Applied Multivariate Methods
Question 1

This question is based on the data used in the following article:

Hopkins, WE and Hopkins, SA (1997). Strategic planning – financial performance relationships in banks: a causal examination. *Strategic Management Journal*, Vol. 18:8, 635–652.

The paper can be downloaded from the MI455 Moodle page. You should read the article for information on the research questions, data collection and details of the individual variables.

The data described in Table 1 on p645 of the article have been transcribed for your convenience into text files which you can obtain from the MI455 Moodle page. These are:

- hh97_intensity_correlation.txt
- hh97_correlation.txt (or hh97_correlation_14.txt – see explanation below).

Answer the following questions:

1. The paper is based on the importance of the latent variable 'strategic planning intensity' for a model explaining financial performance in banks. Conduct some exploratory factor analyses to assess the adequacy of the measurement model for 'strategic planning intensity'. The file hh97_intensity_correlation.txt contains the correlation matrix for the seven items used to capture this construct. Fit a one-factor model and a two-factor model to these data. Comment on the fit of the models, and on the patterns of factor loadings. For the two-factor model, consider the unrotated solution as well as orthogonal (varimax-rotated) and oblique (promax-rotated) solutions.

2. The final model derived in the paper is depicted in Figure 3 on p647[1]. Fit the model depicted in this figure. You can either use the file hh97_correlation.txt, or the file hh97_correlation_14.txt (see 'Computational hint a)' below).

   Provide the path diagram for the model, as well as the elements of output requested in the overall instructions for this coursework (on p2 of this document). Comment on the size, sign and significance of the parameter estimates, in relation to the expectations from the paper. Comment also on the reliability of the observed items. Finally, comment on the overall fit of the model.

   Computational hints:

   a) The file hh97_correlation.txt is the correlation matrix for all 16 observed variables in the data set. If you are analysing the data using a full registered version of LISREL, you can use this file. However, the student version of LISREL only allows analyses of up to 15 observed variables at once. If you are using this student version, you can use the file hh97_correlation_14.txt instead, which is the correlation matrix for just the 14 variables you need for this analysis – i.e. it excludes COMPX and CHNGE.
   b) Your parameter estimates are unlikely to look the same as those expressed in the paper (in Table 3). Do not worry about this: the technical details of the model that the authors used are not explicitly given.
   c) The default in LISREL is to allow exogenous latent variables to be correlated with each other. To override this, you need to include the command:

Set Covariance of [latent variable] – [latent variable] To 0

d) The authors set the scales of the latent variables by means of reference variables. You may use this or another method of your choice.

3. Would you suggest any modifications to the model you have just fitted? If so, explain them (there is no need, however, to actually fit any new models to answer this question). If not, justify your decision to leave the model as it is.

Question 2

The table below reports the correlation between seven variables for 767 eleventh-grade males.

INTELLNC	1.000						
SIBLINGS	-0.100	1.000					
FATHEEDUC	0.277	-0.152	1.000				
FATHOCCU	0.250	-0.108	0.611	1.000			
GRADES	0.572	-0.105	0.294	0.248	1.000		
EDUCEXP	0.489	-0.213	0.446	0.410	0.597	1.000	
OCCUASP	0.335	-0.153	0.303	0.331	0.478	0.651	1.000

x_1 intelligence (INTELLNC)
x_2 number of siblings (SIBLINGS)
x_3 father's education (FATHEEDUC)
x_4 father's occupation (FATHOCC)
y_1 grades (GRADES)
y_2 educational expectation (EDUCEXP)
y_3 occupational aspiration (OCCUASP)

1. Fit a full recursive model in which educational expectation depends on grades and occupational aspiration depends on grades and educational expectation, each dependent variable (*y*-variables) possibly depending on all explanatory variables (*x*-variables). Comment on the significance of the model parameters and the direction of effects. What are the main conclusions drawn from the analysis?

2. Modify this model in stages by testing each of the following hypotheses sequentially:
   a) GRADES does not depend on FATHEDUC and FATHOCCU
   b) SIBLINGS does not have any direct effect on any of the dependent variables (*y*-variables).
   c) There are no direct effects of any explanatory variables (x-variables) on OCCUASP. These effects are only mediated via GRADES AND EDUCEXP.

3. Based on your selected model estimate the direct, indirect and total effects of any of the *x*-variables on EDUCEXP. Comment on your results.

4. Consider fitting an alternative path model (structural equation model with latent variables) to the correlation matrix above. That alternative model will treat x_3 and x_4 as measurements of a common underlying variable (factor) that could be called for example 'father's influence'.

   Path relationships to be fitted:
   - Variables x_3 and x_4 (*x*-type variables) will be indicators of a latent variable ξ, (exogenous latent variable).
   - The y_1 and y_2 variables are *y*-type variables. The exogenous latent variable ξ will affect y_1 and y_2 variables. In addition variable y_1 will affect y_2.

   Fit a path model that accounts for the relationships given above (note that this model does not employ all variables in the data set). Comment on the parameters estimates obtained and the goodness-of-fit of the fitted model. Integrating the results of this model with the one you selected as your final model in Question 2, suggest an alternative model that could be used to explain the relationships among the full set of seven observed variables. Discuss this model briefly without fitting it.

Computational hints:

a) Begin by creating a text file of the correlation matrix (or downloading it as assign1.txt from the Moodle page); alternatively you can read the correlation matrix inside the SIMPLIS command file using the syntax command:

```
Correlation matrix
   1.000
  -0.100   1.000
   0.277  -0.152  1.000
   0.250  -0.108  0.611   1.000
   0.572  -0.105  0.294   0.248   1.000
   0.489  -0.213  0.446   0.410   0.597   1.000
   0.335  -0.153  0.303   0.331   0.478   0.651   1.000
```

b) To allow for the path between y1 and y2 (y-type variables) and the latent variable ξ as well as the path between y1 and y2 introduce in the path diagram two endogenous latent variables eta1 and eta 2 for y1 and y2 respectively.
Use the following commands for introducing the y-type variables into the model:

```
y1 = eta1
y2 = eta2
Let the Error Variances of  y1 to be zero
Let the Error Variances of  y2 to be zero
```