3 Using the data in RDCHEM.RAW, the following equation was obtained by OLS:

$$\hat{rdintens} = 2.613 + .00030 \, sales - .0000000070 \, sales^2$$

$$(.429) \quad (.00014) \quad (.0000000037)$$

$n = 32$, $R^2 = .1484$.

(i) At what point does the marginal effect of $sales$ on $rdintens$ become negative?
(ii) Would you keep the quadratic term in the model? Explain.
(iii) Define $salesbil$ as sales measured in billions of dollars: $salesbil = sales/1,000$.

Rewrite the estimated equation with $salesbil$ and $salesbil^2$ as the independent variables. Be sure to report standard errors and the $R$-squared. [Hint: Note that $salesbil^2 = sales^2/(1,000)^2$.]
(iv) For the purpose of reporting the results, which equation do you prefer?

6 When $atndr^2$ and $ACT\cdot atndr$ are added to the equation estimated in (6.19), the $R$-squared becomes .232. Are these additional terms jointly significant at the 10% level? Would you include them in the model?

7 The following three equations were estimated using the 1,534 observations in 401K.RAW:

$$\hat{prate} = 80.29 + 5.44 \, mrate + .269 \, age - .00013 \, totemp$$

$$(.78) \quad (.52) \quad (.045) \quad (.00004)$$

$R^2 = .100$, $R^2 = .098$.

$$\hat{prate} = 97.32 + 5.02 \, mrate + .314 \, age - 2.66 \, \log(totemp)$$

$$(1.95) \quad (0.51) \quad (.044) \quad (.28)$$

$R^2 = .144$, $R^2 = .142$.

$$\hat{prate} = 80.62 + 5.34 \, mrate + .290 \, age - .00043 \, totemp$$

$$(.78) \quad (.52) \quad (.045) \quad (.00009)$$

$+ .0000000039 \, totemp^2$

$$(.0000000010)$$

$R^2 = .108$, $R^2 = .106$.

Which of these three models do you prefer? Why?
C1 Use the data in KIELMC.RAW, only for the year 1981, to answer the following questions. The data are for houses that sold during 1981 in North Andover, Massachusetts; 1981 was the year construction began on a local garbage incinerator.

(i) To study the effects of the incinerator location on housing price, consider the simple regression model

\[ \log(\text{price}) = \beta_0 + \beta_1 \log(\text{dist}) + u, \]

where \( \text{price} \) is housing price in dollars and \( \text{dist} \) is distance from the house to the incinerator measured in feet. Interpreting this equation causally, what sign do you expect for \( \beta_1 \) if the presence of the incinerator depresses housing prices? Estimate this equation and interpret the results.

(ii) To the simple regression model in part (i), add the variables \( \log(\text{intst}) \), \( \log(\text{area}) \), \( \log(\text{land}) \), \( \text{rooms} \), \( \text{baths} \), and \( \text{age} \), where \( \text{intst} \) is distance from the home to the interstate, \( \text{area} \) is square footage of the house, \( \text{land} \) is the lot size in square feet, \( \text{rooms} \) is total number of rooms, \( \text{baths} \) is number of bathrooms, and \( \text{age} \) is age of the house in years. Now, what do you conclude about the effects of the incinerator? Explain why (i) and (ii) give conflicting results.

(iii) Add \( [\log(\text{intst})]^2 \) to the model from part (ii). Now what happens? What do you conclude about the importance of functional form?

(iv) Is the square of \( \log(\text{dist}) \) significant when you add it to the model from part (iii)?

C2 Use the data in WAGE1.RAW for this exercise.

(i) Use OLS to estimate the equation

\[ \log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 \text{exper}^2 + u \]

and report the results using the usual format.

(ii) Is \( \text{exper}^2 \) statistically significant at the 1% level?

(iii) Using the approximation

\[ \% \Delta \text{wage} \approx 100(\hat{\beta}_2 + 2\hat{\beta}_3 \text{exper}) \Delta \text{exper}, \]

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 \( \text{exper} \) does additional experience actually lower predicted \( \log(\text{wage}) \)? How many people have more experience in this sample?
Use the housing price data in HPRICE1.RAW for this exercise.

(i) Estimate the model

\[ \log(\text{price}) = \beta_0 + \beta_1 \log(\text{lotsize}) + \beta_2 \log(\text{sqrft}) + \beta_3 \text{bdrms} + u \]

and report the results in the usual OLS format.

(ii) Find the predicted value of \( \log(\text{price}) \), when \( \text{lotsize} = 20,000 \), \( \text{sqrft} = 2,500 \), and \( \text{bdrms} = 4 \). 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

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