Multilevel structural equation modeling

1. Introduction

In practice, many multivariate data sets are obtained by using complex sampling designs. Cluster sampling is a very popular sampling method if the target population has a multilevel (hierarchical) structure such as educational target populations for which students are nested within schools that are again nested in education departments. It is a well-known fact that alternative statistical methods are required to analyze multilevel data sets in an appropriate fashion. Statistical software applications such as MLWin, SAS PROC MIXED, HLM, and the statistical application MULTILEV8 of LISREL 8.80 for Windows (Jöreskog & Sörbom 2006) are used to fit linear and nonlinear models appropriately to hierarchical data sets. However, these models are intended for observed rather than latent variables. Although HLM does allow for latent variables, it requires unrealistic assumptions about the underlying measurement model.

The multilevel SEM algorithm of the statistical application LISREL88 of LISREL 8.80 for Windows allows the user to fit latent variable models to two-level hierarchical multivariate data sets by using the Maximum Likelihood (ML) method for complete multilevel data sets or the Full Information Maximum Likelihood (FIML) method for incomplete multilevel data. More details about this multilevel SEM algorithm are available in Du Toit & Du Toit (2001) while comprehensive technical details are provided in Du Toit & Du Toit (2003).

In this note, the multilevel SEM algorithm of LISREL88 is used to fit a between group Confirmatory Factor Analysis (CFA) model and a within group CFA model to a 2-level data set. This data set is described in the next section. Thereafter, the between and within school CFA models are described. In Section 4, the between and within school models are fitted to the two-level data.

2. The data

The data set forms part of the data library of the Multilevel Project at the University of London, and emanates from the Junior School Project (Mortimore et al, 1988). Mathematics and language tests were administered in three consecutive years to 1192 students from 50 primary schools, which were randomly selected from primary schools maintained by the Inner London Education Authority. The data are provided in the location ***C:\LISREL 8.8 Student Examples\WORKSHOP*** as the file ***JSP1.PSF***. The first portion of this file is shown in the following PSF window.

JSP1.PSF

	school	gender	math1	math2	math3	eng1	eng2
1	1.000	0.000	23.000	24.000	23.000	72.000	80.
2	1.000	1.000	14.000	11.000	-9.000	7.000	17.
3	1.000	1.000	36.000	32.000	39.000	88.000	89.
4	1.000	1.000	24.000	26.000	32.000	12.000	25.
5	1.000	0.000	22.000	23.000	-9.000	67.000	78.
6	1.000	0.000	19.000	23.000	11.000	52.000	76.
7	1.000	1.000	22.000	22.000	26.000	37.000	68.
8	1.000	0.000	18.000	29.000	28.000	57.000	86.
9	1.000	1.000	30.000	31.000	-9.000	42.000	59.
10	1.000	0.000	29.000	29.000	-9.000	46.000	79.
11	1.000	0.000	31.000	28.000	32.000	69.000	84.
12	1.000	0.000	18.000	26.000	-9.000	54.000	74.
13	1.000	0.000	23.000	-9.000	27.000	63.000	-9.
14	1.000	0.000	39.000	35.000	36.000	83.000	88.
15	1.000	0.000	24.000	30.000	33.000	37.000	44.

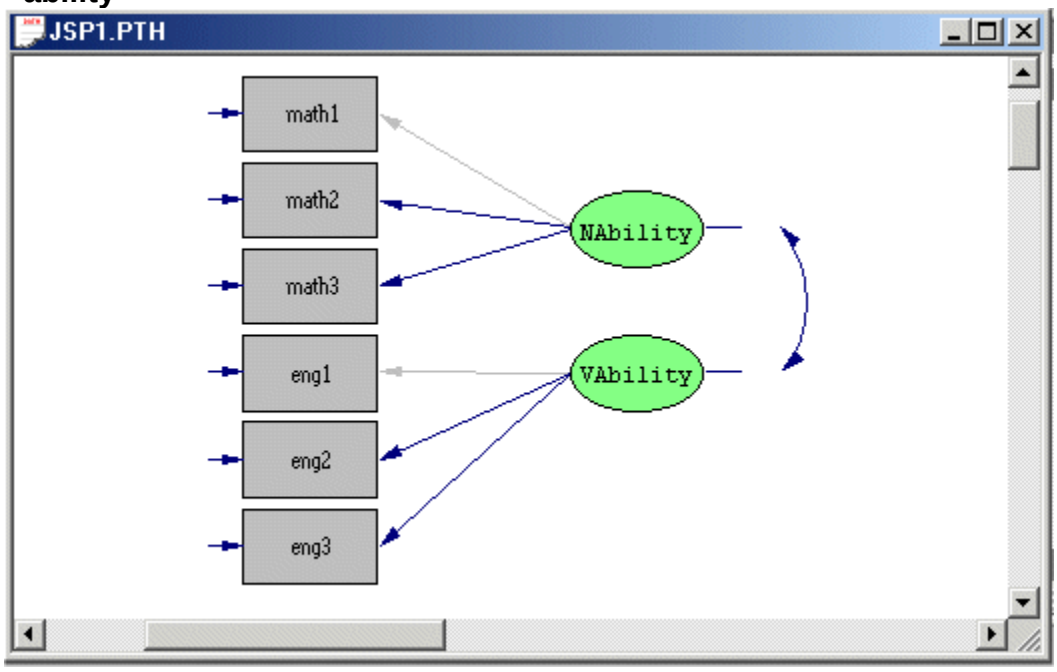
Note that the –9.000 entries represent missing values.

3. The models

The between school model

We consider mathematics and language tests, which were administered in three consecutive years to 1192 students from 50 primary schools. The three mathematics scores are regarded as indicators of the latent variable ***numerical ability*** while the three English scores are regarded as indicators of the latent variable ***verbal ability***. A path diagram for the between school model for numerical and verbal ability is shown in Figure 1.

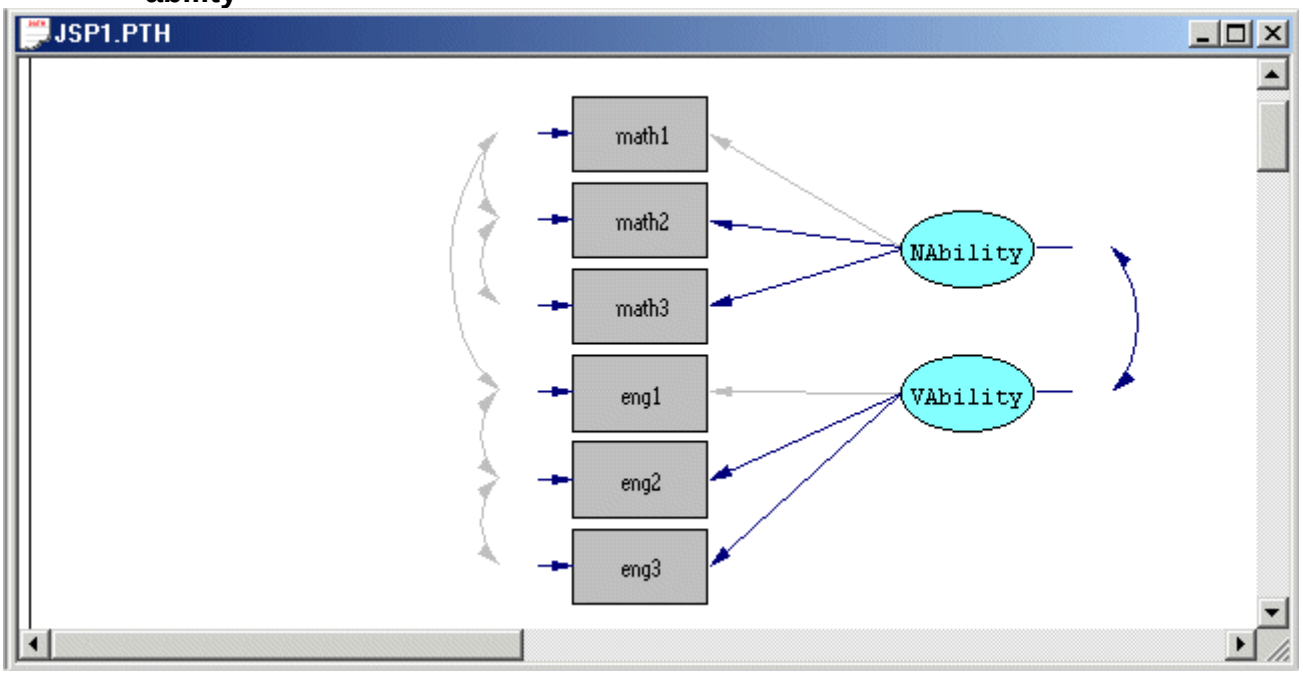
Figure 1: A path diagram for the between school model of numerical and verbal ability



The within school model

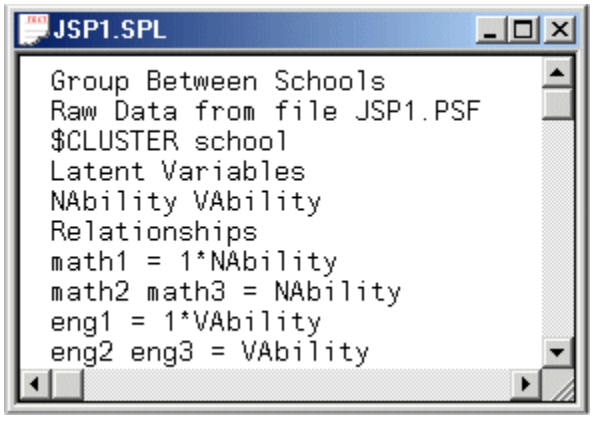
We consider mathematics and language tests, which were administered in three consecutive years to 1192 students from 50 primary schools. The three mathematics scores are regarded as indicators of the latent variable ***numerical ability*** while the three english scores are regarded as indicators of the latent variable ***verbal ability***. In addition, it is assumed that the successive measurement errors for the mathematics and english scores are correlated within schools. It is also hypothesized that the initial errors for the mathematics and english scores are correlated within schools. A path diagram for the within school model for numerical and verbal ability is shown in Figure 2.

Figure 2: A path diagram for the Within School Model of numerical and verbal ability



4. Fitting the between and within school models

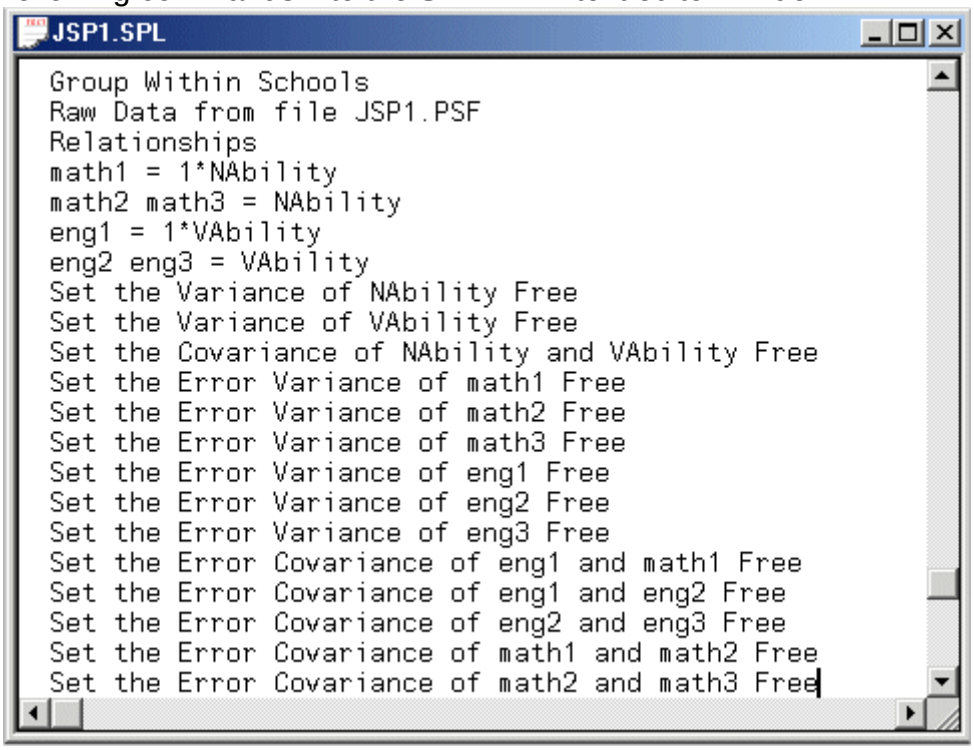
- Use the ***New*** option on the ***File*** menu of the main window to load the ***New*** dialog box.
- Select the ***Syntax Only*** option from the list box on the ***New*** dialog box to open the ***SYNTAX1*** text editor window.
- Enter the following commands into the ***SYNTAX1*** text editor window.



Line 1 specifies that the model that follows is the ***Between School Model***.
Line 2 specifies the raw data source.
Line 3 specifies the cluster variable. This specification invokes the multilevel SEM algorithm of the statistical application LISREL88.
Lines 4-5 specify labels for the latent variables of the model.

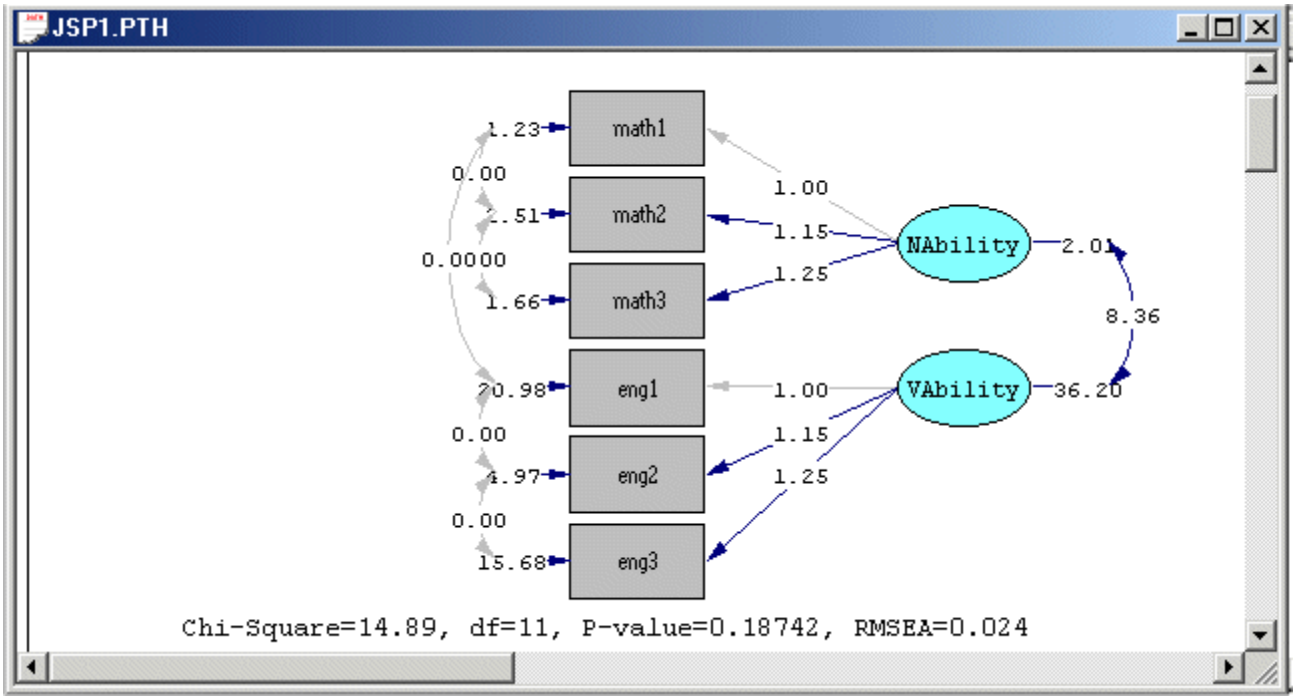
Lines 6-10 specify the model shown in Figure 1.

- Add the following commands into the ***SYNTAX1*** text editor window.



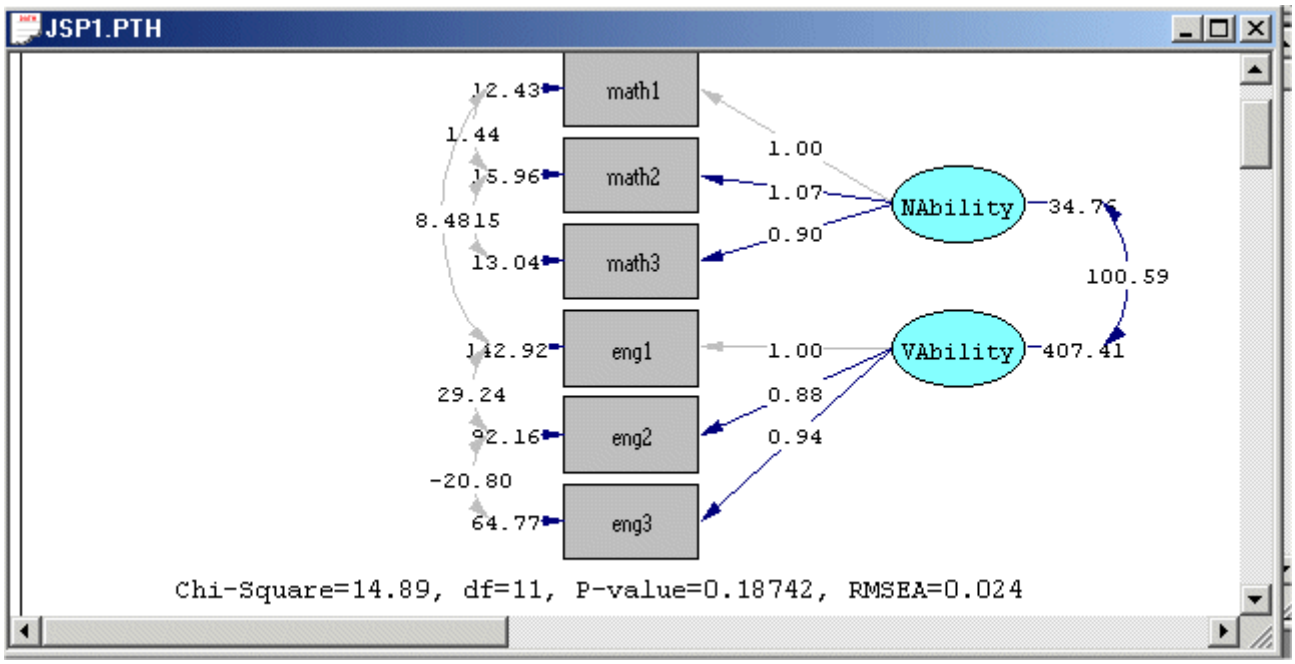
Line 1 specifies that the model that follows is the ***Within School Model***.
Line 2 specifies the raw data source.
Lines 3-21 specify the model shown in Figure 2.
Line 17 specifies that the measurement errors for the initial mathematics and english scores be correlated.
Lines 18-21 specify that the measurement errors for the consecutive mathematics and english scores be correlated.

- Use the ***Save As*** option on the ***File*** menu to load the ***Save As*** dialog box.
- Enter the name ***JSP1.SPL*** in the ***File name*** string field.
- Click on the ***Save*** push button to create the text editor window for ***JSP1.SPL***.
- Click on the ***Run LISREL*** icon to produce the following PTH window.



The results above are those for the Between School Model.

- Select the ***Group Within Schools*** option from the ***Groups*** dropdown list box on the PTH window to display the following PTH window.



References

Du Toit, M. & Du Toit, S.H.C. (2001).
Interactive LISREL: User's Guide.
Lincolnwood, IL: Scientific Software International, Inc.

Du Toit, S.H.C. & Du Toit, M. (2007).
Multilevel Structural Equation Modeling.
In I. Kreft & J. de Leeuw (Eds.): *Multilevel Modeling*, In Press.

Jöreskog, K.G. & Sörbom, D. (2006).
LISREL 8.80 for Windows [Computer Software].
Lincolnwood, IL: Scientific Software International, Inc.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & Ecob, R. (1988).
School Matters: The Junior Years. Wells: Open Books.