ADVANCED METHODS DOCTORAL EXAM--FALL 1998

CHOOSE EITHER 1 OR 2

1. Assessing model fit is one of the most discussed issues in the area of structural equation models (SEMs). What are the primary issues in determining model fit in SEMs as discussed in the literature? What, if any, issues of model fit do you feel that this literature neglects? Contrast the model fit issue in SEMs with assessing model fit in a regression model with a continuous dependent variable.

2. Structural Equation Modeling is sometimes described as having several stages: (1) Model Specification, (2) Implied Covariance (Moment) Matrix, (3) Identification, (4) Estimation, (5) Model Fit, and (6) Respecification. The completeness of our knowledge and the procedures that we have for each of these six steps varies. In your opinion, which of these steps is the one that is the most developed? Which is the least developed? Justify your choices.

CHOOSE EITHER 3 OR 4

3. Please describe the main features of the sample design used in most household surveys. Explain why OLS is not appropriate given the data collected in such a survey. Briefly describe some ways that sample design can be taken into account in a linear regression analysis.

4. Briefly define the concept of "internal validity" (as used by Cook and Campbell). Discuss some of the potential threats to internal validity. How do these threats relate to the assumptions underlying OLS?

CHOOSE EITHER 5 OR 6

5. (Selection bias.) There has been much concern in recent years about the dangers of selectivity bias in sociological research. Discuss the following points, citing the relevant literature as appropriate:

   a. What is the nature of the selectivity problem? Under what circumstances can sample selectivity threaten internal and/or external validity? What kind of sample selection does NOT create validity problems?

   b. How can sample selectivity be corrected (or controlled) in empirical research? How successful are methods of correction for sample selection? Are there problems associated with using these methods?
c. How much merit is there in the following charge: "The sample selectivity 'problem' has been greatly overblown. In actual empirical research, it is rare to find convincing instances where selectivity bias is really a problem. Most of the time, the consequences of selectivity are perfectly ignorable." Discuss.

6. (Selection bias.) Table 6 shows regression results obtained from a sample of bilingual Hispanics in the High School and Beyond base year survey in 1980. This question focuses on the results for dependent variables $Z_1$, educational expectations (years of schooling the respondent expects to achieve) and $Z_2$, grade point average. (Note that the table shows standardized coefficients, as well as t-ratios in parentheses.) The researchers point out two unexpected findings. One finding is that Hispanicity ($Y_1$), the degree to which a respondent both knows Spanish and uses the language at home, has a positive effect on school achievement as measured by educational expectations and GPA, while many people would have predicted a negative effect of this variable on school achievement, because respondent with greater Hispanicity would be expected to have more language difficulties in school. The other unexpected finding is that the effects of English proficiency ($Y_2$) on $Z_1$ and $Z_2$ are non-significant. Many people would expect that English proficiency would have a positive effect on school achievement.

A critic of the paper claims that these unexpected results may be due to sample selection bias. The critic argues that these respondents are seniors (in grade 12) in high school, and therefore this sample is selected because of the high dropout rate of Hispanic students. It can be estimated from other sources that as many as 30 percent of Hispanic students "who would normally be in school at the time of the survey" might have dropped out of school and have been excluded from the survey. Since nothing is known about the dropouts, techniques to correct for sample selection bias cannot be used. If you were the author of the study, how would you answer that criticism? In particular, can a model of sample selection explain the positive effect of Hispanicity on school achievement? Can selection bias explain the insignificant effect of English proficiency? Can the same model of sample selection explain both findings simultaneously?

CHOOSE EITHER 7 OR 8

7. John, a sociology graduate student, is interested in studying the circumstances under which a country adopts the democratic form of government. John has event history data (date of change in the form of government and time-varying information that may influence the change of the form of government) over the past 200 years for about 100 countries in the world. Lay out the main steps John has to go through in data analysis. Be sure to discuss the potential pitfalls at each step and the ways these pitfalls can be avoided.

8. The results of logit and probit regression may be interpreted in many ways. Describe the ways the results may be interpreted and discuss the weaknesses and strengths of each way of interpretation.