

Homework #5
Path Analysis

Listed below is a homework exercise for *Path Analysis*. This exercise's main purpose is to help students consolidate the material learned in this section. Use *Mplus* to determine the answers to the following questions. Make sure that you submit to the Instructor *via e-mail* by Tuesday at 4pm: your typed answers to the questions below, as well as your syntax and output.

Please refer to the attached Mplus syntax file, labeled "HW5PathAnalysis.inp". In this example, we do not have access to the raw data. Rather, we have access to a correlation matrix, which we can read in Mplus in order to analyze the model. The syntax file thus has only the required information in order to run the correlation matrix, which is found in the attached "HW5PathAnalysis.dat" file. Make sure these two files are in the same folder, in order for the syntax to run properly.

In this dataset, measures of exercise, hardiness, fitness, stress and illness were administered to 373 university students.

Let us assume that the data do *not* violate the assumptions of normality, independence of observations, homoscedasticity or homogeneity of variance, multicollinearity, singularity or linearity, and have no missing values.

Specify a path analysis model (ML estimation), in which exercise and hardiness are both exogenous variables that are set to covary (i.e., have an unanalyzed association). Exercise is specified to predict fitness, and fitness should in turn to predict illness. Hardiness should predict stress, and stress should in turn predict illness.

- 1) Model specification
 - a. Draw the model. Remember to include in your drawing all observed variables, paths, variances, and disturbances, and fix any parameters that require fixing
- 2) Model identification
 - a. How many observations are in this model, and how did you arrive at this number?
 - b. How many parameters are in this model, and how did you arrive at this number?
 - c. Is this model recursive or non-recursive, and why?
 - d. Is the model identified, and why?
- 3) Analysis of the model
 - a. Which standardized path coefficients are the strongest, and what are these coefficients (provide the numbers)? What do these coefficients mean (how can they be interpreted in terms of association between the predictor and dependent variable)?
 - b. How much variance in fitness is not accounted for by the direct effect of exercise on it?
- 4) Respecification of the model
 - a. If you respecify the model just as above, but also include indirect effect of a) exercise to fitness to illness, and b) hardiness to stress to illness... are these

indirect effects statistically significant (and what are the standardized coefficients)?

(Note: Normally a hypothesized path model would be assessed for how well it fit the data - using goodness of fit indices. However, model fit will not be discussed in detail until the next lecture – on confirmatory factor analysis. This is because not enough time was available during the path analysis lecture to cover model fit).