(ii) Is exper² statistically significant at the 1% level?
(iii) Using the approximation

\[ \% \text{Δwage} = 100(\hat{β}_2 + 2\hat{β}_3 \cdot \text{exper}) \cdot \text{Δexper}, \]

find the approximate return to the fifth year of experience. What is the approximate return to the twentieth year of experience?

(iv) At what value of exper does additional experience actually lower predicted log(wage)? How many people have more experience in this sample?

**C6.3** Consider a model where the return to education depends upon the amount of work experience (and vice versa):

\[ \log(\text{wage}) = β_0 + β_1 \cdot \text{educ} + β_2 \cdot \text{exper} + β_3 \cdot \text{educ-exper} + u. \]

(i) Show that the return to another year of education (in decimal form), holding exper fixed, is $β_1 + β_2 \cdot \text{exper}$.
(ii) State the null hypothesis that the return to education does not depend on the level of exper. What do you think is the appropriate alternative?
(iii) Use the data in WAGE2.RAW to test the null hypothesis in (ii) against your stated alternative.
(iv) Let $θ_1$ denote the return to education (in decimal form), when $\text{exper} = 10$: $θ_1 = β_1 + 10β_2$. Obtain $\hat{θ}_1$ and a 95% confidence interval for $θ_1$. (Hint: Write $β_1 = θ_1 - 10β_2$ and plug this into the equation; then rearrange. This gives the regression for obtaining the confidence interval for $θ_1$.)

**C6.4** Use the data in GPA2.RAW for this exercise.

(i) Estimate the model

\[ \text{sat} = β_0 + β_1 \cdot \text{hsize} + β_2 \cdot \text{hsizes}^2 + u, \]

where hsize is the size of the graduating class (in hundreds), and write the results in the usual form. Is the quadratic term statistically significant?
(ii) Using the estimated equation from part (i), what is the “optimal” high school size? Justify your answer.
(iii) Is this analysis representative of the academic performance of all high school seniors? Explain.
(iv) Find the estimated optimal high school size, using log(sat) as the dependent variable. Is it much different from what you obtained in part (ii)?

**C6.5** Use the housing price data in HPRICE1.RAW for this exercise.

(i) Estimate the model

\[ \log(\text{price}) = β_0 + β_1 \cdot \log(\text{lotsize}) + β_2 \cdot \log(\text{sqrft}) + β_3 \cdot \text{bdrms} + u \]

and report the results in the usual OLS format.
(ii) Find the predicted value of log(price), when \( \text{lotsize} = 20,000 \), \( \text{sqrft} = 2,500 \), and \( \text{hdrms} = 4 \). Using the methods in Section 6.4, find the predicted value of \text{price} at the same values of the explanatory variables.

(iii) For explaining variation in \text{price}, decide whether you prefer the model from part (i) or the model

\[
\text{price} = \beta_0 + \beta_1 \text{lotsize} + \beta_2 \text{sqrft} + \beta_3 \text{hdrms} + u.
\]

C6.6 Use the data in VOTE1.RAW for this exercise.

(i) Consider a model with an interaction between expenditures:

\[
\text{voteA} = \beta_0 + \beta_1 \text{prtystrA} + \beta_2 \text{expendA} + \beta_3 \text{expendB} + \beta_4 \text{expendA} \cdot \text{expendB} + u.
\]

What is the partial effect of \text{expendB} on \text{voteA}, holding \text{prtystrA} and \text{expendA} fixed? What is the partial effect of \text{expendA} on \text{voteA}? Is the expected sign for \( \beta_4 \) obvious?

(ii) Estimate the equation in part (i) and report the results in the usual form. Is the interaction term statistically significant?

(iii) Find the average of \text{expendA} in the sample. Fix \text{expendA} at 300 (for $300,000). What is the estimated effect of another $100,000 spent by Candidate B on \text{voteA}? Is this a large effect?

(iv) Now fix \text{expendB} at 100. What is the estimated effect of \( \Delta \text{expendA} = 100 \) on \text{voteA}? Does this make sense?

(v) Now, estimate a model that replaces the interaction with \text{shareA}, Candidate A’s percentage share of total campaign expenditures. Does it make sense to hold both \text{expendA} and \text{expendB} fixed, while changing \text{shareA}?

(vi) (Requires calculus) In the model from part (v), find the partial effect of \text{expendB} on \text{voteA}, holding \text{prtystrA} and \text{expendA} fixed. Evaluate this at \text{expendA} = 300 and \text{expendB} = 0 and comment on the results.

C6.7 Use the data in ATTEND.RAW for this exercise.

(i) In the model of Example 6.3, argue that

\[
\Delta \text{stdfnl}/\Delta \text{priGPA} = \beta_2 + 2\beta_3 \text{priGPA} + \beta_4 \text{atndrte}.
\]

Use equation (6.19) to estimate the partial effect when \( \text{priGPA} = 2.59 \) and \( \text{atndrte} = 82 \). Interpret your estimate.

(ii) Show that the equation can be written as

\[
\text{stdfnl} = \theta_0 + \beta_1 \text{atndrte} + \theta_2 \text{priGPA} + \beta_3 \text{ACT} + \beta_4 (\text{priGPA} - 2.59)^2
\]
\[
+ \beta_5 \text{ACT}^2 + \beta_6 \text{priGPA}(\text{atndrte} - 82) + u,
\]

where \( \theta_2 = \beta_3 + 2\beta_4 (2.59) + \beta_5 (82) \). (Note that the intercept has changed, but this is unimportant.) Use this to obtain the standard error of \( \hat{\theta}_2 \) from part (i).

(iii) Suppose that, in place of \( \text{priGPA}(\text{atndrte} - 82) \), you put \( (\text{priGPA} - 2.59)(\text{atndrte} - 82) \). Now how do you interpret the coefficients on \text{atndrte} and \text{priGPA}?
C6.8 Use the data in HPRICE1_RAW for this exercise.

(i) Estimate the model

\[ price = \beta_0 + \beta_1 \text{lotsize} + \beta_2 \text{sqft} + \beta_3 \text{bdrms} + u \]

and report the results in the usual form, including the standard error of the regression. Obtain predicted price, when we plug in lotsize = 10,000, sqft = 2,300, and bdrms = 4; round this price to the nearest dollar.

(ii) Run a regression that allows you to put a 95% confidence interval around the predicted value in part (i). Note that your prediction will differ somewhat due to rounding error.

(iii) Let price\(^*\) be the unknown future selling price of the house with the characteristics used in parts (i) and (ii). Find a 95% CI for price\(^*\) and comment on the width of this confidence interval.

C6.9 The data set NBASAL_RAW contains salary information and career statistics for 269 players in the National Basketball Association (NBA).

(i) Estimate a model relating points-per-game (points) to years in the league (exper), age, and years played in college (coll). Include a quadratic in exper; the other variables should appear in level form. Report the results in the usual way.

(ii) Holding college years and age fixed, at what value of experience does the next year of experience actually reduce points-per-game? Does this make sense?

(iii) Why do you think coll has a negative and statistically significant coefficient? (Hint: NBA players can be drafted before finishing their college careers and even directly out of high school.)

(iv) Add a quadratic in age to the equation. Is it needed? What does this appear to imply about the effects of age, once experience and education are controlled for?

(v) Now regress log(wage) on points, exper, exper\(^2\), age, and coll. Report the results in the usual format.

(vi) Test whether age and coll are jointly significant in the regression from part (v). What does this imply about whether age and education have separate effects on wage, once productivity and seniority are accounted for?

C6.10 Use the data in BWGHT2_RAW for this exercise.

(i) Estimate the equation

\[ \log(\text{bwght}) = \beta_0 + \beta_1 \text{npvis} + \beta_2 \text{npvis}^2 + u \]

by OLS, and report the results in the usual way. Is the quadratic term significant?

(ii) Show that, based on the equation from part (i), the number of prenatal visits that maximizes log(bwght) is estimated to be about 22. How many women had at least 22 prenatal visits in the sample?

(iii) Does it make sense that birth weight is actually predicted to decline after 22 prenatal visits? Explain.

(iv) Add mother's age to the equation, using a quadratic functional form. Holding npvis fixed, at what mother's age is the birth weight of the child maximized? What fraction of women in the sample are older than the "optimal" age?