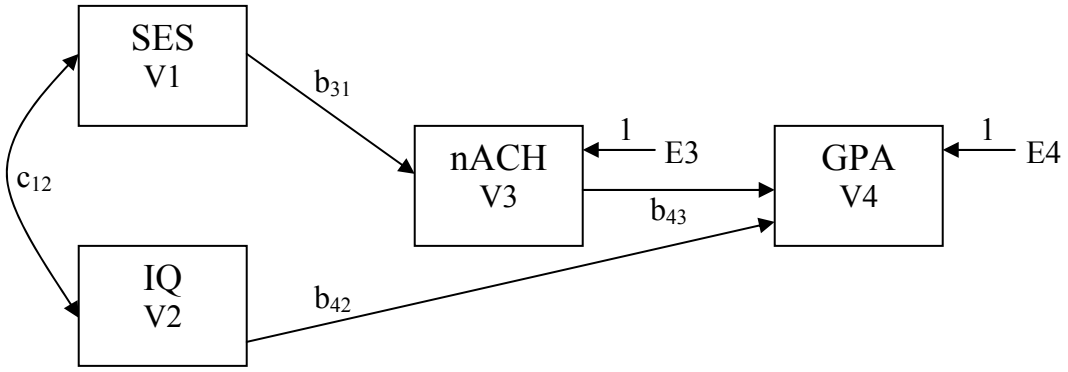
EDMS 722
Assignment 2: Modern measured variable path analysis

Grading: This assignment is worth a total of 41 points.
Format: Although word processed work is preferred, *neatly* handwritten work is acceptable. Show all work and explain answers where required. This assignment requires a computer printout generated outside of class using SIMPLIS.
Due date: This is due by _____. Work should be submitted on time for full earned credit. Late work will be accepted for full earned credit *if and only if* special arrangements are made with me *prior to the due date*; otherwise 5% of the total points possible will be deducted for each day the assignment is late.

Most of the questions on this assignment pertain to the model below, with variables Socioeconomic Status (SES), Intelligence Quotient (IQ), Need for Achievement (nACH), and Grade Point Average (GPA). Note that the variance terms are omitted from the diagram.



1. As defined in class and in the readings, which elements (variables and errors) in the above model are exogenous and which are endogenous? **(2 points)**

2. Write out the structural equations associated with this theoretical model. Use the path labels as they appear in the original diagram above, but use the V variable labels for the variable names. Be sure to include the errors in your equations where appropriate. **(3 points)**

3. In the population there exists a variance/covariance matrix of the form below (where σ_i^2 is the variance of variable V*i* and σ_{ij} is the covariance of variables V*i* and V*j*).

$$\begin{pmatrix} \sigma_1^2 & & & \\ \sigma_{21} & \sigma_2^2 & & \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{pmatrix}$$

If the theory expressed by the path model shown in the diagram is correct, the population variances and covariances of some of the variables can be expressed as functions of the unstandardized path coefficients, variances, and covariances of exogenous elements. Other population variances and covariances are simply functions of themselves, given that they do not depend on anything else within the model (at least according to the theory expressed in the path model).

Recalling the rules given in class for unstandardized path tracing, complete a chart like that below with the appropriate algebraic expressions. Be sure to include all direct effect paths (from variables *and errors*), all indirect effect paths, and all nonstructural associations (spurious and unanalyzed), and the appropriate variance multipliers as dictated by the unstandardized path tracing rules.

First do the covariances. This section decomposes all the possible sources of covariation between the given pair of variables; note that not all sources are necessarily present in each decomposition. **(6 points)**

COVARIANCES

σ_{21}	=	direct effect	+	indirect effect(s)	+	nonstructural relation(s)
σ_{31}	=	direct effect	+	indirect effect(s)	+	nonstructural relation(s)
σ_{32}	=	direct effect	+	indirect effect(s)	+	nonstructural relation(s)
σ_{41}	=	direct effect	+	indirect effect(s)	+	nonstructural relation(s)
σ_{42}	=	direct effect	+	indirect effect(s)	+	nonstructural relation(s)
σ_{43}	=	direct effect	+	indirect effect(s)	+	nonstructural relation(s)

For the variances, use *all* legitimate paths from a variable back to itself. (I offer two important reminders. First, a path from a variable back to itself may follow the same circuit in two different flow directions. Second, multiple backward steps are allowed as long as no path tracing rules are violated.) **(4 points)**

VARIANCES

σ_1^2 =

σ_2^2 =

σ_3^2 =

σ_4^2 =

4. a. According to our class, the unknowns requiring estimation should only be variances of exogenous elements (variables and errors), covariances among exogenous elements appearing in the model, and all direct paths. I bet in some of your decomposition expressions, however, you have a variance term from an *endogenous* variable as part of the right hand side, even though we said that the variance of an exogenous variable is not one of the things for which we need to solve. Because the variance of an exogenous variable is itself decomposed into other things upon which it is dependent, you should substitute your decomposition expression for the exogenous variable's variance in any other decomposition expression where the variance of that exogenous variable appeared. Rewrite only those decomposition expressions from question 3 in which this substitution was needed to be made. **(2 points)**

b. How many different parameters are there that need to be solved for? List them out. They should be only the variances of independent elements, covariances among exogenous elements appearing in the model, and all direct paths. Verify that this is the case. **(2 points)**

5. The expressions from questions 3 and 4 make up the elements in the theoretical model-implied variance/covariance matrix ($\hat{\Sigma}$) – that is, the matrix that should be equal to the population variance/covariance matrix *if* the theory expressed in the path model is correct. Of course, we almost never have the privilege of knowing the numbers in the population matrix. All we know are *sample* variances/covariances. For the above variables, these have been determined from a sample of **n=312** to be as follows:

	V1	**V2**	**V3**	**V4**
V1	2.250			
V2	6.840	231.040		
V3	5.658	22.374	84.640	
V4	0.346	6.065	3.220	0.490

Write out the equations that need to be solved by equating the model-implied variance/covariance expressions with the statistics observed in the sample. **(2 points)**

6. How many "pieces of information" (unique variances and covariances, u) are in the sample variance/covariance matrix? How many population parameters in total (t) need to be estimated? Assuming that the model can be solved, does this model appear to be under-, just-, or over-identified? Explain.

The number of degrees of freedom associated with this model is equal to the number of pieces of information minus the number of parameters to be estimated ($u-t$). How many degrees of freedom should this model have? **(2 points)**

Write a SIMPLIS program to do these analyses (using maximum likelihood estimation, which is the default), and run the program. Use your SIMPLIS output and path diagram to answer the remaining questions, and append the output to the back of this homework.

7. Based on your output, write the unstandardized solutions for each of the population parameters to be estimated: direct paths, variances of exogenous elements (variables and errors), and covariances among exogenous elements. [Note: SIMPLIS reports something called the "Reduced Form Equations" – ignore those.] **(3 points)**

8. By lots of plugging into the theoretical model-implied variance/covariance matrix, the parameter estimates from question 7 can be used to generate the reproduced variance/covariance matrix. In a sudden and thoroughly unexpected fit of kindness, I'll only have you do one element here. Based on your decomposition of the covariance between SES and GPA (question 3 and/or 4), use the appropriate parameter estimates from question 7 to compute the covariance estimated to exist between these two variables based on the theoretical model. Verify your number (within rounding error) against that appearing in the implied variance/covariance matrix in your printout — see the matrix labeled "`Covariance Matrix of Latent Variables`". **(2 points)**

9. a. Using criteria discussed in class, what do you conclude about the overall fit of this model based on its goodness of fit summary information? Explain. **(2 points)**

b. Using the equation given in class, derive the value for the Comparative Fit Index based on the independence model χ^2 ("`Chi-Square for Independence Model`") and the theoretical model χ^2 ("`Minimum Fit Function Chi-Square`") information. **(1 point)**

10. Provide a complete interpretation of the *unstandardized* path from IQ to GPA. **(1 point)**

11. In the end, path models are often reported in their standardized form. Redraw the original path model and label all measured variable paths with their standardized solutions, and include the R^2 values for endogenous variables. You may also cut and paste the LISREL path diagram into your homework, if possible; but you will need to add some things to it (R^2 values, and asterisks as per below).

The z-value for the significance test of each parameter estimate appears in the output; it can also be made to appear in the path diagram by choosing "T-values" in the "Estimates:" box. Attach one asterisk (*) to the standardized path coefficient of any path between variables (not involving errors) that is statistically significant at the .05 level (i.e., having a z-value whose absolute value equals or exceeds 1.96). [Don't worry about putting asterisks by statistically significant error vartiances.] **(5 points)**

12. Provide a complete interpretation of the standardized path from IQ to GPA. **(1 point)**

13. Using a *standardized* path model (assume the paths have symbols r_{12}, p_{31}, p_{42}, and p_{43} instead of c_{12}, b_{31}, b_{42}, and b_{43}, respectively), write out the algebraic decomposition for the direct, indirect, and then total structural relations (but *not* the nonstructural relations) in a table of the form below. Note that you may have some blank/zero entries in this table. **(3 points)**

path	direct effect	indirect effect(s)	total effect
SES, IQ			
SES, nACH			
SES, GPA			
IQ, nACH			
IQ, GPA			
nACH, GPA			