

Chapter 23

45. OECD (introduced in Chapter 19) An analyst at the United Nations is developing a model that describes GDP (gross domestic product per capita, a measure of the overall production in an economy per citizen) among developed countries. She is using national data for 29 countries from the 2005 report of the Organization for Economic Cooperation and Development (OECD). She started with the equation (estimated by least squares)

Estimated per capita GDP = \$26,714 + \$1.441Trade Balance

The trade balance is measured as a percentage of GDP. Exporting countries tend to have large positive trade balances. Importers have negative balances. This equation explains only 37% of the variation in per capita GDP, so she added a second explanatory variable, the number of kilograms of municipal waste per person.

- (a) Examine scatterplots of the response versus the two explanatory variables as well as the scatterplot between the explanatory variables. Do you notice any unusual features in the data? Do the relevant plots appear straight enough for multiple regression?
- (b) Do you think, before fitting the multiple regression, that the partial slope for trade balance will be the same as in the equation shown? Explain.
- (c) Fit the multiple regression that expands the one-predictor equation by adding the second explanatory variable to the model. Summarize the estimates obtained for the fitted model.
- (d) Does the estimated model appear to meet the conditions for the use of the MRM?
- (e) Draw the path diagram for this estimated model. Use it to explain why the estimated slope for the trade balance has become smaller than in the simple regression shown.
- (f) Give a confidence interval, to presentation precision, for the slope of the municipal waste variable. Does this interval imply that countries can increase their GDP by encouraging residents to produce more municipal waste?

47. Promotion (introduced in Chapter 19) These data describe promotional spending by a pharmaceutical company for a cholesterol-lowering drug. The data cover 39 consecutive weeks and isolate the area around Boston. The variables in this collection are shares. Marketing research often describes the level of promotion in terms of voice. In place of the level of spending, *voice* is the share of advertising devoted to a specific product.

The column *Market Share* is sales of this product divided by total sales for such drugs in the Boston area. The column *Detail Voice* is the ratio of detailing for this drug to the amount of detailing for all cholesterol-lowering drugs in Boston. Detailing counts the number of promotional visits made by representatives of a pharmaceutical company to doctors' offices. Similarly, *Sample Voice* is the share of samples in this market that are from this manufacturer. Formulate the MRM with y given by the Market Share and x given by Detail Voice and Sample Voice.

- (a) Examine scatterplots of the response versus the two explanatory variables as well as the scatterplot between the explanatory variables. Do you notice any unusual features in the data? Do the relevant plots appear straight enough for multiple regression?
- (b) Fit the indicated multiple regression and show a summary of the estimated features of the model.
- (c) Does the estimated model appear to meet the conditions for the use of the MRM?
- (d) Does this estimated model explain statistically significant variation in the market share?
- (e) At a fixed level of sampling, do periods with increased detailing show significantly larger market share?
- (f) Does the fit of the multiple regression imply that the pharmaceutical company should no longer invest in detailing and only rely on sampling? Discuss briefly.