**Educational Psychology 960**

**Structural Equation Modeling**

**Problem Set 1: Path Analysis**

For this problem set, you are asked to estimate and test a path analysis model for reading literacy. The data come from the OECD/Programme for International Student Assessment (PISA).

PISA focuses on young people’s ability to use their knowledge and skills to meet real-life challenges. This orientation reflects a change in the goals and objectives of curricula themselves, which are increasingly concerned with what students can do with what they learn at school and not merely with whether they have mastered specific curricular content. PISA’s unique features include its:

• Policy orientation, which highlights differences in performance patterns and identifies features common to high-performing students, schools and education systems by linking data on learning outcomes with data on student characteristics and other key factors that shape learning in and outside of school.

• Innovative concept of “literacy”, which refers both to students’ capacity to apply knowledge and skills in key subject areas and to their ability to analyse, reason and communicate effectively as they pose, interpret and solve problems in a variety of situations.

• Relevance to lifelong learning, which goes beyond assessing students’ competencies in school subjects by asking them to report on their motivation to learn, their beliefs about themselves and their learning strategies.

• Regularity, which enables countries to monitor their progress in meeting key learning objectives.

• Breadth of geographical coverage and collaborative nature, which, in PISA 2009, encompasses the 34 OECD member countries and 41 partner countries and economies.

The data come from the US sample of PISA 2009 (N = 5,233). The data are contained in the file “PISA09.PathAnalysis.csv”. The path diagram and description of the variables is given in the file “PathModel.docx”. For the write-up, please address the following issues.

1. Using summary statistics programs in R, determine if the data meet the assumptions of multivariate normality. Provide evidence to support your conclusion. Note that missing data are coded 9999. For now, please use listwise deletion.

2. Estimate the model using ML. Compute all indirect and total effects in the model. Interpret the results

3. Provide and report on any evidence of the fit (or lack thereof) of the model.

4. Modify the model using the modification indices and expected parameter change statistics. Justify the modification on substantive grounds. (I recognize that this is not your data – make something up that is sensible).

5. Choose a final model based on the AIC.

ESCS

Immigrant status

Gender (female)

MEMOR

CSTRAT

ELAB

Reading

**ESCS**: The PISA index of economic, social and cultural status (ESCS) was derived from the following three indices: highest occupational status of parents (HISEI), highest educational level of parents in years of education according to ISCED (PARED), and home possessions (HOMEPOS). The index of home possessions (HOMEPOS) comprises all items on the indices of WEALTH, CULTPOSS and HEDRES, as well as books in the home recoded into a four-level categorical variable (0-10 books, 11-25 or 26-100 books, 101-200 or 201-500 books, more than 500 books).

The PISA index of economic, social and cultural status (ESCS) was derived from a principal component analysis of standardized variables (each variable has an OECD mean of zero and a standard deviation of one), taking the factor scores for the first principal component as measures of the index of economic, social and cultural status. The final values on the PISA index of economic, social and cultural status (ESCS) have an OECD mean of 0 and a standard deviation of 1.

**Gender**: it is coded to have 1 if female and 0 if male.

**Immig**: The index on immigrant background (IMMIG) is coded to have 1 if either (2) second-generation students (those born in the country of assessment but whose parents were born in another country) or (3) first-generation students (those born outside the country of assessment and whose parents were also born in another country), and to have 0 if native students (those students born in the country of assessment, or those with at least one parent born in that country; students who were born abroad with at least one parent born in the country of assessment are also classified as ‘native’ students).

**Approaches to learning**

How students approach learning is based on student responses in ST27 and measured through the following three indices: memorisation (**MEMOR**), elaboration (**ELAB**) and control strategies (**CSTRAT**). The index of memorisation (**MEMOR**) was derived from the frequency with which students did the following when they were studying: i) try to memorize everything that is covered in the text; ii) try to memorize as many details as possible; iii) read the text so many times that they can recite it; and iv) read the text over and over again. The index of elaboration (**ELAB**) was derived from the frequency with which students did the following when they were studying: i) try to relate new information to prior knowledge acquired in other subjects; ii) figure out how the information might be useful outside school; iii) try to understand the material better by relating it to my own experiences; and iv) figure out how the text information fits in with what happens in real life.

The index of control strategies (**CSTRAT**) was derived from students’ reports on how often they did the following statements: i) when I study, I start by figuring out what exactly I need to learn; ii) when I study, I check if I understand what I have read; iii) when I study, I try to figure out which concepts I still haven’t really understood; iv) when I study, I make sure that I remember the most important points in the text; and v) when I study and I don’t understand something, I look for additional information to clarify this. Higher values on the index indicate higher importance attached to the given strategy.

**Reading**

The final outcome variable is an estimate of reading competency. Higher values indicate higher levels of reading competency.



ESCS

Immigrant status

Gender (female)

MEMOR

CSTRAT

ELAB

Reading