C3 Refer to Computer Exercise C2 in Chapter 3. Now, use the log of the housing price as the dependent variable:

$$\log(price) = \beta_0 + \beta_1 sqrft + \beta_2 bdrms + u.$$

(i) You are interested in estimating and obtaining a confidence interval for the percentage change in *price* when a 150-square-foot bedroom is added to a house. In decimal form, this is $\theta_1 = 150\beta_1 + \beta_2$. Use the data in HPRICE1 to estimate θ_1.

(ii) Write β_2 in terms of θ_1 and β_1 and plug this into the log(*price*) equation.

(iii) Use part (ii) to obtain a standard error for $\hat{\theta}_1$ and use this standard error to construct a 95% confidence interval.

C8 The data set 401KSUBS contains information on net financial wealth (*nettfa*), age of the survey respondent (*age*), annual family income (*inc*), family size (*fsize*), and participation in certain pension plans for people in the United States. The wealth and income variables are both recorded in thousands of dollars. For this question, use only the data for single-person households (so $fsize = 1$).

(i) How many single-person households are there in the data set?

(ii) Use OLS to estimate the model

$$nettfa = \beta_0 + \beta_1 inc + \beta_2 age + u,$$

and report the results using the usual format. Be sure to use only the single-person households in the sample. Interpret the slope coefficients. Are there any surprises in the slope estimates?

(iii) Does the intercept from the regression in part (ii) have an interesting meaning? Explain.

(iv) Find the p-value for the test $H_0: \beta_2 = 1$ against $H_1: \beta_2 < 1$. Do you reject H_0 at the 1% significance level?

(v) If you do a simple regression of *nettfa* on *inc*, is the estimated coefficient on *inc* much different from the estimate in part (ii)? Why or why not?

C10 Use the data in ELEM94_95 to answer this question. The findings can be compared with those in Table 4.1. The dependent variable *lavgsal* is the log of average teacher salary and *bs* is the ratio of average benefits to average salary (by school).

(i) Run the simple regression of *lavgsal* on *bs*. Is the estimated slope statistically different from zero? Is it statistically different from -1?

(ii) Add the variables *lenrol* and *lstaff* to the regression from part (i). What happens to the coefficient on *bs*? How does the situation compare with that in Table 4.1?

(iii) Why is the standard error on the *bs* coefficient smaller in part (ii) than in part (i)? (*Hint:* What happens to the error variance versus multicollinearity when *lenrol* and *lstaff* are added?)

(iv) How come the coefficient on *lstaff* is negative? Is it large in magnitude?

(v) Now add the variable *lunch* to the regression. Holding other factors fixed, are teachers being compensated for teaching students from disadvantaged backgrounds? Explain.

(vi) Overall, is the pattern of results that you find with ELEM94_95 consistent with the pattern in Table 4.1?