CHAPTER
9
FUNCTIONAL FORMS OF REGRESSION MODELS

QUESTIONS

9.1. **(*a*)** In a log-log model the dependent and all explanatory variables are in the logarithmic form.

(*b*) In the log-lin model the dependent variable is in the logarithmic form but the explanatory variables are in the linear form.

(*c*) In the lin-log model the dependent variable is in the linear form, whereas the explanatory variables are in the logarithmic form.

(*d*) It is the percentage change in the value of one variable for a (small) percentage change in the value of another variable. For the log-log model, the slope coefficient of an explanatory variable gives a direct estimate of the elasticity coefficient of the dependent variable with respect to the given explanatory variable.

(*e*) For the lin-lin model, elasticity = slope$\left(\frac{X}{Y}\right)$. Therefore the elasticity will depend on the values of *X* and *Y*. But if we choose $\bar{X}$ and $\bar{Y}$, the mean values of *X* and *Y*, at which to measure the elasticity, the elasticity at mean values will be: slope$\left(\frac{\bar{X}}{\bar{Y}}\right)$.

9.2. The slope coefficient gives the rate of change in (mean) *Y* with respect to *X*, whereas the elasticity coefficient is the percentage change in (mean) *Y* for a (small) percentage change in *X*. The relationship between two is: Elasticity = slope$\left(\frac{X}{Y}\right)$. For the log-linear, or log-log, model only, the elasticity and slope coefficients are identical.

9.3. <u>Model 1</u>: $\ln Y_i = B_1 + B_2 \ln X_i$: If the scattergram of ln *Y* on ln *X* shows a linear relationship, then this model is appropriate. In practice, such models are used to estimate the elasticities, for the slope coefficient gives a direct estimate of the elasticity coefficient.

Model 2: $\ln Y_i = B_1 + B_2\ X_i$: Such a model is generally used if the objective of the study is to measure the rate of growth of *Y* with respect to *X*. Often, the *X* variable represents time in such models.

Model 3: $Y_i = B_1 + B_2 \ln X_i$: If the objective is to find out the absolute change in *Y* for a relative or percentage change in *X*, this model is often chosen.

Model 4: $Y_i = B_1 + B_2(1/X_i)$: If the relationship between *Y* and *X* is curvilinear, as in the case of the Phillips curve, this model generally gives a good fit.

9.4. (***a***) Elasticity.

(***b***) The absolute change in the mean value of the dependent variable for a proportional change in the explanatory variable.

(***c***) The growth rate.

(***d***) $\dfrac{dY}{dX}\left(\dfrac{X}{Y}\right)$

(***e***) The percentage change in the quantity demanded for a (small) percentage change in the price.

(***f***) Greater than 1; less than 1.

9.5. (***a***) *True*. $\dfrac{d \ln Y}{d \ln X} = \dfrac{dY}{dX}\left(\dfrac{X}{Y}\right)$, which, by definition, is elasticity.

(***b***) *True*. For the two-variable linear model, the slope equals B_2 and the elasticity = slope$\left(\dfrac{X}{Y}\right) = B_2\left(\dfrac{X}{Y}\right)$, which varies from point to point. For the log-linear model, slope $= B_2\left(\dfrac{Y}{X}\right)$, which varies from point to point while the elasticity equals B_2. This can be generalized to a multiple regression model.

(*c*) *True*. To compare two or more R^2s, the dependent variable must be the same.

(*d*) *True*. The same reasoning as in (*c*).

(*e*) *False*. The two r^2 values are not directly comparable.

9.6. The elasticity coefficients for the various models are:

(*a*) $B_2(X_i / Y_i)$ **(*b*)** - $B_2(1/X_iY_i)$ **(*c*)** B_2

(*d*) - $B_2(1/X_i)$ **(*e*)** $B_2(1/Y_i)$ **(*f*)** $B_2(1/X_i)$

Model (*a*) assumes that the income elasticity is dependent on the levels of both income and consumption expenditure. If $B_2 > 0$, Models (*b*) and (*d*) give negative income elasticities. Hence, these models may be suitable for "inferior" goods. Model (*c*) gives constant elasticity at all levels of income, which may not be realistic for all consumption goods. Model (*e*) suggests that the income elasticity is independent of income, *X*, but is dependent on the level of consumption expenditure, *Y*. Finally, Model (*f*) suggests that the income elasticity is independent of consumption expenditure, *Y*, but is dependent on the level of income, *X*.

9.7 **(*a*)** Instantaneous growth: 3.02%; 5.30%; 4.56%; 1.14%.

(*b*) Compound growth: 3.07%; 5.44%; 4.67%; 1.15%.

(*c*) The difference is more apparent than real, for in one case we have annual data and in the other we have quarterly data. A quarterly growth rate of 1.14% is *about* equal to an annual growth rate of 4.56%.

PROBLEMS

9.8. **(*a*)** MC = $B_2 + 2B_3X_i + 3B_4X_i^2$

(*b*) AVC = $B_2 + B_3X_i + B_4X_i^2$

(*c*) AC = $B_1\left(\dfrac{1}{X_i}\right) + B_2 + B_3X_i + B_4X_i^2$

By way of an example based on actual numbers, the MC, AVC, and AC from Equation (9.33) are as follows:

MC = 63.4776 – 25.9230 X_i + 2.8188 X_i^2

AVC = 63.4776 – 12.9615 X_i + 0.9396 X_i^2

AC = 141.7667$\left(\dfrac{1}{X_i}\right)$ + 63.4776 – 12.9615 X_i + 0.9396 X_i^2

(*d*) The plot will show that they do indeed resemble the textbook U-shaped cost curves.

9.9. **(*a*)** $\left(\dfrac{1}{Y_i}\right) = B_1 + B_2 X_i$ **(*b*)** $\left(\dfrac{X_i}{Y_i}\right) = B_1 + B_2 X_i^2$

9.10. **(*a*)** In Model A, the slope coefficient of -0.4795 suggests that if the price of coffee per pound goes up by a dollar, the average consumption of coffee per day goes down by about half a cup. In Model B, the slope coefficient of -0.2530 suggests that if the price of coffee per pound goes up by 1%, the average consumption of coffee per day goes down by about 0.25%.

(*b*) Elasticity = -0.4795$\left(\dfrac{1.11}{2.43}\right)$ = -0.2190

(*c*) -0.2530

(*d*) The demand for coffee is price inelastic, since the absolute value of the two elasticity coefficients is less than 1.

(*e*) Antilog (0.7774) = 2.1758. In Model B, if the price of coffee were $1, on average, people would drink approximately 2.2 cups of coffee per day. [*Note*: Keep in mind that ln(1) = 0].

(*f*) We cannot compare the two r^2 values directly, since the dependent variables in the two models are different.

9.11. **(*a*)** *Ceteris paribus*, if the labor input increases by 1%, output, on average, increases by about 0.34%. The computed elasticity is different from 1, for

$$t = \frac{0.3397 - 1}{0.1857} = -3.5557$$

For 17 d.f., this *t* value is statistically significant at the 1% level of significance (two-tail test).

(***b***) *Ceteris paribus*, if the capital input increases by 1%, on average, output increases by about 0.85 %. This elasticity coefficient is statistically different from zero, but not from 1, because under the respective hypothesis, the computed *t* values are about 9.06 and -1.65, respectively.

(***c***) The antilog of -1.6524 = 0.1916. Thus, if the values of $X_2 = X_3 = 1$, then Y = 0.1916 or (0.1916)(1,000,000) = 191,600 pesos.

Of course, this does not have much economic meaning [*Note*: ln(1) = 0].

(***d***) Using the R^2 variant of the *F* test, the computed *F* value is:

$$F = \frac{0.995/2}{(1-0.995)/17} = 1{,}691.50$$

This *F* value is obviously highly significant. So, we can reject the null hypothesis that $B_2 = B_3 = 0$. The critical *F* value is $F_{2,17} = 6.11$ for $\alpha = 1\%$.

Note: The slight difference between the calculated *F* value here and the one shown in the text is due to rounding.

9.12. (***a***) *A priori*, the coefficients of ln(*Y* / *P*) and $\ln \sigma_{\text{BP}}$ should be positive and the coefficient of $\ln \sigma_{\text{EX}}$ should be negative. The results meet the prior expectations.

(***b***) Each partial slope coefficient is a partial elasticity, since it is a log-linear model.

(***c***) As the 1,120 observations are quite a large number, we can use the normal distribution to test the null hypothesis. At the 5% level of significance, the critical (standardized normal) *Z* value is 1.96. Since, in absolute value, each estimated *t* coefficient exceeds 1.96, each estimated coefficient is statistically different from zero.

(***d***) Use the *F* test. The author gives the *F* value as 1,151, which is highly statistically significant. So, reject the null hypothesis.

9.13. (***a***) If (1 / *X*) goes up by a unit, the average value of *Y* goes up by 8.7243.

(***b***) Under the null hypothesis, $t = \dfrac{8.7243}{2.8478}$ = 3.0635, which is statistically significant at the 5% level. Hence reject the null hypothesis.

(***c***) Under the null hypothesis, $F_{1,15} = t_{15}^2$, which is the case here, save the rounding errors.

(***d***) For this model: slope = $-B_2\left(\frac{1}{X_t^2}\right) = -8.7243\left(\frac{1}{2.25}\right) = -3.8775$.

(***e***) Elasticity = $-B_2\left(\frac{1}{X_t Y_t}\right) = -8.7243\left(\frac{1}{(1.5)(4.8)}\right) = -1.2117$

Note: The slope and elasticity are evaluated at the mean values of *X* and *Y*.

(***f***) The computed *F* value is 9.39, which is significant at the 1% level, since for 1 and 15 d.f. the critical *F* value is 8.68. Hence reject the null hypothesis that $r^2 = 0$.

9.14. (***a***) The results of the four regressions are as follows:

	Dependent Variable	**Intercept**	**Independent Variable**	**Goodness of Fit**
1	$\hat{Y}_t$ =	38.9690	+ 0.2609 X_t	$r^2 = 0.9423$
	t =	(10.105)	(15.655)	
2	$\widehat{\ln} Y_t$ =	1.4041	+ 0.5890 ln X_t	$r^2 = 0.9642$
	t =	(8.954)	(20.090)	
3	$\widehat{\ln} Y_t$ =	3.9316	+ 0.0028 X_t	$r^2 = 0.9284$
	t =	(84.678)	(13.950)	
4	$\hat{Y}_t$ =	-192.9661	+ 54.2126 ln X_t	$r^2 = 0.9543$
	t =	(-11.781)	(17.703)	

(***b***) In Model (1), the slope coefficient gives the absolute change in the mean value of *Y* per unit change in *X*. In Model (2), the slope gives the elasticity coefficient. In Model (3) the slope gives the (instantaneous) rate of growth in (mean) *Y* per unit change in *X*. In Model (4), the slope gives the absolute change in mean *Y* for a relative change in *X*.

(***c***) 0.2609; 0.5890(*Y* / *X*); 0.0028(*Y*); 54.2126(1 / *X*).

(***d***) 0.2609(*X* / *Y*); 0.5890; 0.0028(*X*); 54.2126(1 / *Y*).

For the first, third and the fourth model, the elasticities at the mean values are, respectively, 0.5959, 0.6165, and 0.5623.

(*e*) The choice among the models ultimately depends on the end use of the model. Keep in mind that in comparing the r^2 values of the various models, the dependent variable must be in the same form.

9.15. **(*a*)** $\hat{\dfrac{1}{Y_i}} = 0.0130 + 0.0000833\,X_i$

$t = (17.206) \qquad (5.683) \qquad\qquad r^2 = 0.8015$

The slope coefficient gives the rate of change in mean (1 / *Y*) per unit change in *X*.

(*b*) $\dfrac{dY}{dX} = -\dfrac{B_2}{(B_1 + B_2 X_i)^2}$

At the mean value of *X*, $\bar{X}$ = 38.9, this derivative is -0.3146.

(*c*) Elasticity = $\dfrac{dY}{dX}\left(\dfrac{X}{Y}\right)$. At $\bar{X}$ = 38.9 and $\bar{Y}$ = 63.9, this elasticity coefficient is -0.1915.

(*d*) $\hat{Y}_i = 55.4871 + 112.1797\left(\dfrac{1}{X_i}\right)$

$t = (17.409) \qquad (4.245) \qquad\qquad r^2 = 0.6925$

(*e*) No, because the dependent variables in the two models are different.

(*f*) Unless we know what Y and X stand for, it is difficult to say which model is better.

9.16. For the linear model, r^2 = 0.99879, and for the log-lin model, r^2 = 0.99965. Following the procedure described in the problem, r^2 = 0.99968, which is comparable with the r^2 = 0.99879.

9.17. **(*a*)** *Log-linear model*: The slope and elasticity coefficients are the same. *Log-lin model*: The slope coefficient gives the growth rate.

Lin-log model: The slope coefficient gives the absolute change in GNP for a percentage in the money supply.

Linear-in-variable model: The slope coefficient gives the (absolute) rate of change in mean GNP for a unit change in the money supply.

(*b*) The elasticity coefficients for the four models are:

Log-linear:	0.9882
Log-lin (Growth):	1.0007 (at $\bar{X}$ = 1,755.667)
Lin-log:	0.9260 (at $\bar{Y}$ = 2,791.473)
Linear (LIV):	0.9637 (at $\bar{X}$ = 1,755.667 and $\bar{Y}$ = 2,791.473).

(*c*) The r^2s of the log-linear and log-lin models are comparable, as are the r^2s of the lin-log and linear (LIV) models.

(*d*) Judged by the usual criteria of the t test, r^2 values, and the elasticities, all the models more or less give similar results.

(*e*) From the log-linear model, we observe that for a 1% increase in the money supply, on the average, GNP increases by about 1%, the coefficient 0.9882 being statistically equal to 1. Perhaps this model supports the monetarist view. Since the elasticity coefficients of the other models are similar, it seems all the models support the monetarists.

9.18. **(*a*)** $\hat{Y}_t = 28.3407 + 0.9817 X_{2t} - 0.2595 X_{3t}$

se = (1.4127) (0.0193) (0.0152)

t = (20.0617) (50.7754) (-17.0864) $R^2 = 0.9940$

p value = (0.0000)* (0.0000)* (0.0000)* $\bar{R}^2 = 0.9934$

* Denotes a very small value.

(*b*) Per unit change in the real GDP index, on average, the energy demand index goes up by about 0.98 points, *ceteris paribus*. Per unit change in the energy price, the energy demand index goes down about 0.26 points, again holding all else constant.

(*c*) From the *p* values given in the above regression, all the partial regression coefficients are individually highly statistically significant.

(*d*) The values required to set up the ANOVA table are: TSS = 6,746.9887; ESS = 6,706.2863, and RSS = 40.7024. The computed *F* value is 1,647.638 with a p value of almost zero. Therefore, we can reject the null hypothesis that there is no relationship between energy demand, real GDP, and energy prices (*Note*: These ANOVA numbers can easily be calculated with the regression options in *Excel*).

(*e*) Mean value of demand = 84.370; mean value of real GDP = 89.626, and the mean value of energy price = 123.135, all in index form. Therefore, at the mean values, the elasticity of demand with respect to real GDP is 1.0428 and with respect to energy price, it is -0.3787.

(*f*) This is straightforward.

(*g*) The normal probability plot will show that the residuals from the regression model lie approximately on a straight line, indicating that the error term in the regression model seems to be normally distributed. The Anderson-Darling normality test gives an A^2 value of 0.502, whose *p* value is about 0.188, thereby supporting the normality assumption.

(*h*) The normality plot will show that the residuals do not lie on a straight line, suggesting that the normality assumption for the error term may not be tenable for the log-linear model. The computed Anderson-Darling A^2 is 1.020 with a p value of about 0.009, which is quite low.

Note: Any minor coefficient differences between this log-linear regression and the log-linear regression (9.12) are due to rounding. Regarding the Anderson-Darling test, it is available in *MINITAB*. If you do not have access to *MINITAB*, you can use the normal probability plots in *EViews* and *Excel* for a visual inspection, as described above. *EViews* also has the Jarque-Bera normality test, but you should avoid using it here because it is a large sample asymptotic test and the present data set has only 23 observations. In fact, the Jarque-Bera test will show that the residuals of both the linear and the log-linear regressions satisfy the normality assumption, which is not the case based on the Anderson-Darling A^2 and the normal probability plots.

(*i*) Since the linear model seems to satisfy the normality assumption, this model may be preferable to the log-linear model.

9.19. **(*a*)** This will make the model linear in the parameters.

(***b***) The slope coefficients in the two models are, respectively:

$$\frac{dY}{dt} = -\frac{B}{(A+Bt)^2} \qquad \text{and} \qquad -\frac{1}{Y^2}\frac{dY}{dt} = B$$

(***c***) In models (1) and (2) the slope coefficients are negative and are statistically significant, since the *t* values are so high. In both models the reciprocal of the loan amount has been decreasing over time. From the slope coefficients already given, we can compute the rate of change of loans over time.

(***d***) Divide the estimated coefficients by their *t* values to obtain the standard errors.

(***e***) Suppose for Model 1 we postulate that the true *B* coefficient is -0.14. Then, using the *t* test, we obtain:

$$t = \frac{-0.20-(-0.14)}{0.0082} = -7.3171$$

This *t* value is statistically significant at the 1% level. Hence, it seems there is a difference in the loan activity of New York and non-New York banks. [*Note*: s.e. = (-0.20)/(-24.52) = 0.0082].

9.20. (***a***) For the reciprocal model, as Table 9-11 shows, the slope coefficient (i.e., the rate of change of *Y* with respect to *X* is $-B_2(1/X^2)$. In the present instance B_2 = 0.0549. Therefore, the value of the slope will depend on the value taken by the *X* variable.

(***b***) For this model the elasticity coefficient is $-B_2(1/XY)$. Obviously, this elasticity will depend on the chosen values of *X* and *Y*. Now, $\bar{X}$ = 28.375 and $\bar{Y}$ = 0.4323. Evaluating the elasticity at these means, we find it to be equal to -0.0045.

9.21. We have the following variable definitions:

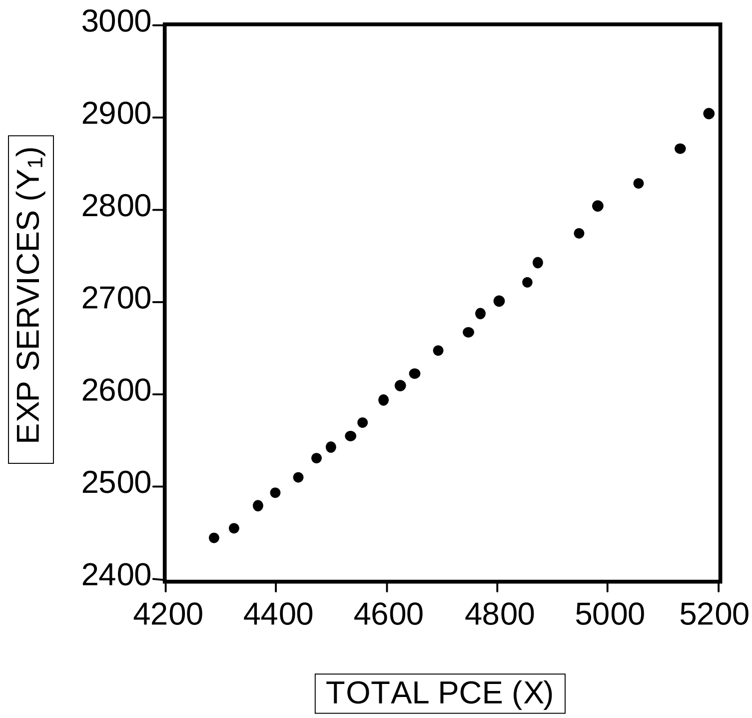
TOTAL PCE (X) = Total personal consumption expenditure;

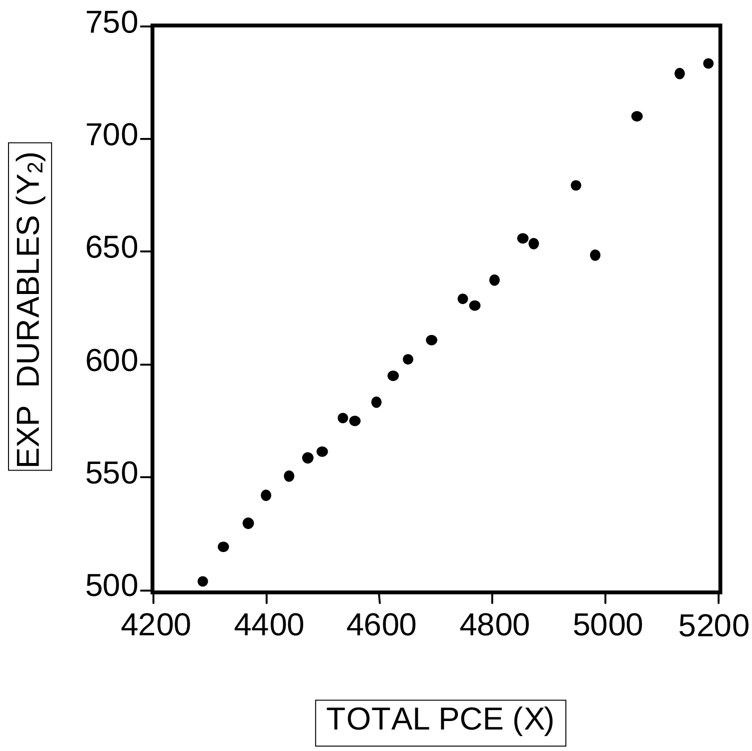
EXP SERVICES (Y_1) = Expenditure on services;

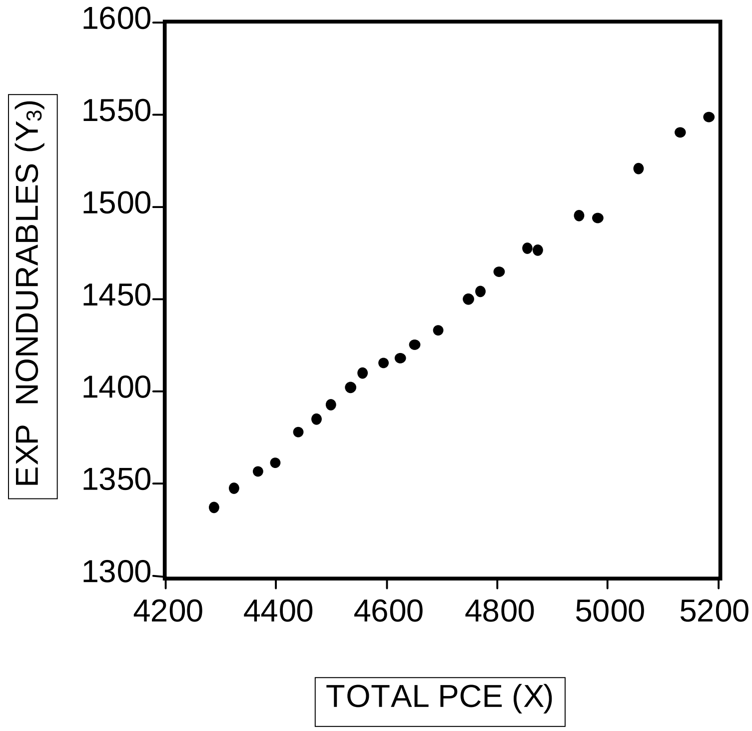
EXP DURABLES (Y_2) = Expenditure on durable goods;

EXP NONDURABLES (Y_3) = Expenditure on nondurable goods.

Plotting the data, we obtain the following scatter graphs:







It seems that the relationship between the various expenditure categories and total personal consumption expenditure is approximately linear. Hence, as a first step one could apply the linear (in variables) model to the various categories. The regression results are as follows: (the independent variable is TOTAL PCE and figures in the parentheses are the estimated t values).

Dependent variable	**EXP SERVICES (Y_1)**	**EXP DURABLES (Y_2)**	**EXP NONDURABLES (Y_3)**
Intercept	222.5759 (11.9281)	-554.5943 (-16.8744)	335.7624 (24.7647)
Slope (TOTAL PCE)	0.5164 (129.8600)	0.2484 (35.4682)	0.2345 (81.1599)
R^2	0.9988	0.9836	0.9968

Judged by the usual criteria, the results seem satisfactory. In each case the slope coefficient represents the marginal propensity of expenditure (MPE) that is the additional expenditure for an additional dollar of TOTAL PCE.

This is highest for services, followed by durable and nondurable goods expenditures. By fitting a double-log model one can obtain the various elasticity coefficients.

9.22. The *EViews* results for the first model are as follows:

Dependent Variable: Y Sample: 1971 1980				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.279719	7.688560	0.166445	0.8719
X	1.069084	0.238315	4.486004	0.0020
R-squared	0.715548			

The output for the regression-through-the-origin model is:

Dependent Variable: Y Sample: 1971 1980				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	1.089912	0.191551	5.689922	0.0003
R-squared	0.714563			

- This R^2 may not be reliable.

Since the intercept in the first model is not statistically significant, we can choose the second model.

9.23. Using the raw r^2 formula, we obtain :

$$\text{raw } r^2 = \frac{(\sum X_i Y_i)^2}{\sum X_i^2 \sum Y_i^2} = \frac{(11{,}344.28)^2}{(10{,}408.44)(15{,}801.41)} = 0.7825$$

You can compare this with the intercept-present R^2 value of 0.7155.

9.24. Computations will show that the raw r^2 is 0.7318. The one in Equation (9.40) is 0.7353. There is not much difference between the two values. Any minor differences between regressions (9.39) and (9.40) in the text and the same regressions based on Table 6-12 are due to rounding.