the annual cost of attending law school, and rank is a law school ranking (with rank = 1 being the best).

(i) Explain why we expect $\beta_4 \leq 0$.
(ii) What signs do you expect for the other slope parameters? Justify your answers.
(iii) Using the data in LAWSCH85.RAW, the estimated equation is

$$\log(salary) = 8.34 + .0047 LSAT + .248 GPA + .095 \log(libvol)$$

$$+ .038 \log(cost) - .033 \text{rank}$$

$$n = 136, R^2 = .842.$$  

What is the predicted ceteris paribus difference in salary for schools with a median GPA different by one point? (Report your answer as a percentage.)

(iv) Interpret the coefficient on the variable $\log(libvol)$.
(v) Would you say it is better to attend a higher ranked law school? How much is a difference in ranking of 20 worth in terms of predicted starting salary?

3.5 In a study relating college grade point average to time spent in various activities, you distribute a survey to several students. The students are asked how many hours they spend each week in four activities: studying, sleeping, working, and leisure. Any activity is put into one of the four categories, so that for each student, the sum of hours in the four activities must be 168.

(i) In the model

$$GPA = \beta_0 + \beta_1 \text{study} + \beta_2 \text{sleep} + \beta_3 \text{work} + \beta_4 \text{leisure} + u,$$

does it make sense to hold sleep, work, and leisure fixed, while changing study?

(ii) Explain why this model violates Assumption MLR.3.

(iii) How could you reformulate the model so that its parameters have a useful interpretation and it satisfies Assumption MLR.3?

3.6 Consider the multiple regression model containing three independent variables, under Assumptions MLR.1 through MLR.4:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u.$$

You are interested in estimating the sum of the parameters on $x_1$ and $x_2$; call this $\theta_1 = \beta_1 + \beta_2$.

(i) Show that $\hat{\theta}_1 = \hat{\beta}_1 + \hat{\beta}_2$ is an unbiased estimator of $\theta_1$.

(ii) Find $\text{Var}(\hat{\theta}_1)$ in terms of $\text{Var}(\hat{\beta}_1)$, $\text{Var}(\hat{\beta}_2)$, and $\text{Corr}(\hat{\beta}_1, \hat{\beta}_2)$.

3.7 Which of the following can cause OLS estimators to be biased?

(i) Heteroskedasticity.

(ii) Omitting an important variable.

(iii) A sample correlation coefficient of .95 between two independent variables both included in the model.
3.8 Suppose that average worker productivity at manufacturing firms (avgprod) depends on two factors, average hours of training (avgtrain) and average worker ability (avgabil):

\[ \text{avgprod} = \beta_0 + \beta_1 \text{avgtrain} + \beta_2 \text{avgabil} + u. \]

Assume that this equation satisfies the Gauss-Markov assumptions. If grants have been given to firms whose workers have less than average ability, so that avgtrain and avgabil are negatively correlated, what is the likely bias in \( \hat{\beta}_2 \) obtained from the simple regression of avgprod on avgtrain?

3.9 The following equation describes the median housing price in a community in terms of amount of pollution (nox for nitrous oxide) and the average number of rooms in houses in the community (rooms):

\[ \log(\text{price}) = \beta_0 + \beta_1 \log(\text{nox}) + \beta_2 \text{rooms} + u. \]

(i) What are the probable signs of \( \beta_1 \) and \( \beta_2 \)? What is the interpretation of \( \beta_2 \)? Explain.

(ii) Why might nox [or more precisely, log(nox)] and rooms be negatively correlated? If this is the case, does the simple regression of log(price) on log(nox) produce an upward or a downward biased estimator of \( \beta_2 \)?

(iii) Using the data in HPRICE2.RAW, the following equations were estimated:

\[ \log(\text{price}) = 11.71 - 1.043 \log(\text{nox}), \quad n = 506, \quad R^2 = .264. \]

\[ \log(\text{price}) = 9.23 - .718 \log(\text{nox}) + .306 \text{rooms}, \quad n = 506, \quad R^2 = .514. \]

Is the relationship between the simple and multiple regression estimates of the elasticity of price with respect to nox what you would have predicted, given your answer in part (ii)? Does this mean that \(-.718\) is definitely closer to the true elasticity than \(-1.043\)?

3.10 Suppose that you are interested in estimating the ceteris paribus relationship between \( y \) and \( x_1 \). For this purpose, you can collect data on two control variables, \( x_2 \) and \( x_3 \). (For concreteness, you might think of \( y \) as final exam score, \( x_1 \) as class attendance, \( x_2 \) as GPA up through the previous semester, and \( x_3 \) as SAT or ACT score.) Let \( \hat{\beta}_1 \) be the simple regression estimate from \( y \) on \( x_1 \) and let \( \hat{\beta}_1 \) be the multiple regression estimate from \( y \) on \( x_1, x_2, x_3 \).

(i) If \( x_1 \) is highly correlated with \( x_2 \) and \( x_3 \) in the sample, and \( x_2 \) and \( x_3 \) have large partial effects on \( y \), would you expect \( \beta_1 \) and \( \hat{\beta}_1 \) to be similar or very different? Explain.

(ii) If \( x_1 \) is almost uncorrelated with \( x_2 \) and \( x_3 \), but \( x_2 \) and \( x_3 \) are highly correlated, will \( \hat{\beta}_1 \) and \( \hat{\beta}_1 \) tend to be similar or very different? Explain.

(iii) If \( x_1 \) is highly correlated with \( x_2 \) and \( x_3 \), and \( x_2 \) and \( x_3 \) have small partial effects on \( y \), would you expect \( \text{se}(\hat{\beta}_1) \) or \( \text{se}(\hat{\beta}_1) \) to be smaller? Explain.

(iv) If \( x_1 \) is almost uncorrelated with \( x_2 \) and \( x_3 \), \( x_2 \) and \( x_3 \) have large partial effects on \( y \), and \( x_2 \) and \( x_3 \) are highly correlated, would you expect \( \text{se}(\hat{\beta}_1) \) or \( \text{se}(\hat{\beta}_1) \) to be smaller? Explain.
3.11 Suppose that the population model determining \( y \) is
\[
y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u,
\]
and this model satisfies Assumptions MLR.1 through MLR.4. However, we estimate the model that omits \( x_3 \). Let \( \hat{\beta}_0, \hat{\beta}_1, \) and \( \hat{\beta}_2 \) be the OLS estimators from the regression of \( y \) on \( x_1 \) and \( x_2 \). Show that the expected value of \( \hat{\beta}_1 \) (given the values of the independent variables in the sample) is
\[
E(\hat{\beta}_1) = \beta_1 + \frac{\sum_{i=1}^{n} \hat{r}_{i1}x_{1i}}{\sum_{i=1}^{n} \hat{r}_{i1}^2},
\]
where the \( \hat{r}_{i1} \) are the OLS residuals from the regression of \( x_1 \) on \( x_2 \). [Hint: The formula for \( \beta_1 \) comes from equation (3.22). Plug \( y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i \) into this equation. After some algebra, take the expectation treating \( x_{3i} \) and \( \hat{r}_{i1} \) as nonrandom.]

3.12 The following equation represents the effects of tax revenue mix on subsequent employment growth for the population of counties in the United States:
\[
growth = \beta_0 + \beta_1 \text{share}_p + \beta_2 \text{share}_t + \beta_3 \text{share}_s + \text{other factors},
\]
where \( growth \) is the percentage change in employment from 1980 to 1990, \( \text{share}_p \) is the share of property taxes in total tax revenue, \( \text{share}_t \) is the share of income tax revenues, and \( \text{share}_s \) is the share of sales tax revenues. All of these variables are measured in 1980. The omitted \( \text{share}_p \), includes fees and miscellaneous taxes. By definition, the four shares add up to one. Other factors would include expenditures on education, infrastructure, and so on (all measured in 1980).

(i) Why must we omit one of the tax share variables from the equation?
(ii) Give a careful interpretation of \( \beta_1 \).

3.13 (i) Consider the simple regression model \( y = \beta_0 + \beta_1 x + u \) under the first four Gauss-Markov assumptions. For some function \( g(x) \), for example \( g(x) = x^2 \) or \( g(x) = \log(1 + x^2) \), define \( z_i = g(x_i) \). Define a slope estimator as
\[
\hat{\beta}_1 = \left( \sum_{i=1}^{n} (z_i - \bar{z})x_i \right) / \left( \sum_{i=1}^{n} (z_i - \bar{z})^2 \right).
\]
Show that \( \hat{\beta}_1 \) is linear and unbiased. Remember, because \( E(u|x) = 0 \), you can treat both \( x \) and \( z \) as nonrandom in your derivation.

(ii) Add the homoskedasticity assumption, MLR.5. Show that
\[
\text{Var}(\hat{\beta}_1) = \sigma^2 \left( \sum_{i=1}^{n} (z_i - \bar{z})^2 \right) / \left( \sum_{i=1}^{n} (z_i - \bar{z})^2 \right)^2.
\]

(iii) Show directly that, under the Gauss-Markov assumptions, \( \text{Var}(\hat{\beta}_1) \leq \text{Var}(\beta_1) \), where \( \hat{\beta}_1 \) is the OLS estimator. [Hint: The Cauchy-Schwartz inequality in Appendix B implies that]
\[
\left( n^{-1} \sum_{i=1}^{n} (z_i - \bar{z})(x_i - \bar{x}) \right)^2 \leq \left( n^{-1} \sum_{i=1}^{n} (z_i - \bar{z})^2 \right) \left( n^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right);
\]

notice that we can drop \( \bar{x} \) from the sample covariance.

---

**COMPUTER EXERCISES**

**C3.1** A problem of interest to health officials (and others) is to determine the effects of smoking during pregnancy on infant health. One measure of infant health is birth weight; a birth weight that is too low can put an infant at risk for contracting various illnesses. Since factors other than cigarette smoking that affect birth weight are likely to be correlated with smoking, we should take those factors into account. For example, higher income generally results in access to better prenatal care, as well as better nutrition for the mother. An equation that recognizes this is

\[
bwght = \beta_0 + \beta_1 cigs + \beta_2 faminc + u.
\]

(i) What is the most likely sign for \( \beta_2 \)?

(ii) Do you think \( cigs \) and \( faminc \) are likely to be correlated? Explain why the correlation might be positive or negative.

(iii) Now, estimate the equation with and without \( faminc \), using the data in BWGHT.RAW. Report the results in equation form, including the sample size and R-squared. Discuss your results, focusing on whether adding \( faminc \) substantially changes the estimated effect of \( cigs \) on \( bwght \).

**C3.2** Use the data in HPRICE1.RAW to estimate the model

\[
price = \beta_0 + \beta_1 sqrt + \beta_2 bdrms + u,
\]

where \( price \) is the house price measured in thousands of dollars.

(i) Write out the results in equation form.

(ii) What is the estimated increase in price for a house with one more bedroom, holding square footage constant?

(iii) What is the estimated increase in price for a house with an additional bedroom that is 140 square feet in size? Compare this to your answer in part (ii).

(iv) What percentage of the variation in price is explained by square footage and number of bedrooms?

(v) The first house in the sample has \( sqrt = 2,438 \) and \( bdrms = 4 \). Find the predicted selling price for this house from the OLS regression line.

(vi) The actual selling price of the first house in the sample was $300,000 (so \( price = 300 \)). Find the residual for this house. Does it suggest that the buyer underpaid or overpaid for the house?

**C3.3** The file CEOSAL2.RAW contains data on 177 chief executive officers and can be used to examine the effects of firm performance on CEO salary.
(i) Estimate a model relating annual salary to firm sales and market value. Make the model of the constant elasticity variety for both independent variables. Write the results out in equation form.

(ii) Add profits to the model from part (i). Why can this variable not be included in logarithmic form? Would you say that these firm performance variables explain most of the variation in CEO salaries?

(iii) Add the variable ceoten to the model in part (ii). What is the estimated percentage return for another year of CEO tenure, holding other factors fixed?

(iv) Find the sample correlation coefficient between the variables log(mkval) and profits. Are these variables highly correlated? What does this say about the OLS estimators?

C3.4 Use the data in ATTEND.RAW for this exercise.

(i) Obtain the minimum, maximum, and average values for the variables andrte, priGPA, and ACT.

(ii) Estimate the model

\[ andrte = \beta_0 + \beta_1 \text{priGPA} + \beta_2 \text{ACT} + u, \]

and write the results in equation form. Interpret the intercept. Does it have a useful meaning?

(iii) Discuss the estimated slope coefficients. Are there any surprises?

(iv) What is the predicted andrte if priGPA = 3.65 and ACT = 20? What do you make of this result? Are there any students in the sample with these values of the explanatory variables?

(v) If Student A has priGPA = 3.1 and ACT = 21 and Student B has priGPA = 2.1 and ACT = 26, what is the predicted difference in their attendance rates?

C3.5 Confirm the partialling out interpretation of the OLS estimates by explicitly doing the partialling out for Example 3.2. This first requires regressing educ on exper and tenure and saving the residuals, \( \hat{r}_1 \). Then, regress log(wage) on \( \hat{r}_1 \). Compare the coefficient on \( \hat{r}_1 \) with the coefficient on educ in the regression of log(wage) on educ, exper, and tenure.

C3.6 Use the data set in WAGE2.RAW for this problem. As usual, be sure all of the following regressions contain an intercept.

(i) Run a simple regression of IQ on educ to obtain the slope coefficient, say, \( \delta_1 \).

(ii) Run the simple regression of log(wage) on educ, and obtain the slope coefficient, \( \beta_1 \). 

(iii) Run the multiple regression of log(wage) on educ and IQ, and obtain the slope coefficients, \( \beta_1 \) and \( \beta_2 \), respectively.

(iv) Verify that \( \beta_1 = \beta_1 + \beta_2 \delta_1 \).

C3.7 Use the data in MEAP93.RAW to answer this question.

(i) Estimate the model

\[ \text{math10} = \beta_0 + \beta_1 \log(\text{expend}) + \beta_2 \text{inchprg} + u, \]
and report the equation in the usual form, including the sample size and $R$-squared. Are the signs of the slope coefficients what you expected? Explain.

(ii) What do you make of the intercept you estimated in part (i)? In particular, does it make sense to set the two explanatory variables to zero? [Hint: Recall that $\log(1) = 0$.]

(iii) Now run the simple regression of $math10$ on $\log(expend)$, and compare the slope coefficient with the estimate obtained in part (i). Is the estimated spending effect now larger or smaller than in part (i)?

(iv) Find the correlation between $expend = \log(expend)$ and $lnchprg$. Does its sign make sense to you?

(v) Use part (iv) to explain your findings in part (iii).

C3.8 Use the data in DISCRIM.RAW to answer this question. These are zip-code-level data on prices for various items at fast-food restaurants, along with characteristics of the zip code population, in New Jersey and Pennsylvania. The idea is to see whether fast-food restaurants charge higher prices in areas with a larger concentration of blacks.

(i) Find the average values of $prpbck$ and $income$ in the sample, along with their standard deviations. What are the units of measurement of $prpbck$ and $income$?

(ii) Consider a model to explain the price of soda, $psoda$, in terms of the proportion of the population that is black and median income:

$$psoda = \beta_0 + \beta_1prpbck + \beta_2income + u.$$  

Estimate this model by OLS and report the results in equation form, including the sample size and $R$-squared. (Do not use scientific notation when reporting the estimates.) Interpret the coefficient on $prpbck$. Do you think it is economically large?

(iii) Compare the estimate from part (ii) with the simple regression estimate from $psoda$ on $prpbck$. Is the discrimination effect larger or smaller when you control for income?

(iv) A model with a constant price elasticity with respect to income may be more appropriate. Report estimates of the model

$$\log(psoda) = \beta_0 + \beta_1prpbck + \beta_2income + u.$$  

If $prpbck$ increases by .20 (20 percentage points), what is the estimated percentage change in $psoda$? [Hint: The answer is 2.xx, where you fill in the "xx." ]

(v) Now add the variable $prppov$ to the regression in part (iv). What happens to $\beta_{prpbck}$?

(vi) Find the correlation between $\log(income)$ and $prppov$. Is it roughly what you expected?

(vii) Evaluate the following statement: "Because $\log(income)$ and $prppov$ are so highly correlated, they have no business being in the same regression."