**Solutions to the Review Questions at the End of Chapter 5**

1. In the same way as we make assumptions about the true value of beta and not the estimated values, we make assumptions about the true unobservable disturbance terms rather than their estimated counterparts, the residuals.

We know the exact value of the residuals, since they are defined by . So we do not need to make any assumptions about the residuals since we already know their value. We make assumptions about the unobservable error terms since it is always the true value of the population disturbances that we are really interested in, although we never actually know what these are.

2. We would like to see no pattern in the residual plot! If there is a pattern in the residual plot, this is an indication that there is still some ‘action’ or variability left in *yt* that has not been explained by our model. This indicates that potentially it may be possible to form a better model, perhaps using additional or completely different explanatory variables, or by using lags of either the dependent or of one or more of the explanatory variables. Recall that the two plots shown on pages 157 and 159, where the residuals followed a cyclical pattern, and where they followed an alternating pattern, are used as indications that the residuals are positively and negatively autocorrelated, respectively.

Another problem if there is a ‘pattern’ in the residuals is that, if it does indicate the presence of autocorrelation, then this may suggest that our standard error estimates for the coefficients could be wrong and hence any inferences we make about the coefficients could be misleading.

3. The *t*-ratios for the coefficients in this model are given in the third row after the standard errors. They are calculated by dividing the individual coefficients by their standard errors.

 = 0.638 + 0.402 *x*2*t* – 0.891 *x*3*t* 

 (0.436) (0.291) (0.763)

*t*-ratios 1.46 1.38 –1.17

The problem appears to be that the regression parameters are all individually insignificant (i.e., not significantly different from zero), although the value of R2 and its adjusted version are both very high, so that the regression taken as a whole seems to indicate a good fit. This looks like a classic example of what we term ‘near multicollinearity’. This is where the individual regressors are very closely related, so that it becomes difficult to disentangle the effect of each individual variable upon the dependent variable.

The solution to near multicollinearity that is usually suggested is that since the problem is really one of insufficient information in the sample to determine each of the coefficients, then one should go out and get more data. In other words, we should switch to a higher frequency of data for analysis (e.g., weekly instead of monthly, monthly instead of quarterly, etc.). An alternative is also to get more data by using a longer sample period (i.e., one going further back in time), or to combine the two independent variables in a ratio (e.g., *x*2*t* / *x*3*t* ).

Other, more *ad hoc* methods for dealing with the possible existence of near multicollinearity were discussed in Chapter 5

* Ignore it: if the model is otherwise adequate, i.e., statistically and in terms of each coefficient being of a plausible magnitude and having an appropriate sign. Sometimes, the existence of multicollinearity does not reduce the *t*-ratios on variables that would have been significant without the multicollinearity sufficiently to make them insignificant. It is worth stating that the presence of near multicollinearity does not affect the BLUE properties of the OLS estimator – i.e., it will still be consistent, unbiased and efficient since the presence of near multicollinearity does not violate any of the CLRM assumptions (1)–(4). However, in the presence of near multicollinearity, it will be hard to obtain small standard errors. This will not matter if the aim of the model-building exercise is to produce forecasts from the estimated model, since the forecasts will be unaffected by the presence of near multicollinearity so long as this relationship between the explanatory variables continues to hold over the forecasted sample.
* Drop one of the collinear variables – so that the problem disappears. However, this may be unacceptable to the researcher if there were strong *a priori* theoretical reasons for including both variables in the model. Also, if the removed variable was relevant in the data generating process for *y*, an omitted variable bias would result.
* Transform the highly correlated variables into a ratio and include only the ratio and not the individual variables in the regression. Again, this may be unacceptable if financial theory suggests that changes in the dependent variable should occur following changes in the individual explanatory variables, and not a ratio of them.

4. (a) The assumption of homoscedasticity is that the variance of the errors is constant and finite over time. Technically, we write .

(b) The coefficient estimates would still be the ‘correct’ ones (assuming that the other assumptions required to demonstrate OLS optimality are satisfied), but the problem would be that the standard errors could be wrong. Hence if we were trying to test hypotheses about the true parameter values, we could end up drawing the wrong conclusions. In fact, for all of the variables except the constant, the standard errors would typically be too small, so that we would end up rejecting the null hypothesis too many times.

(c) There are a number of ways to proceed in practice, including

- Using heteroscedasticity robust standard errors which correct for the problem by enlarging the standard errors relative to what they would have been for the situation where the error variance is positively related to one of the explanatory variables.

- Transforming the data into logs, which has the effect of reducing the effect of large errors relative to small ones.

5. (a) This is where there is a relationship between the *i*th and *j*th residuals. Recall that one of the assumptions of the CLRM was that such a relationship did not exist. We want our residuals to be random, and if there is evidence of autocorrelation in the residuals, then it implies that we could predict the sign of the next residual and get the right answer more than half the time on average!

(b) The Durbin–Watson test is a test for first order autocorrelation. The test is calculated as follows. You would run whatever regression you were interested in, and obtain the residuals. Then calculate the statistic



You would then need to look up the two critical values from the Durbin Watson tables, and these would depend on how many variables and how many observations and how many regressors (excluding the constant this time) you had in the model.

The rejection/non-rejection rule would be given by selecting the appropriate region from the following diagram:



(c) We have 60 observations, and the number of regressors excluding the constant term is 3. The appropriate lower and upper limits are 1.48 and 1.69. respectively, so the Durbin–Watson is lower than the lower limit. It is thus clear that we reject the null hypothesis of no autocorrelation. So it looks like the residuals are positively autocorrelated.