Practice answers

10.27 Table 10.3 gives data on import, GDP, and the Consumer Price Index (CPI) for the United States over the period 1975-2005. You are asked to consider the following model:

$$\ln imports_t = \beta_1 + \beta_2 \ln GDP_t + \beta_3 \ln CPI_t + u_t$$

a. Estimate the parameters of this model using the data

Source	SS	df	MS
Model	17.2842182	2	8.64210911
Residual	.139293473	28	.004974767
Total	17.4235117	30	.580783723

Number of obs	= 31
F(2, 28)	= 1737.19
Prob > F	= 0.0000
R-squared	= 0.9920
Adj R-squared	= 0.9914
Root MSE	= .07053

lnimports	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lngdp	1.850098	.1829119	10.11	0.000	1.47542	2.224777
lncpi	-.873369	.2848058	-3.07	0.005	-1.456767	-.2899708
_cons	1.409415	.2700745	5.22	0.000	.8561926	1.962638

b. Do you suspect that these is multicollinearity in the data?

Judged by the high *R*2 value and the negative coefficient on the log CPI variable, there *might* some multicollinearity in the data.

c. Regress:

$$(1) \ln imports_t = A_1 + A_2 \ln GDP_t$$
$$(2) \ln imports_t = B_1 + B_2 \ln CPI_t$$
$$(3) \ln GDP_t = C_1 + C_2 \ln CPI_t$$

On the basis of these regressions, what can you say about the nature of multicollinearity in the data?

$$(1)\ln imports_t = A_1 + A_2 \ln GDP_t$$

Source	SS	df	MS
Model	17.2374371	1	17.2374371
Residual	.1860746	29	.006416366
Total	17.4235117	30	.580783723

Number of obs	=	31
F(1, 29)	=	2686.48
Prob > F	=	0.0000
R-squared	=	0.9893
Adj R-squared	=	0.9890
Root MSE	=	.0801

lnimports	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lngdp	1.293252	.0249512	51.83	0.000	1.242221	1.344283
_cons	2.00191	.2143083	9.34	0.000	1.5636	2.440219

$$(2)\ln imports_t = B_1 + B_2 \ln CPI_t$$

Source	SS	df	MS
Model	16.7752643	1	16.7752643
Residual	.648247373	29	.022353358
Total	17.4235117	30	.580783723

Number of obs	=	31
F(1, 29)	=	750.46
Prob > F	=	0.0000
R-squared	=	0.9628
Adj R-squared	=	0.9615
Root MSE	=	.14951

lnimports	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncpi	1.986499	.0725145	27.39	0.000	1.83819	2.134808
_cons	3.578294	.3480571	10.28	0.000	2.866437	4.290151

$$(3)\ln GDP_t = C_1 + C_2 \ln CPI_t$$

Source	SS	df	MS
Model	10.1576897	1	10.1576897
Residual	.148692396	29	.005127324
Total	10.3063821	30	.34354607

Number of obs	31
F(1, 29)	1981.09
Prob > F	0.0000
R-squared	0.9856
Adj R-squared	0.9851
Root MSE	.07161

lngdp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncpi	1.545792	.0347295	44.51	0.000	1.474763	1.616822
_cons	1.172304	.1666957	7.03	0.000	.8313734	1.513236

The auxiliary regression of LN_GDP on LN_CPI shows that the two variables are highly correlated, perhaps suggesting that the data suffer from the collinearity problem.

d. Suppose there is multicollinearity in the data but $\hat{\beta}_2$ and $\hat{\beta}_3$ are individually significant at the 5 percent level and the overall F-test is also significant. In this case should we worry about the collinearity problem?

The best solutions here would be to express imports and GDP in real terms by dividing each by CPI (the ratio method). The results are as follows:

Source	SS	df	MS
Model	4.64571883	1	4.64571883
Residual	.139534936	29	.00481155
Total	4.78525377	30	.159508459

Number of obs	31
F(1, 29)	965.53
Prob > F	0.0000
R-squared	0.9708
Adj R-squared	0.9698
Root MSE	.06937

lnimpcpi	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lngdpcpi	1.811942	.0583123	31.07	0.000	1.69268	1.931204
_cons	1.442444	.2210172	6.53	0.000	.9904132	1.894475

Chapter 11

11.1 State with brief reason whether the following statements are true, false, or uncertain

a. In the presence of heteroscedasticity OLS estimators are biased as well as inefficient.
False. The estimators are unbiased but are inefficient.

b. If heteroscedasticity is present, the conventional t and F tests are invalid
True.

c. In the presence of heteroscedasticity the usual OLS method always overestimates the standard errors of estimators.
False. Typically, but not always, will the variance be overestimated.

d. If residuals estimated from an OLS regression exhibit a systematic pattern, it means heteroscedasticity is present in the data
False. Besides heteroscedasticity, such a pattern may result from autocorrelation, model specification errors, etc.

11.15

Table 11.7 gives data on 81 cars about MPG (average miles per gallons), HP (engine horsepower), VOL (cubic feet of cab space), SP (top speed, miles per hour), and WT (vehicle weight in 100 lbs)

a. Consider the following model:

$$MPG_i = \beta_1 + \beta_2 SP_i + \beta_3 HP_i + \beta_4 WT_i + u_i$$

Estimate the parameters of this model and interpret the results.

Source	SS	df	MS
Model	7141.40496	3	2380.46832
Residual	947.498564	77	12.3051762
Total	8088.90352	80	101.111294

Number of obs	= 81
F(3, 77)	= 193.45
Prob > F	= 0.0000
R-squared	= 0.8829
Adj R-squared	= 0.8783
Root MSE	= 3.5079

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sp	-1.271697	.2331174	-5.46	0.000	-1.735894	-.8075013
hp	.3904334	.0762458	5.12	0.000	.2386086	.5422582
wt	-1.903273	.1855158	-10.26	0.000	-2.272682	-1.533864
_cons	189.9597	22.52879	8.43	0.000	145.0991	234.8202

As expected, MPG is positively related to HP and negatively related to speed and weight.

b. Use the white test to find out if the error variance is heteroscedastic.

$H_0 : Homoscedastic$

$H_1 : Heteroscedastic$

White's general test statistic: 37.65619

Compare with chi-square distribution table at degree of freedom 9, the critical value is 16.9190

Reject the null hypothesis

That is, there is heteroscedasticity.

Source	SS	df	MS
Model	7141.40496	3	2380.46832
Residual	947.498564	77	12.3051762
Total	8088.90352	80	101.111294

Number of obs	= 81
F(3, 77)	= 193.45
Prob > F	= 0.0000
R-squared	= 0.8829
Adj R-squared	= 0.8783
Root MSE	= 3.5079

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sp	-1.271697	.2331174	-5.46	0.000	-1.735894	-.8075013
hp	.3904334	.0762458	5.12	0.000	.2386086	.5422582
wt	-1.903273	.1855158	-10.26	0.000	-2.272682	-1.533864
_cons	189.9597	22.52879	8.43	0.000	145.0991	234.8202


```
. whitetst

White's general test statistic :  37.65619  Chi-sq( 9)  P-value =  2.0e-05
```

Chapter 12

12.1State whether the following statements are true or false. Briefly justify your answer

a. When autocorrelation is present, OLS estimators are biased as well as inefficient.
   False. The estimators are unbiased but they are not efficient.

b. The Durbin-Watson d test assumes that the variance of the error term u_t is homoscedastic.
   True. We are still retaining the other assumptions of CLRM.

c. The R-squared values of two models, one involving regression in the first-difference form and another in the level form, are not directly comparable.
   True. To compare *R*2s, the regressand in the two models must be the same.

d. The exclusion of an important variable(s) from a regression model may give a significant d value.
   True.

12.2

For n = 50 and *k*' = 4, and a = 5%, the critical *d* values are:
*d*L = 1.38 4 – *d*L = 2.62
*d*U = 1.72 4 - *d*U = 2.28
(a) positive autocorrelation; (*b*) inconclusive, (*c*) inconclusive; and
(*d*) negative autocorrelation.

12.15

Since the model contains the lagged dependent variable as a regressor, the Durbin-Watson d is not the appropriate test statistic.

12.34 a&b

Source	SS	df	MS
Model	6.5238e+11	1	6.5238e+11
Residual	3.9448e+10	39	1.0115e+09
Total	6.9183e+11	40	1.7296e+10

Number of obs	= 41
F(1, 39)	= 644.98
Prob > F	= 0.0000
R-squared	= 0.9430
Adj R-squared	= 0.9415
Root MSE	= 31804

inventories	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sales	1.183862	.0466152	25.40	0.000	1.089574	1.278151
_cons	66668.97	10895.73	6.12	0.000	44630.28	88707.65

Durbin-Watson d-statistic(2, 41) = .1255528

$$H_0 : No\ \ positive\ autocorrelation$$
$$H_0^* : No\ negative\ autocorrelation$$
$$H_1 : otherwise$$
$$H_1^* : otherwise$$

At 1% significant level

$d_L = 1.198$

$d_U = 1.398$

0.1255 falls in in the area between 0 and 1.198

Reject null hypothesis

Evidence of positive autocorrelation