# **9.** *Quantitative Comparisons for Multivariate Models*

## SUGGESTED COURSE EXTENSIONS

### A. Reviewing

- 1. Find a journal article in your field that presents results of an OLS model with at least one categorical independent variable and at least one continuous independent variable. Use their results and the criteria in chapter 9 of *Writing about Multivariate Analysis, 2nd Edition* to answer the following questions.
  - a. Critique the description of the coefficient on a continuous independent variable in terms of direction, magnitude, statistical significance, and units.
  - b. Critique the description of the coefficient on a categorical independent variable.
  - c. Evaluate whether the authors explained their choice of reference category for that variable, and whether they provided enough substantive and empirical information to justify their choice.
  - d. Rewrite the descriptions of the coefficients to correct any problems you identified in parts a through c of this question.
- 2. Find a journal article that presents results of a logistic regression of a binary dependent variable, with at least one categorical independent variable and at least one continuous independent variable. Use the results to answer the following questions.
  - a. Do they report log-odds or odds ratios? If odds ratios, do they interpret them in terms of multiples of odds or multiples of risk?
  - b. Critique the description of the effect size for a continuous independent variable in terms of direction, magnitude, statistical significance, and units, using the criteria in chapter 9.
  - c. Critique the description of the effect size for a categorical independent variable.
  - d. Rewrite the descriptions to correct any shortcomings you identified in parts b and c.

## **B.** Applying Statistics and Writing

Notes: For the "applying statistics" questions, use variables from your own data or the data sets available with the supplemental online materi-

als to substitute for  $Y_1$ ,  $Y_2$ ,  $X_1$ , DUMMY, and CATEGVAR in the models described below. For example, suppose you want to examine factors that predict income. You might use income in dollars as a continuous dependent variable ( $Y_1$ ), educational attainment in years as a continuous independent variable ( $X_1$ ), gender as a binary independent variable (DUMMY), and residence (urban/suburban/rural) as a multicategory independent variable (CATEGVAR). If you wanted to study factors that predict poverty, you might use poverty status (poor/nonpoor) as a categorical dependent variable ( $Y_2$ ) to estimate logit models with the same set of independent variables.

If possible, choose variables that are part of an ongoing research project. Save the computer output from the models you estimate in questions B.1 through B.3 for use in the exercises for chapters 11, 15, and 16.

- 1. Using data on a continuous dependent variable (denoted  $Y_1$  in the equations below) and a continuous independent variable (denoted  $X_1$  in the equations below),
  - a. Estimate an OLS model of the form  $Y_1 = \beta_0 + \beta_1 X_1$  in the original, untransformed units of both the dependent and independent variables, with unstandardized coefficients.
  - b. Write a sentence interpreting the value of  $\beta_1$ , referring to the specific independent and dependent variables you have used and specifying the units using the guidelines in chapter 9 of *Writing about Multivariate Analysis, 2nd Edition.*
- 2. Using data on the dependent variable used in the preceding question and a binary independent variable (denoted *DUMMY* in the equations below, coded 1 for a specified value and 0 for the reference category),
  - a. Estimate an OLS model of the specification:  $Y_1 = \beta_0 + \beta_1 DUMMY$ .
  - b. Write a sentence interpreting  $\beta_1$ .
  - c. Using the estimated coefficients from part a, calculate predicted values of  $Y_1$  for cases in each category of *DUMMY*. Compare these against the mean value of  $Y_1$  for each of those categories of *DUMMY* from a bivariate calculation.
- 3. Using data on the same variables used in the two preceding questions, estimate an OLS model of the form  $Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 DUMMY$ .
  - a. Write a sentence interpreting the value of  $\beta_1$ , making sure to specify what else was controlled in the model.
  - b. Write a sentence interpreting the value of  $\beta_2$ .
- 4. Using data from your data set on a dichotomous dependent variable  $(Y_2)$ , a continuous independent variable  $(X_1)$ , and a categorical independent variable (*DUMMY*), estimate a logistic regression model of the form logit $(Y_2) = \beta_0 + \beta_1 X_1 + \beta_2 DUMMY$ . See your software manual for instructions on how to specify which category of your

#### 68 CHAPTER NINE

dependent variable to model. Using the guidelines in chapter 9 for writing about odds ratios,

- a. Write a sentence interpreting the value of  $\beta_1$ .
- b. Write a sentence interpreting the value of  $\beta_2$ .
- c. Revising

## C. Revising

- 1. Repeat question A.1 for a results section you have written previously that describes results of an OLS regression.
- 2. Repeat question A.2 for a results section you have written previously that describes results from a logistic regression analysis of a binary dependent variable.