CHAPTER 12 ASDA ANALYSIS EXAMPLES REPLICATION-MPLUS 5.21

GENERAL NOTES ABOUT ANALYSIS EXAMPLES REPLICATION

These examples are intended to provide guidance on how to use the commands/procedures for analysis of complex sample survey data and assume all data management and other preliminary work is done. The relevant syntax for the procedure of interest is shown first along with the associated output for that procedure(s). In some examples, there may be more than one block of syntax and in this case all syntax is first presented followed by the output produced.

In some software packages certain procedures or options are not available but we have made every attempt to demonstrate how to match the output produced by Stata 10+ in the textbook. Check the ASDA website for updates to the various software tools we cover.

NOTES ABOUT ANALYSIS OF LONGITUDINAL DATA IN MPLUS 5.21

The analysis replication examples were all run using Mplus 5.21. Mplus is an advanced modeling tool and offers the ability to correctly account for complex sample survey data for all analytic techniques.

Mplus can perform analysis of longitudinal data using a two-level approach and incorporate level specific weights. In Mplus, the number of levels is assumed to be one more than is used to refer to a model. For example, a two-level model in Mplus is actually performing a three-level analysis. This is due to the way Mplus handles levels in the syntax and setup of the analysis.

The example from Chapter 12 of ASDA uses 4 time points (2000, 2002, 2004, 2006) nested within individuals nested within SECUs, using HRS data from 2000, 2002, 2004 and 2006. The levels, therefore, are time points (level 1) within individuals (level 2) within SECUs (level 3). Through use of "TYPE=TWOLEVEL" along with two variables on the 'CLUSTER' statement we properly account for the multi-level approach as well as the complex sample design. In addition, use of the "WITHIN" weight (level 1 weight) with the "BWEIGHT" (level 2 weight) allows the analyst to use weights for levels 1 (within) and 2 (between). Furthermore, the weights can be scaled as needed. In this example, use of the WTSCALE= ECLUSTER (sum to effective sample size) and BWTSCALE = SAMPLE (product of the within weights and the between weights sum to sample size) options are demonstrated to scale the level weights as needed. This approach is comparable to the Stata gllamm command and provides nearly exact results (slightly different due to variable division/rounding).

Some of the fine points of this approach are use of a unique cluster variable with a different value for each person in the data set, setting any missing data on the weights to zero, dividing the total assets variable expressed in $1000 by 100 to accommodate the Mplus output limits (multiply by 100 to convert to scale presented in ASDA), use of TYPE=COMPLEX TWOLEVEL RANDOM and ESTIMATOR=MLR on the ANALYSIS command, and the declaration of within/between weights. On the MODEL statement, use of the %WITHIN% and %BETWEEN% commands are used to specify the level specific models.

Data preparation in advance of Mplus included extraction of the variables of interest and output to a delimited text file. For additional help and detail, see the Mplus User's Guide.
Title: ANALYSIS EXAMPLE 12.0 LONGITUDINAL DATA EXAMPLE HRS 2000-2006 DATA

Data:
   FILE IS 'F:\brahms\applied_analysis_book\Mplus\testmplus_v3.txt';

Variable:
   NAMES ARE yrssince00 totassets000 newsecu hhidpn baseweight l1weight ;
   Missing are . ;
   weight = l1weight ;
   wtscale= ecluster ;
   bweight = baseweight ;
   bwtscale = sample ;
   usevar = yrssince00 n_tasset ;
   between= ;
   within = yrssince00 ;
   cluster = newsecu hhidpn ;

Define:
   n_tasset = totassets000/100 ;

Analysis:
   type is twolevel random complex ;
   estimator=mlr ;

Model:
   %within%
   n_tasset on yrssince00 ;
   
   %between%
   n_tasset ;
   
*** WARNING
   Variable name contains more than 8 characters.
   Only the first 8 characters will be printed in the output.

   Variable: YRSSINCE00

*** WARNING
Variable name contains more than 8 characters.
Only the first 8 characters will be printed in the output.
Variable: TOTASSETS000

*** WARNING
Variable name contains more than 8 characters.
Only the first 8 characters will be printed in the output.
Variable: BASEWEIGHT

*** WARNING
Data set contains cases with missing on all variables except x-variables. These cases were not included in the analysis.
Number of cases with missing on all variables except x-variables: 124

4 WARNING(S) FOUND IN THE INPUT INSTRUCTIONS

ANALYSIS EXAMPLE 12.0 LONGITUDINAL DATA EXAMPLE HRS 2000-2006 DATA

SUMMARY OF ANALYSIS
Number of groups 1
Number of observations 6840
Number of dependent variables 1
Number of independent variables 1
Number of continuous latent variables 0

Observed dependent variables
Continuous
N_TASSET

Observed independent variables
YRSSINCE

Variables with special functions
Cluster variable NEWSECU HHIDPN
Weight variable (effective cluster-size scaling)
L1WEIGHT
Between weight variable (sample-size scaling)
BASEWEIG
Within variables
YRSSINCE
Estimator MLR

Information matrix OBSERVED

Maximum number of iterations 1000

Convergence criterion 0.100D-05

Maximum number of EM iterations 500

Convergence criteria for the EM algorithm

- Loglikelihood change 0.100D-02
- Relative loglikelihood change 0.100D-05
- Derivative 0.100D-02

Minimum variance 0.100D-03

Maximum number of steepest descent iterations 20

Maximum number of iterations for H1 2000

Convergence criterion for H1 0.100D-03

Optimization algorithm EMA

Input data file(s)

F:\brahms\applied_analysis_book\Mplus\testmplus_v3.txt

Input data format FREE

SUMMARY OF DATA

- Number of missing data patterns 1

COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value 0.100

PROPORTION OF DATA PRESENT

Covariance Coverage

<table>
<thead>
<tr>
<th></th>
<th>N_TASSET</th>
<th>YRSSINCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>N_TASSET</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>YRSSINCE</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

THE MODEL ESTIMATION TERMINATED NORMALLY
TESTS OF MODEL FIT

Chi-Square Test of Model Fit

<table>
<thead>
<tr>
<th>Value</th>
<th>0.002*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degrees of Freedom</td>
<td>0</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.0000</td>
</tr>
<tr>
<td>Scaling Correction Factor</td>
<td>1.000</td>
</tr>
</tbody>
</table>

* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference tests. MLM, MLR and WLSM chi-square difference testing is described in the Mplus Technical Appendices at www.statmodel.com. See chi-square difference testing in the index of the Mplus User's Guide.

Chi-Square Test of Model Fit for the Baseline Model

<table>
<thead>
<tr>
<th>Value</th>
<th>10.538</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degrees of Freedom</td>
<td>1</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

CFI/TLI

<table>
<thead>
<tr>
<th>CFI</th>
<th>1.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLI</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Loglikelihood

<table>
<thead>
<tr>
<th>H0 Value</th>
<th>-31768.011</th>
</tr>
</thead>
<tbody>
<tr>
<td>H0 Scaling Correction Factor</td>
<td>100.064</td>
</tr>
<tr>
<td>for MLR</td>
<td></td>
</tr>
<tr>
<td>H1 Value</td>
<td>-31768.010</td>
</tr>
<tr>
<td>H1 Scaling Correction Factor</td>
<td>100.064</td>
</tr>
<tr>
<td>for MLR</td>
<td></td>
</tr>
</tbody>
</table>
Information Criteria

Number of Free Parameters              4
Akaike (AIC)                   63544.022
Bayesian (BIC)                 63571.344
Sample-Size Adjusted BIC       63558.633
(n* = (n + 2) / 24)

RMSEA (Root Mean Square Error Of Approximation)

Estimate                           0.000

SRMR (Standardized Root Mean Square Residual)

Value for Within                  0.000
Value for Between                  0.000

MODEL RESULTS

Two-Tailed

Within Level

N_TASSET   ON
YRSSINCE00  0.813      0.130      6.258      0.000

Residual Variances

N_TASSET         440.268    146.438      3.007      0.003

Between Level

Means

N_TASSET           7.435      0.790      9.413      0.000

Variances

N_TASSET         351.063    152.408      2.303      0.021

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix              0.449E-01
(ratio of smallest to largest eigenvalue)