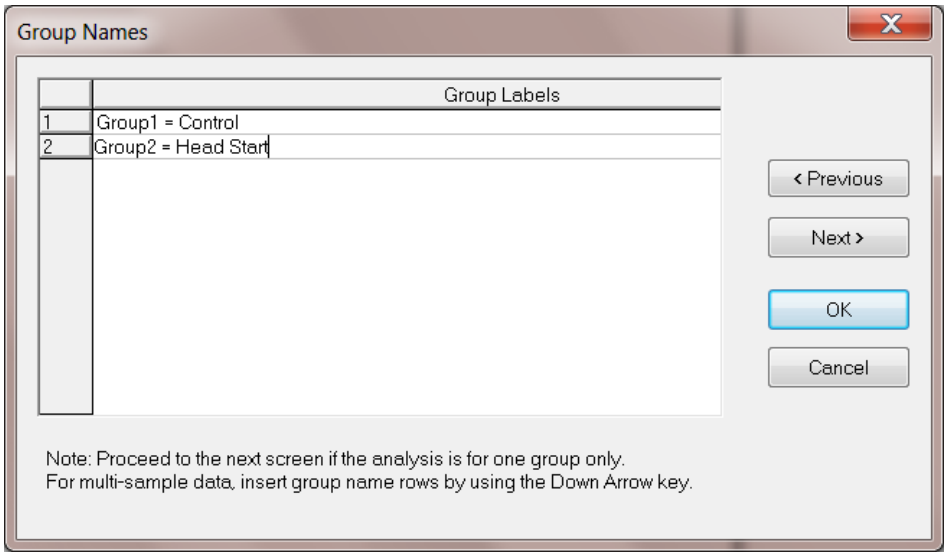
```
Raw data from file control.lsf
```

If a *.**lsf** file contains missing data, the user should ensure that a global missing value code is assigned. This can be done by using the **Data**, **Define Variables** option. Subsequently, we illustrate how to build the SIMPLIS (or LISREL) syntax by drawing a path diagram.

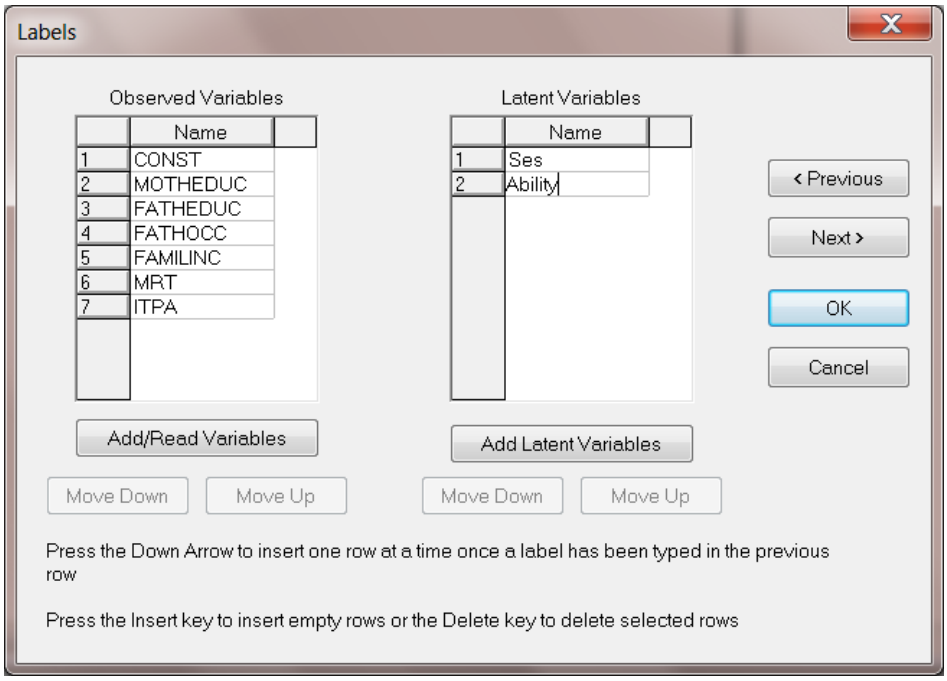
Select the **New** option from the **File** menu. On the **New** dialog box, select **Path Diagram** and save the path diagram as **mgroup.pth**. These steps are described in detail in earlier sections of this guide. From the **Setup** menu, select **Title and Comments** and provide a title and any optional comments. When done, click **Next** to go to the **Group Names** dialog box.



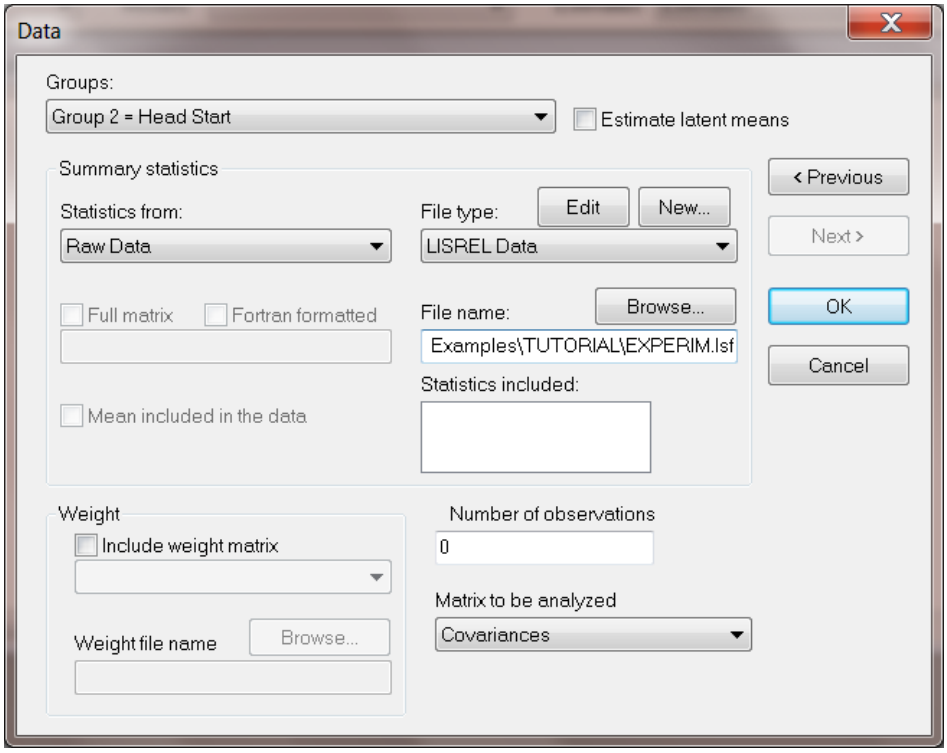
Use the instructions on the bottom of the **Group Names** dialog box to enter group names, then click **Next** to go to the **Labels** dialog box.



The files **control.lsf** and **experim.lsf** contain the Group1 and Group2 data respectively. Click the **Add/Read Variables** button, select LISREL **System File** from the **Add/Read Variables** dialog box and use the **Browse** button to locate the file **control.lsf**. Click **OK** when done to return to the **Labels** dialog box. Use the **Add Latent Variables** button to insert the names Ses and Ability.

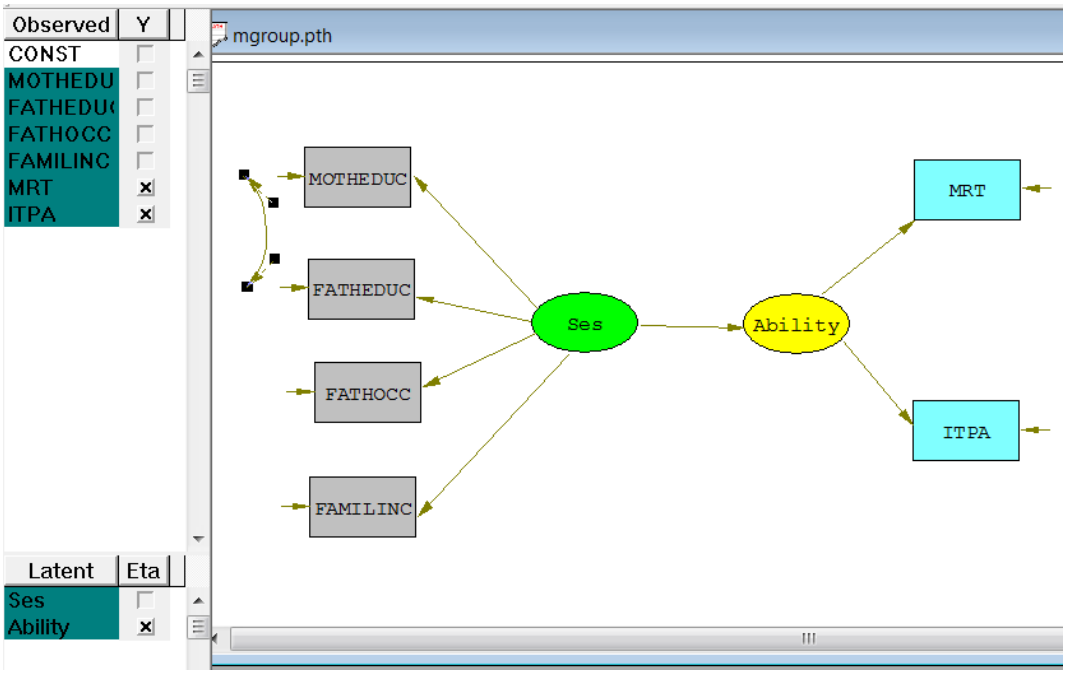


Click the **Next** button to go to the **Data** dialog box. Since a ***.lsf** file was selected in the previous step, the **Data** dialog box shows the **File type:** as LISREL System Data File and the filename as **control.lsf**. Use the **Groups:** drop-down list box to select Group2=Head Start. Once this is done, use the **Browse** button to locate the file **experim.lsf**. Click **Open** to return to the **Data** dialog box.



Click **OK** when done. Select the variables MRT and ITPA as Y-variables and Ability as an Eta-variable. Drag variable names to the path diagram screen in the order MRT, ITPA, Ability, Ses, MOTHEDUC, FATHEDUC, FATHOCC, and FAMILINC. Use the draw toolbar to draw the paths as shown below.

A more detailed description of this process is given in preceding sections of this guide.



When the path diagram is completed, use the **Build SIMPLIS syntax** option from the **Setup** menu to create the SIMPLIS syntax file which is displayed in the SIMPLIS project window, as shown below.

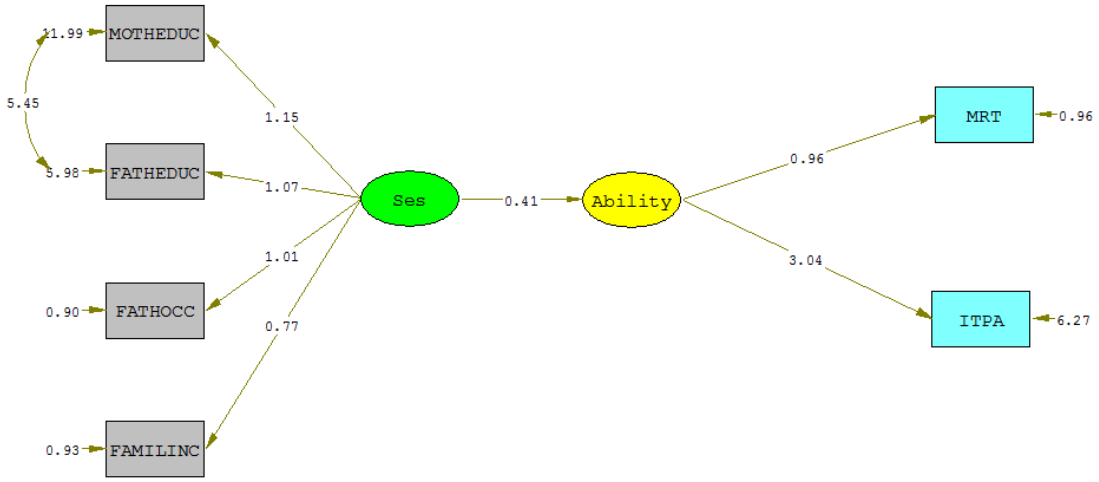
mgroup.SPJ

```
Group 1 = Control
! See examples missex1a.lis and missex1a.spl in the missingex folder
Raw Data from file 'C:\LISREL9 Examples\TUTORIAL\CONTROL.lsf'
Latent Variables  Ability Ses
Relationships
MRT = Ability
ITPA = Ability
MOTHEDUC = Ses
FATHEDUC = Ses
FATHOCC = Ses
FAMILINC = Ses
Ability = Ses
Set the Error Covariance of FATHEDUC and MOTHEDUC Free
Path Diagram

Group 2 = Head Start
Raw Data from file 'C:\LISREL9 Examples\TUTORIAL\EXPERIM.lsf'
Latent Variables  Ability Ses
Relationships
End of Problem
```

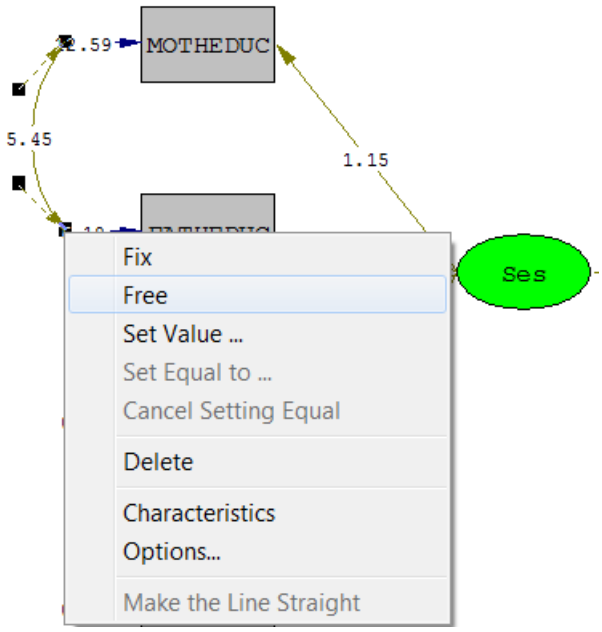
Note that no relationships are given for Group2 under the Relationships keyword. This implies that all parameters are constrained to be equal across groups.

Click the **Run LISREL** icon to create an output file and the path diagram shown below.



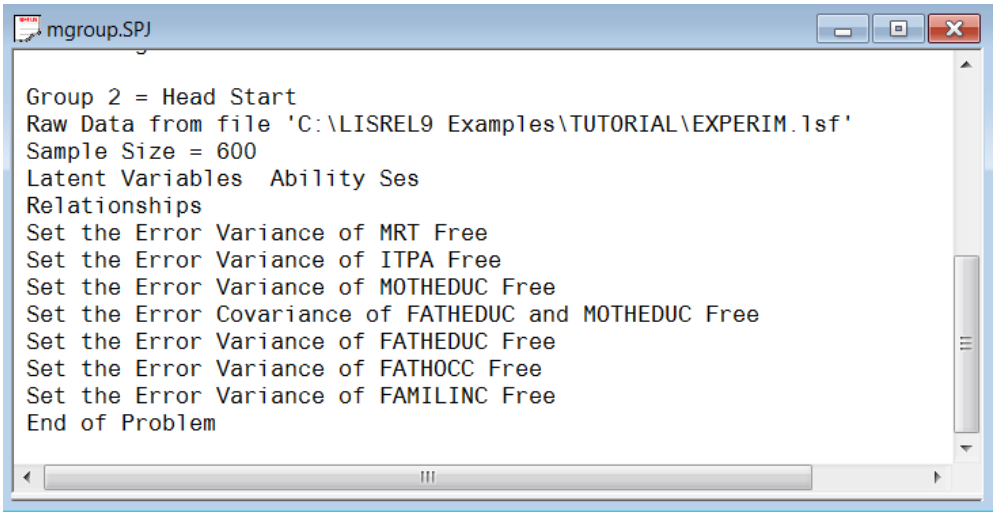
Chi-Square=72.69, df=27, P-value=0.00000, RMSEA=0.054

Since it is not realistic to assume equal error variances across groups, we can set each error variance free by moving the mouse to the appropriate arrow (e.g. MRT ← 0.96). Right click to obtain the menu below and select the **Free** option. Proceed in a similar way to free the error variances of ITPA, MOTHEDUC, FATHEDUC, FATHOCC and FAMILINC as well as the covariance (5.45) between MOTHEDUC and FATHEDUC.

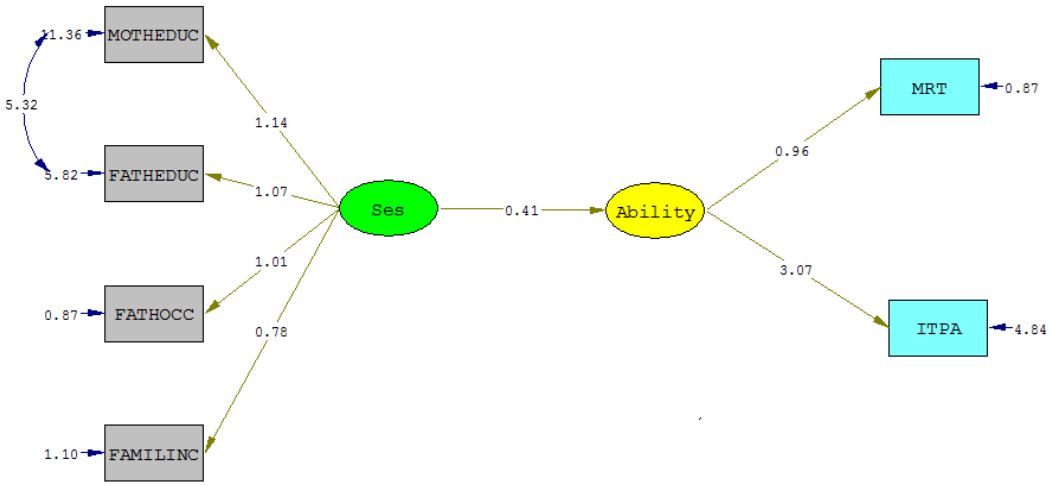


Also right click on the arrow between Ses and Ability to free that path. When this is done, rebuild the SIMPLIS syntax using the **Build SIMPLIS syntax** option on the **File** menu.

A portion of the SIMPLIS project file showing the resultant relationships for group 2 is shown below.



Click the **Run LISREL** icon button to produce the path diagram shown below.



Chi-Square=31.87, df=20, P-value=0.04468, RMSEA=0.032

When the missing value code, sample size and raw data file information are read in, the EM algorithm for estimating the means and covariances under the unrestricted model is started. From this portion of the output a percentage of 14.61% missing cases are reported for group 1 and 9.92% for group 2.

Note that the estimated means and covariances are used to obtain starting values for the FIML procedure. In addition, a $-2\ln L$ value is reported for each group. This value is minus two times the log likelihood value obtained when no restrictions are imposed on means and covariance matrices. From the output it follows that the sum of the $-2\ln L$ values for the groups equals 11358.595 + 12716.065 = 24074.660.

The FIML procedure converged in 6 iterations. Portions of the output file are given below.

(i) LISREL Estimates (Maximum Likelihood) for the control group

Number of Iterations = 6
Measurement Equations

MRT = 0.86*Ability, Errorvar.= 1.06 , R**2 = 0.20
(0.11) (0.084)
7.78 12.58

ITPA = 2.74*Ability, Errorvar.= 7.89 , R**2 = 0.26
(0.69)
11.41

MOTHEDUC = 1.14*Ses, Errorvar.= 12.70, R**2 = 0.093
(0.14) (0.85)
8.42 14.92

FATHEDUC = 1.07*Ses, Errorvar.= 6.15 , R**2 = 0.16
(0.098) (0.43)
10.90 14.26

FATHOCC = 1.01*Ses, Errorvar.= 0.96 , R**2 = 0.51
(0.055) (0.11)
18.35 8.39

FAMILINC = 0.78*Ses, Errorvar.= 0.74 , R**2 = 0.45
(0.047) (0.076)
16.65 9.81

(ii) Global Goodness of Fit Statistics, Missing Data Case

-2ln(L) for the saturated model = 24074.660
-2ln(L) for the fitted model = 24106.532
Degrees of Freedom = 20
Full Information ML Chi-Square = 31.872 (P = 0.0447)
Root Mean Square Error of Approximation (RMSEA) = 0.032

The FIML χ^2 is obtained as the difference between -2 ln L (24105.18) for the fitted model and -2 ln L for the unrestricted model and equals 31.872.