Question 5.1

(a) True. The t test is based on variables with a normal distribution. Since the estimators of β_1 and β_2 are linear combinations of the error u_i, which is assumed to be normally distributed under CLRM, these estimators are also normally distributed.

(b) True. So long as $\mathrm{E}(u_i) = 0$, the OLS estimators are unbiased. No probabilistic assumptions are required to establish unbiasedness.

(c) True. In this case the Eq. (1) in App. 3A, Sec. 3A.1, will be absent. This topic is discussed more fully in Chap. 6, Sec. 6.1.

(d) True. The p value is the smallest level of significance at which the null hypothesis can be rejected. The terms level of significance and size of the test are synonymous.

(e) True. This follows from Eq. (1) of App. 3A, Sec. 3A.1.

(f) False. All we can say is that the data at hand does not permit us to reject the null hypothesis.

(g) False. A larger σ^2 may be counterbalanced by a larger $\sum x_i^2$. It is only if the latter is held constant, the statement can be true.

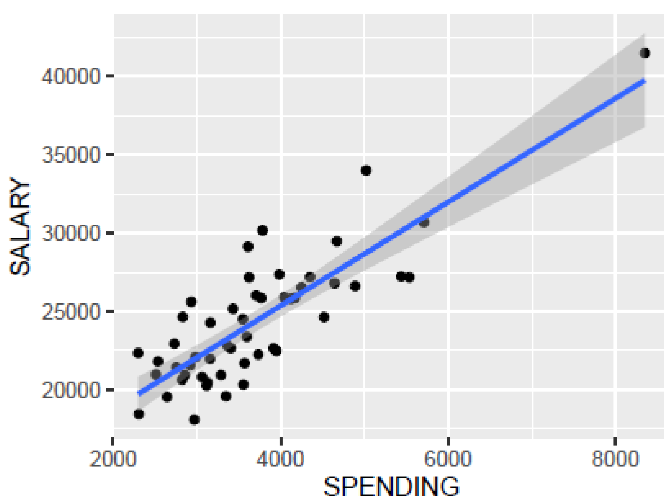
(h) False. The conditional mean of a random variable depends on the values taken by another (conditioning) variable. Only if the two variables are independent, that the conditional and unconditional means can be the same.

(i) True. This is obvious from Eq. (3.1.7).

(j) True. Refer of Eq. (3.5.2). If X has no influence on $Y, \hat{\beta}_2$ will be zero, in which case $\sum y_i^2 = \sum \hat{u}_i^2$

Question 5.9

(a)



(b)

Table 1:

	Dependent variable:
	SALARY
SPENDING	3.308***
	(0.312)
Constant	12,129.370***
	(1,197.351)
Observations	51
R^2	0.697
Adjusted R^2	0.691
Residual Std. Error	2,324.779 (df = 49)
F Statistic	112.600*** (df = 1; 49)
Note:	*p<0.1; **p<0.05; ***p<0.01

$$ESS = 608560533 \quad RSS = 264678860$$

$$\begin{aligned} \text{(b) Pay }_i &= 12129.37 + 3.3076 \text{ Spend} \\ \text{se } &= (1197.351)(0.3117) \\ r^2 &= 0.6968 \text{RSS} = 2.65\text{E}+08 \end{aligned}$$

(c) If the spending per pupil increases by a dollar, the average pay increases by about \$3.31. The intercept term has no viable economic meaning.

(d) The 95%CI for β_2 is: $3.3076 \pm 2(0.3117) = (2.6842, 3.931)$ Based on this CI you will not reject the null hypothesis that the true slope coefficient is 3 .

(e) The mean and individual forecast values are the same, namely, $12129.37 + 3.3076(5000) \approx 28,667$. The standard error of the mean forecast value, using eq.(5.10.2), is 520.5117 (dollars) and the standard error of the individual forecast, using Eq.(5.10.6), is 2382.337. The confidence intervals are:

Mean Prediction: $28,667 \pm 2(520.5117)$, that is $(\$27,626 - \$29,708)$ Individual Prediction: $28667 \pm 2(2382.337)$, that is $(\$23,902 - \$33,432)$

As expected, the latter interval is wider than the former.

Question 6.1

True. Note that the usual OLS formula to estimate the intercept is

$\hat{\beta}_1 =$ (mean of the regressand $-\hat{\beta}_2$ mean of the regressor). But when Y and X are in deviation form, their mean values are always zero. Hence in this case the estimated intercept is also zero.

Question 6.3

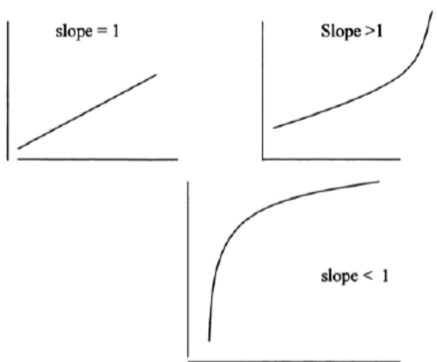
(a) Since the model is linear in the parameters, it is a linear regression model.

(b) Define $Y^* = (1/Y)$ and $X^* = (1/X)$ and do an OLS regression of Y^* on X^*.

(c) As X tends to infinity, Y tends to $(1/\beta_1)$.

(d) Perhaps this model may be appropriate to explain low consumption of a commodity when income is large, such as an inferior good.

Question 6.4



Question 6.6

We can write the first model as: $\ln(\mathrm{w}_1 \mathrm{Y}_\mathrm{i}) = \alpha_1 + \alpha_2 \ln(\mathrm{w}_2 \mathrm{X}_\mathrm{i}) + \mathrm{u}_\mathrm{i}^*$, that is,

$\ln \mathrm{w}1 + \ln \mathrm{Y}_\mathrm{i} = \alpha_1 + \alpha_2 \ln \mathrm{w}_2 + \alpha_2 \ln \mathrm{X}_\mathrm{i} + \mathrm{u}_\mathrm{i}$, using properties of the logarithms. Since the w's are constants, collecting terms, we can simplify this model as:

$\ln Y_i = (\alpha_1 + \alpha_2 - \ln w_1) + \alpha_2 X_i + u_i^*$ where $\mathrm{A} = (\alpha_1 + \alpha_2 \ln \mathrm{w}_2 - \ln \mathrm{w}_1)$

Comparing this with the second model, you will see that except for the intercept terms, the two models are the same. Hence the estimated slope coefficients in the two models will be the same, the only difference being in the estimated intercepts.

(b) The r^2 values of the two models will be the same.

Question 6.14

The regression results are as follows:

Table 1:

	Dependent variable:
	$log(\frac{V}{L})$
$log(wage)$	1.334**
	(0.447)
Constant	−0.453
	(1.351)
Observations	15
R^2	0.407
Adjusted R^2	0.361
Residual Std. Error	0.190 (df = 13)
F Statistic	8.915** (df = 1; 13)
Note:	*p<0.1; **p<0.05; ***p<0.01

To test the null hypothesis, use the t test as follows:

$$t = (1.3338 - 1)/0.4470 = 0.7468$$

For 13 df, the 5% (two-tail) critical t value is 2.16. Therefore, do not reject the hypothesis that the true elasticity of substitution between capital and labor is 1.

Question 7.1

The regression results are:

Table 2:

	Dependent variable:		
	Y		
	(1)	(2)	(3)
X_2	3.50		1.00
	(0.87)		
X_3		-1.36^*	-1.00
		(0.12)	
Constant	-3.00	4.00^{**}	2.00
	(1.87)	(0.27)	
Observations	3	3	3
R^2	0.94	0.99	1.00
Adjusted R^2	0.88	0.98	
Residual Std. Error	1.22 (df = 1)	0.46 (df = 1)	
F Statistic	16.33 (df = 1; 1)	120.33^* (df = 1; 1)	
Note:		*p<0.1; **p<0.05; ***p<0.01	

$\hat{\alpha}_1 = -3.00; \hat{\alpha}_2 = 3.50\ \hat{\lambda}_1 = 4.00; \hat{\lambda}_2 = -1.357\ \hat{\beta}_1 = 2.00; \hat{\beta}_2 = 1.00; \hat{\beta}_3 = -1.00$

(a) No. Given that model (3) is the true model, $\hat{\alpha}_2$ is a biased estimator of β_2

(b) No. $\hat{\lambda}_3$ is a biased estimator of β_3, for the same reason as in (a). The lesson here is that misspecifying an equation can lead to biased estimation of the parameters of the true model.

Question 7.8

If you leave out the years of experience (X_3) from the model, the coefficient of education (X_2) will be biased, the nature of the bias depending on the correlation between X_2 and X_3. The standard error, the residual sum of squares, and R^2 will all be affected as a result of this omission. This is an instance of the omitted variable bias.

Question 7.14

(a) As discussed in Sec. 6.9, to use the classical normal linear regression model (CNLRM), we must assume that $\ln u_i \sim N\left(0, \sigma^2\right)$ After estimating the Cobb-Douglas model, obtain the residuals and subject them to normality test, such as the Jarque-Bera test.

(b) No. As discussed in Sec. 6.9,

$u_i \sim$ log-normal $\left[e^{\sigma 2/2}, e^{\sigma 2/2}\left(e^{\sigma 2/2} - 1\right)\right]$

Question 7.15

(a) The normal equations would be:

$\sum Y_i X_{2i} = \beta_2 \sum X_{2i}^2 + \beta_3 \sum X_{2i} X_{3i}$

$\sum Y_i X_{3i} = \beta_2 \sum X_{2i} X_{3i} + \beta_3 \sum X_{3i}^2$

(b) No, for the same reason as the two-variable case.
(c) Yes, these conditions still hold.
(d) It will depend on the underlying theory.
(e) This is a straightforward generalization of the normal equations given above.

Question 7.18

(a) The regression results are:

Table 3:

	Dependent variable:
	Y
X_2	0.018**
	(0.006)
X_3	−0.284
	(0.457)
X_4	1.343***
	(0.259)
X_5	6.332*
	(3.030)
Constant	19.443***
	(3.406)
Observations	20
R^2	0.978
Adjusted R^2	0.972
Residual Std. Error	4.880 (df = 15)
F Statistic	163.725*** (df = 4; 15)
Note:	*p<0.1; **p<0.05; ***p<0.01

(b) A priori, all the slope coefficients are expected to be positive. Except the coefficient for US military sales, all the other variables have the expected signs and are statistically significant at the 5% level.
(c) Overall federal outlays and some form of trend variable may be valuable.

Question 7.20

(a) Ceteris paribus, on average, a 1% increase in the unemployment rate leads to a 0.34% increase in the quite rate, a 1% increase in the percentage of employees under 25 leads to a 1.22% increase in the quite rate, and 1% increase in the relative manufacturing employment leads to 1.22% increase in the quite rate, a 1% increase in the percentage of women employees leads to a 0.80% increase in the quite rate, and that over the time period under study, the quite rate declined at the rate of 0.54% per year.

(b) Yes, quite rate and the unemployment rate are expected to be negatively related.

(c) As more people under the age of 25 are hired, the quite rate is expected to go up because of turnover among younger workers.

(d) The decline rate is 0.54%. As working conditions and pensions benefits have increased over time, the quit rate has probably declined.

(e) No. Low is a relative term.

(f) Since the t values are given, we can easily compute the standard errors. Under the null hypothesis that the true J_i is zero, we have the relationship: $\text{t} = \frac{\widehat{\beta}_i}{se\left(\widehat{\beta}_i\right)} \to \text{se}\left(\hat{\beta}_i\right) = \frac{t}{\widehat{\beta}_i}$

Question 7.23

The regression results are as follows:

Table 4:

	Dependent variable:	
	Log(wage)	
	(1)	(2)
Log(education)	−0.968	0.935***
	(1.315)	(0.090)
Log(education)^2	0.403	
	(0.278)	
Log(education^2)		
Constant	2.034	−0.167
	(1.532)	(0.221)
Observations	13	13
R^2	0.924	0.908
Adjusted R^2	0.909	0.899
Residual Std. Error	0.104 (df = 10)	0.109 (df = 11)
F Statistic	60.614*** (df = 2; 10)	108.267*** (df = 1; 11)
Note:		*p<0.1; **p<0.05; ***p<0.01

(b) Here you will not be able to estimate the model because of perfect collinearity. This is easy to see: $\log(\text{education}^2) = 2\log(\text{education})$

5.3. (a) se of the slope coefficient is: 0:7241/10.406 = 0.069

the t value under Ho: $\beta_1 = 0$, is: -0:01445/0.875 = -.017

(b) On average, mean hourly wage goes up by about 72 cents for an additional year of schooling.

(c) Here n = 13, so df= 11. If the null hypothesis were true, the estimated t value is 9.6536. The probability of obtaining such a t value is extremely small; the p value is practically zero. Therefore, one can reject the null hypothesis that education has no effect on hourly earnings.

(d) The ESS = 74.9389; RSS = 8.8454; numerator df = 1, denominator df = 11. F = 108.3. The p value of such an F under the null hypothesis that there is no relationship between the two variables is 4.958e-07, which is extremely small. We can thus reject the null hypothesis with great confidence. Note that the F value is approximately the square of the t value under the same null hypothesis.

(e) In the bivariate case, given Ho: $\beta_2 = 0$, there is the following relationship between the t value and r^2

$r^2 = t^2 / [t2+(n-2)]$. Since the t value is given as 10.406, we obtain: $r^2 = 10.406^2 / [(10.406^2) + 11] = 0.907$.