

# Problem Set #1

## Statistics for Psychosocial Research II

### **Part A: Causality**

Consider the following system of equations:

$$Y_1 = \gamma_{11}X_1 + \varepsilon_1$$

$$Y_2 = \gamma_{22}X_2 + \gamma_{21}X_1 + \beta_{21}Y_1 + \varepsilon_2$$

1. Draw the path diagram associated with the system of equations.
2.
  - a. Which variables are endogenous?
  - b. Which variables are exogenous?
3. Think of a substantive example which is consistent with the above model. Describe each variable and your hypothesis as to the sign of the coefficients in the model.
4. Assume the model above is correctly specified. What if you were to remove  $Y_1$  from the model? What assumption of causal models would you be violating? Describe in WORDS the effect of omitting  $Y_1$  from the model.
5. Extra Credit:
  - a. Calculate the covariance between  $X_1$  and  $\varepsilon_2$  (i.e.  $\varepsilon_2^*$ ) if  $Y_1$  is omitted from the model specified above.
  - b. Calculate the covariance between  $X_2$  and  $\varepsilon_2$  (i.e.  $\varepsilon_2^*$ ) if  $Y_1$  is omitted from the model specified above.
  - c. Interpret the findings from a. and b.

## **Part B: Path Analysis**

This section of the problem set concerns intergenerational mobility– that is, the degree to which the status attainment of a prior generation predicts the status attainment of the children. “Status” is the ranking of the individual along dimensions of the reward and power structure of society. The problem set will ask you to replicate the classic path model from the study by Peter Blau and Otis Dudley Duncan, using data from the Baltimore ECA Followup. The variables pertaining to occupation are percentile rankings, from 0 to 99, with 99 being more prestigious occupations. The Blau-Duncan variables are available in the dataset BIOS658 in the class area. There are six variables of interest for this problem:

*Fathed*: Father’s educational attainment, in years of schooling achieved

*Fathocc*: Father’s occupational attainment, measured in a percentile prestige score

*Grade*: Respondent’s educational attainment in years of schooling

*Firstjob*: Occupational attainment score of the son’s first job

*Jobnow*: Son’s eventual occupational attainment.

*Income*: Income of the household of the respondent, in 22 ordinal categories from less than \$1,000 to \$150,000 or more

1. Create a specification for the process of intergenerational mobility: that is, draw a path diagram which reasonably represents the process. How many parameters need to be estimated? Is the model identified? Include non-causal associations (i.e. correlations as depicted with curved arrows) between variables in your path model if that makes sense to you. In words, compare your specification to a just-identified (saturated) recursive model of these data.
2. Write the structural equations that represent the specification. Use betas to represent the path coefficients and abbreviated names of the variables as above. Use the path formula to show how the correlations relate to the structural equations.
4. Estimate the model. Comment on the fit of the model.
5. Compute one or more indirect effects in your model and comment on their substantive interpretation. If your model did not include indirect effects, respecify and estimate it again.
6. Eliminate any path coefficients which are not statistically different from zero. If there are none in your estimated model, eliminate one of your choice. Re-estimate the model, and comment on any changes in path coefficients in comparison to the original model. Is the fit of the model improved or does it weaken? How much, numerically, and how important is the change in fit?
7. Specify a model in which it cannot be assumed that respondent’s educational attainment preceded his first job, and estimate it. Comment on its interpretability. Compare the fit of this model with your original specification.
8. Respecify your model to one in which son’s first job *predicts to* father’s educational level. What in the model and diagnostics, if anything, tells you that the model is not correctly specified.