

Sociology 601

Homework #9: due November 19, 2009.

STATA problem using the ECLS-K* – Multivariate regression

Directions: You are interested in parents' school involvement. The variables available for this analysis are as follows: school involvement (*schinv*), respondent relationship (mother, father), SES (*w3sesl*), maternal employment status (*momft*), and marital status (*married*).

School involvement is an index of six items asked of parent respondents, "Since the beginning of this school year, [has the child's mother or father] done the following: Attended an open house or a back-to-school night, attended a meeting of a PTA, PTO, or Parent-Teacher Student Organization, gone to a regularly-scheduled parent-teacher conference with {CHILD}'s teacher or meeting with {CHILD}'s teacher, attended a school or class event, such as a play, sports event, or science fair, volunteered at the school or served on a committee, or participated in fundraising for (CHILD)'s school?" The variable index ranges from zero to six, representing the number of different involvements a parent had with the child's school.

W3SESL is a standardized variable, with a mean of 0 and a s.d. of 1, representing a composite of each child's mother's and father's occupation, income, and education. However, with missing data and the sample selected for these analyses, the mean and s.d. is slightly different.

Mother is coded 1 when the child's mother is the respondent attending events, and is coded 0 when the child's father is the respondent attending events.

*This is child-level data. So, each observation represents one child's characteristics, including responses from his/her parents. For this example, "gender" in this data is the child's gender, while the respondent relationship (mother, father) represents parent gender. Most of the parents in this sample are mothers.

Using STATA when necessary, answer the following questions.

You first decide that parents at private schools are more involved in children's education than parents at public schools. Then, you decide that parents' gender will also affect levels of involvement.

- A) Identify the response and explanatory variables. Describe the variables. Is school involvement (*schinv*) an interval scale variable? Is it normally distributed? What kind of variable is SES (*w3sesl*)? Is it normally distributed?
- B) Plot (using a scattergram) the relationship between *schinv* and *w3sesl*. Show your graph results. What kind of relationship do you see?

- C) You also think that the parents' gender might affect school involvement. What proportion of the sample is mothers (mother)? What is the mean school involvement for mothers, and the mean level of school involvement for fathers?
- D) You decide to test these bivariate statistics using OLS regression. Provide the multiple regression equation for this sample, given these key variables, and then provide the prediction equation using the beta estimates provided by STATA.
- E) What are your findings on the relationship of SES to levels of parent involvement? What about respondent relationship (mother) to parent involvement? Interpret your results.
- F) [Reeve: you should go over the "predict" command if we are going to have this question on there?] Find the predicted average level of school involvement for a student with the median SES score, controlling for respondent relationship; then, find the average level of expected school involvement for a mother, controlling for w3sesl.
- G) [Have we gone over this? – either using standardized beta or T values?] What explanatory variables appear to have the largest effects on schinv?
- H) Do a confidence interval for the effects of SES on school involvement, controlling for parent respondent relationship.
- I) How much variation in levels of school involvement do these two explanatory variables account for?
- J) The following variables are also available in this data:
 MOMAGE: the mother's age in years
 MARRIED: whether the respondent is married (=1) or not married (=0).
 READING: the child's reading test score (reading ability). Again, a standardized variable.
 PRIVATE: the child's school type, where private=1 and private=0 (if the child attends a public school).

Imagine that you believe a spurious or intervening relationship exists. Choose one of the above four variables that might produce a spurious or intervening relationship, defend your decision, and then test the outcome using multiple regression. Your hypothesis doesn't need to be right, but rather logically defended. Show your regression results. What are your findings?

Questions from Agresti and Finlay, Chapter 11

2.

5. a-f, i, j

10. a

13