Homework 2
Due May 21, 2020, 10AM (pdf preferred)

For all the following problems, please type your answers and provide the relevant printouts. Each answer should be no longer than 2-3 paragraphs in length. You may use Mplus or lavaan for any of the structural models requested below.

The data set for this assignment comes from the same aging and control study used in HW 1 but contains a set of different variables. Download agingcontrol2.dat (for use with Mplus) or agingcontrol.sav (for use with R) from the data page http://web.pdx.edu/~newsomj/data.htm. The variables, which are from the Center for Epidemiologic Studies-Depression (CES-D) measure were rated on a 7-point scale of how often each of the following symptoms applied. Three of the items are positively worded and have been reverse scored (so that higher scores indicate greater depression).

sleep "trouble going/staying sleep?"
effort "felt that everything was an effort?"
getgo "felt could not get going?"
enjoy "enjoyed life? (reversed)"
hopeful "felt hopeful about future (reversed)?"
energy "had lots of energy (reversed)?"

One additional variable, strenex, is the strenuous exercise variable used in HW 1.

1. Conduct a confirmatory factor analysis of the six depression items to test whether a single factor model fits the data (no recoding of the variables should be done in this case). Report the standardized loadings, their significance, the chi-square for the model, the CFI, and the SRMR. Does this model fit the data well? Why or why not? Do the factor loadings suggest there are any poor items? Why or why not?

2. Use the information contained in the printout from the one-factor model to compute Bollen’s IFI by hand (please show your work). Interpret the result and compare to the value of the CFI obtained from Mplus or lavaan.

3. Use Mplus or lavaan to conduct a confirmatory factor analysis that investigates whether the items instead belong to two related factors in which sleep, effort, getgo load on one factor and enjoy, hopeful, and energy load on a second factor. Report the standardized loadings, their significance, the chi-square for the model, the CFI, and the SRMR. Does this model fit the data well? Why or why not?

4. Compare the fit of the one-factor and two-factor models using a chi-square difference test. Report and interpret your findings. Use the Excel spreadsheet available on the class website to obtain Cohen's $w$ and the $\Delta$Mc Centrality Index to gauge the magnitude of effect of the difference.

5. For the two-factor model tested above, show by hand how the degrees of freedom are derived (show your work).

6. Test the two-factor model again (same item assignments as before), but this time set the variance of the latent variables to 1 instead of using a marker (referent) variable for identification of each factor.

---

1 The cases are a random subset of 448 from the full study.
7. For the two-factor model tested above, re-run the model and request modification indices. Report the three largest modification indices for correlated measurement residuals (correlated errors) and their associated standardized expected parameter change values (you do not need to repeat the write-up of the two-factor model results). What do the results suggest about changes that might be made to improve the model?

8. Add one correlated error suggested by the largest modification index and retest the model. Report the standardized loadings, their significance, the chi-square for the model, the CFI, and the SRMR. Does this model fit the data well? Why or why not?

9. a. Use your own data set and test a one-factor confirmatory factor model (using Mplus or lavaan) with at least 6 continuous measured variables (should have 5 or more ordinal values) that could be hypothesized to belong to a single factor (note the Mplus demo version has a limit of six variables). Briefly describe your variables and measure (2-3 sentences at most). Report the standardized loadings, their significance, the chi-square for the model, the CFI, and the SRMR. Does this model fit the data well? Why or why not?

b. Test a two-factor model using the same indicators as in 9a above (even if the two-factor model does not make much sense, assign at three indicators to another factor). Report the standardized loadings, their significance, the chi-square for the model, the CFI, and the SRMR. Does this model fit the data well? Why or why not? Conduct a likelihood ratio test comparing the chi-square value for the two-factor model and the one-factor model. Report and interpret your findings.

c. For whichever model fit the data better, request modification indices and discuss what changes you might make to the model to improve the fit.

10. Use the aging and control data set and the two-factor model tested above (with correlated error added) to investigate the independent (direct) effects of the two latent variables in predicting the strenuous exercise. Report the standardized paths, their significance, the chi-square for the model, the CFI, the SRMR, and the R-squared value for the depression measure.