**Chapter 7 Summary**

We have covered a lot of ground in this chapter on a variety of practical topics in econometric modeling.

If we omit a relevant variable(s) from a regression model, the estimated coefficients and standard errors of OLS estimators in the reduced model are biased as well as inconsistent. We considered the RESET and Lagrange Multiplier tests to detect the omission of relevant variables bias.

If we add unnecessary variables to a model, the OLS estimators of the expanded model are still BLUE. The only penalty we pay is the loss of efficiency (i.e. increased standard errors) of the estimated coefficients.

The appropriate functional form of a regression model is a commonly encountered question in practice. In particular, we often face a choice between a linear and a log-linear model. We showed how we can compare the two models in making the choice, using the Cobb–Douglas production function data for the 50 states in the USA and Washington, DC, as an example.

Errors of measurement are a common problem in empirical work, especially if we depend on secondary data. We showed that the consequences of such errors can be very serious if they exist in explanatory variables, for in that case the OLS estimators are not even consistent. Errors of measurement do not pose a serious problem if they are in the dependent variable. In practice, however, it is not always easy to spot the errors of measurement. The method of instrumental variables, discussed in Chapter 19, is often suggested as a remedy for this problem.

Generally we use the sample data to draw inferences about the relevant population. But if there are “unusual observations” or outliers in the sample data, inferences based on such data may be misleading. Therefore we need to pay special attention to outlying observations. Before we throw out the outlying observations, we must be very careful to find out why the outliers are present in the data. Sometimes they may result from human errors in recording or transcribing the data. We illustrated the problem of outliers with data on cigarette smoking and deaths from lung cancer in a sample of 11 countries.

One of the assumptions of the classical normal linear regression model is that the error term included in the regression model follows the normal distribution. This assumption cannot always be maintained in practice. We showed that as long the assumptions of the classical linear regression model (CLRM) hold, and *if the sample* *size is large*, we can still use the *t* and *F* tests of significance even if the error term is not normally distributed.

Finally, we discussed the problem of simultaneity bias which arises if we estimate an equation that is embedded in a system of simultaneous equations by the usual OLS. If we blindly apply OLS in this situation, the OLS estimators are biased as well as inconsistent. There are alternative methods of estimating simultaneous equations, such as the methods of indirect least-squares (ILS) or the two-stage least squares (2SLS). In this chapter we showed how ILS can be used to estimate the consumption expenditure function in the simple Keynesian model of determining aggregate income.