6.4 Consider the wage equation

$$\ln(WAGE) = \beta_1 + \beta_2 EDUC + \beta_3 EDUC^2 + \beta_4 EXPER + \beta_5 EXPER^2 + \beta_6(EDUC \times EXPER) + \beta_7 HRSWK + e$$

where the explanatory variables are years of education, years of experience and hours worked per week. Estimation results for this equation, and for modified versions of it obtained by dropping some of the variables, are displayed in Table 6.4. These results are from the 1000 observations in the file *cps4c_small.dat*.

(a) Using an approximate 5% critical value of $t_c = 2$, what coefficient estimates are not significantly different from zero?

(b) What restriction on the coefficients of Eqn (A) gives Eqn (B)? Use an F-test to test this restriction. Show how the same result can be obtained using a t-test.

(c) What restrictions on the coefficients of Eqn (A) give Eqn (C)? Use an F-test to test these restrictions. What question would you be trying to answer by performing this test?

(d) What restrictions on the coefficients of Eqn (B) give Eqn (D)? Use an F-test to test these restrictions. What question would you be trying to answer by performing this test?

(e) What restrictions on the coefficients of Eqn (A) give Eqn (E)? Use an F-test to test these restrictions. What question would you be trying to answer by performing this test?

(f) Based on your answers to parts (a) to (e), which model would you prefer? Why?

(g) Compute the missing AIC value for Eqn (D) and the missing SC value for Eqn (A). Which model is favored by the AIC? Which model is favored by the SC?

6.5* Consider the wage equation

$$\ln(WAGE) = \beta_1 + \beta_2 EDUC + \beta_3 EDUC^2 + \beta_4 EXPER + \beta_5 EXPER^2 + \beta_6 HRSWK + e$$

(a) Suppose you wish to test the hypothesis that a year of education has the same effect on ln ($WAGE$) as a year of experience. What null and alternative hypotheses would you set up?
(b) What is the restricted model, assuming that the null hypothesis is true?
(c) Given that the sum of squared errors from the restricted model is $SSE_R =$ 254.1726, test the hypothesis in (a). (For SSE_U use the relevant value from Table 6.4. The sample size is $N = 1{,}000$.)

Table 6.4 Wage Equation Estimates for Exercises 6.4 and 6.5

Variable	Coefficient Estimates and (Standard Errors)				
	Eqn (A)	Eqn (B)	Eqn (C)	Eqn (D)	Eqn (E)
C	1.055	1.252	1.573	1.917	0.904
	(0.266)	(0.190)	(0.188)	(0.080)	(0.096)
$EDUC$	0.0498	0.0289	0.0366		0.1006
	(0.0397)	(0.0344)	(0.0350)		(0.0063)
$EDUC^2$	0.00319	0.00352	0.00293		
	(0.00169)	(0.00166)	(0.00170)		
$EXPER$	0.0373	0.0303		0.0279	0.0295
	(0.0081)	(0.0048)		(0.0054)	(0.0048)
$EXPER^2$	−0.000485	−0.000456		−0.000470	−0.000440
	(0.000090)	(0.000086)		(0.000096)	(0.000086)
$EXPER \times EDUC$	−0.000510				
	(0.000482)				
$HRSWK$	0.01145	0.01156	0.01345	0.01524	0.01188
	(0.00137)	(0.00137)	(0.00136)	(0.00151)	(0.00136)
SSE	222.4166	222.6674	233.8317	280.5061	223.6716
AIC	−1.489	−1.490	−1.445		−1.488
SC		−1.461	−1.426	−1.244	−1.463

Does the omission of *CIT* lead to omitted-variable bias? Can you suggest why?

6.15 The file *stockton4.dat* contains data on 1500 houses sold in Stockton, California, during 1996–1998. Variable descriptions are in the file *stockton4.def*.

(a) Estimate the following model and report the results:

$$SPRICE = \beta_1 + \beta_2 LIVAREA + \beta_3 AGE + \beta_4 BEDS + \beta_5 BATHS + e$$

(b) Xiaohui wants to buy a house. She is considering two that have the same living area, the same number of bathrooms, and the same number of bedrooms. One is two years old and the other is ten years old. What price difference can she expect between the two houses? What is a 95% interval estimate for this difference?

(c) Wanling's house has a living area of 2000 square feet. She is planning to extend her living room by 200 square feet. What is the expected increase in price she will get from this extension? Test as an alternative hypothesis that the increase in price will be at least \$20,000. Use $\alpha = 0.05$.

(d) Xueyan's house has a living area of 1800 square feet. She is planning to add another bedroom of size 200 square feet. What is the expected increase in price she will get from this extension? Find a 95% interval estimate for the expected price increase.

(e) Does RESET suggest that the model is a reasonable one?

6.16 Reconsider the data and model estimated in Exercise 6.15.

(a) Add the variables $LIVAREA^2$ and AGE^2 to the model, re-estimate it, and report the results.

(b) Does an F-test suggest that the addition of $LIVAREA^2$ and AGE^2 has improved the model? Use $\alpha = 0.05$.

(c) Answer parts (b)–(e) of Exercise 6.15 using the new specification.

6.22* In Chapter 5.7 we used the data in file *pizza4.dat* to estimate the model

$$PIZZA = \beta_1 + \beta_2 AGE + \beta_3 INCOME + \beta_4 (AGE \times INCOME) + e$$

(a) Test the hypothesis that age does not affect pizza expenditure—that is, test the joint hypothesis $H_0{:}\beta_2 = 0,\ \beta_4 = 0$. What do you conclude?
(b) Construct point estimates and 95% interval estimates of the marginal propensity to spend on pizza for individuals of ages 20, 30, 40, 50, and 55. Comment on these estimates.
(c) Modify the equation to permit a "life-cycle" effect in which the marginal effect of income on pizza expenditure increases with age, up to a point, and then falls. Do so by adding the term $(AGE^2 \times INC)$ to the model. What sign do you anticipate on this term? Estimate the model and test the significance of the coefficient for this variable. Did the estimate have the expected sign?
(d) Using the model in (c), construct point estimates and 95% interval estimates of the marginal propensity to spend on pizza for individuals of ages 20, 30, 40, 50 and 55. Comment on these estimates. In light of these values, and of the range of age in the sample data, what can you say about the quadratic function of age that describes the marginal propensity to spend on pizza?
(e) For the model in part (c), are each of the coefficient estimates for AGE, $(AGE \times INC)$ and $(AGE^2 \times INC)$ significantly different from zero at a 5% significance level? Carry out a joint test for the significance of these variables. Comment on your results.
(f) Check the model used in part (c) for collinearity. Add the term $(AGE^3 \times INC)$ to the model in (c) and check the resulting model for collinearity.