Assignment 2

SEM 2: Structural Equation Modeling

Please hand in a `.pdf` file containing your report and a `.R` containing your codes or screenshots of every Jasp analysis. The deadline of this assignment is Tuesday May 22 15:00.

Assignment

Part 1

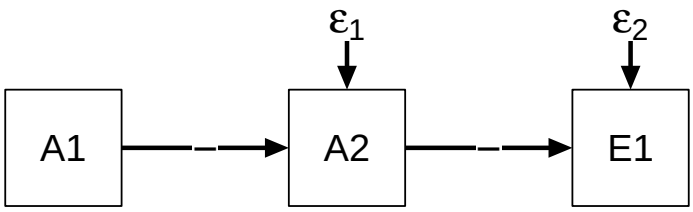
In this assignment we revisit the Big 5 dataset we used in SEM 1, now including variables on age, gender and education:

Item label	Item description
A1	Am indifferent to the feelings of others
A2	Inquire about others' well-being
A3	Know how to comfort others
A4	Love children
A5	Make people feel at ease
C1	Am exacting in my work
C2	Continue until everything is perfect
C3	Do things according to a plan
C4	Do things in a half-way manner
C5	Waste my time
E1	Don't talk a lot
E2	Find it difficult to approach others
E3	Know how to captivate people
E4	Make friends easily
E5	Take charge
N1	Get angry easily
N2	Get irritated easily
N3	Have frequent mood swings
N4	Often feel blue
N5	Panic easily
O1	Am full of ideas
O2	Avoid difficult reading material
O3	Carry the conversation to a higher level
O4	Spend time reflecting on things
O5	Will not probe deeply into a subject
gender	Males = 1 Females = 2
education	1 = HS, 2 = finished HS, 3 = some college, 4 = college graduate 5 = graduate degree
age	Age in years

You can obtain the full dataset (named `bfi`) in R as follows:

```
library("psych")
data("bfi")
```

Of note, these items are *not* from the Big Five Inventory. See `?bfi` for details. For simplicity, treat all variables as continuous. We can think of theories on how these variables interact with one-another. For example, I can reason that being "indifferent to the feelings of others" (A1) causes someone not to "inquire about others' well-being" (A2), and that in turn not inquiring about others well-being causes to "don't talk a lot" (E1).




```
library("lavaan")
Model <- '
E1 ~ A2
A2 ~ A1
'
fit <- sem(Model, bfi, meanstructure = TRUE, missing = "fiml")
fit

## lavaan (0.5-23.1097) converged normally after  21 iterations
##
##   Number of observations                          2800
##
##   Number of missing patterns                         5
##
##   Estimator                                         ML
##   Minimum Function Test Statistic                3.850
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.050

round(fitMeasures(fit)[c("rmsea","cfi","tli","srmr")],2)

## rmsea   cfi   tli  srmr
##  0.03  0.99  0.98  0.01

parameterEstimates(fit)

##   lhs op rhs    est    se       z pvalue ci.lower ci.upper
## 1  E1  ~  A2 -0.291 0.026 -11.246      0   -0.342   -0.240
## 2  A2  ~  A1 -0.284 0.015 -19.045      0   -0.314   -0.255
## 3  E1 ~~  E1  2.544 0.068  37.247      0    2.410    2.678
## 4  A2 ~~  A2  1.214 0.033  37.208      0    1.150    1.278
## 5  A1 ~~  A1  1.981 0.000      NA     NA    1.981    1.981
## 6  E1 ~1      4.373 0.128  34.169      0    4.122    4.623
## 7  A2 ~1      5.489 0.042 131.765      0    5.407    5.570
## 8  A1 ~1      2.413 0.000      NA     NA    2.413    2.413
```

This model fits well!

Question 1 (1 point) What parameter corresponds to $\mathscr{E}(\text{A1})$? ■

Question 2 (2 points) Give $\mathscr{E}(\text{E1})$ (the model implied mean of E1) using the above supplied parameter estimates. Compare your answer to `inspect(fit, "sample")$mean` (the sample means). Your answer might differ slightly due to rounding. ■

Question 3 (1 point) Suppose we intervene on this system and make someone very indifferent to the feelings of others (such that that person responds to item A1 with the maximum of 5). What is the expected result of our intervention, $\mathscr{E}(\text{E1} \mid \text{Do}(\text{A1} = 5))$, assuming the above model is true? Tip: replace the expected value of A1 in the expression you used above with the result of this intervention. ■

Question 4 (1 point) Suppose we intervene on this system and make someone talk a lot (such that that person responds to item E1 with the minimum of 1). What is the expected result of for how indifferent this person is to the feelings of others, $\mathscr{E}(\text{A1} \mid \text{Do}(\text{E1} = 1))$, assuming the above model is true? Tip: think about this first before you calculate anything. ■

Question 5 (2 points) List all possible equivalent models to the model I fitted, using only the same variables (no latent variables). For each model, describe if you think $\mathscr{E}(\text{E1} \mid \text{Do}(\text{A1} = 5))$ and $\mathscr{E}(\text{A1} \mid \text{Do}(\text{E1} = 1))$ will be *higher* or *lower* than in the original model (you do not have to fit the models and compute these values). ■

Question 6 (3 points) Think of causal model using variables from the `bfi` dataset. Use between 4 to 6 variables, and make sure your theorized model is *acyclic*. Draw the path diagram of the model, including the expected sign (- or +) of each regression (you may simplify the path diagram by not drawing exogenous variables as I did above). ■

Question 7 (2 points) Fit your hypothesized model using Lavaan and judge the fit. You do not have to modify the model to improve fit. ■

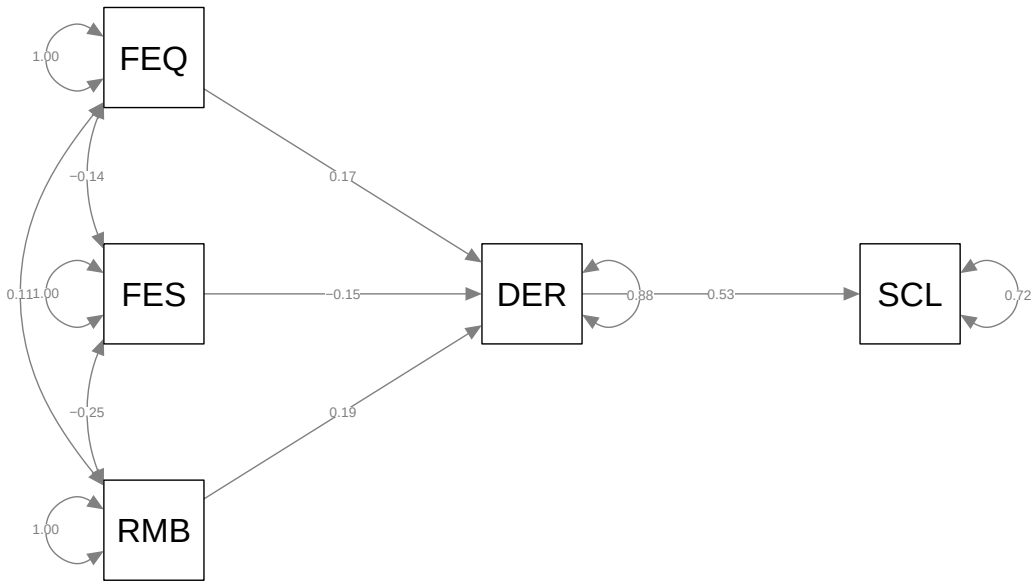
Question 8 (3 points) Is an equivalent model possible using the same variables? Why (not)? If you can identify an equivalent model, confirm the model is equivalent using Lavaan. If not, think of a model that, while different, might fit about as well as your theorized model, and compare model fit of the alternative model to your theorized model. ■

Part 2

Given the following SEM analysis (original article on blackboard):

```
library("semPlot")
library("lavaan")
suveg.r <- c(
'1.00
-0.25 1.00
 0.11 -0.14 1.00
 0.25 -0.22 0.21 1.00
0.18 -0.15 0.19 0.53 1.00')

suveg.r <- getCov(suveg.r, names = c("RMBI", "FES", "FEQN", "DERS", "SCL90ANX"))
sd.suveg <- c(0.33, 0.62, 1.00, 0.54, 0.47)
suveg <- cor2cov(R = suveg.r, sds = sd.suveg)

mod1 <- '

ENTER MODEL HERE

'

fit1 <- sem(mod1, sample.cov = suveg, sample.nobs = 676, fixed.x = FALSE)
```

```
library("semPlot")
semPaths(fit1,"mod","std", layout = "tree2", rotation = 2,
         sizeMan = 10, curve = 2)
```



Question 9 (1 point) Fill in the model to replicate the above analysis ■

Question 10 (2 points) Given the above path diagram, which of these statements are true?

- FES ⫫ SCL
- FES ⫫ SCL | DER
- RMB ⫫ FEQ | DER

■

We can obtain the implied variances and covariances using:

```
lavInspect(fit1, "sigma")
```

```
##           DERS   SCL90A RMBI   FES    FEQN
## DERS       0.291
## SCL90ANX   0.134  0.221
## RMBI       0.044  0.021  0.109
## FES       -0.074 -0.034 -0.051  0.384
## FEQN       0.113  0.052  0.036 -0.087  0.999
```

Question 11 (2 points) Using the Schur complement and the *model implied* variances and covariances, compute the conditional covariance between FES and SCL given DER. Round your result to 3 digits. ■

Question 12 (1 point) Can the arrow DER → SCL be changed into DER ← SCL? Why (not)? ■