**Chapter 17 Summary**

Panel data regression models are based on panel data, which are observations on the same cross-sectional, or individual, units over several time periods.

Panel data have several advantages over purely cross-sectional or purely time series data. These include: (a) increase in the sample size, (b) study of dynamic changes in cross-sectional units over time, and (c) study of more complicated behavioural models, including study of time-invariant variables.

However, panel models pose several estimation and inference problems, such as heteroscedasticity, autocorrelation, and cross-correlation in cross-sectional units at the same point in time.

The two prominently used methods to deal with one or more of these problems are the fixed effects model (FEM) and the random effects model (REM), also known as the error components model (ECM).

In FEM, the intercept in the regression model is allowed to differ among individuals to reflect the unique feature of individual units. This is done by using dummy variables, provided we take care of the dummy variable trap. The FEM using dummy variables is known as the least squares dummy variable model (LSDV). FEM is appropriate in situations where the individual-specific intercept may be correlated with one or more regressors. A disadvantage of the LSDV is that it consumes a lot of degrees of freedom when *N* (the number of cross-sectional units) is very large.

An alternative to LSDV is to use the within-group (WG) estimator. Here we subtract the (group) mean values of the regressand and regressor from their individual values and run the regression on the mean-corrected variables. Although it is economical in terms of the degrees of freedom, the mean-corrected variables wipe out time-invariant variables (such as gender and race) from the model.

An alternative to FEM is REM. In REM we assume that the intercept value of an individual unit is a random drawing from a much larger population with a constant mean. The individual intercept is then expressed as a deviation from the constant mean value. REM is more economical than FEM in terms of the number of parameters estimated. REM is appropriate in situations where the (random) intercept of each cross-sectional unit is uncorrelated with the regressors. Another advantage of REM is that we can introduce time-invariant regressors. This is not possible in FEM because all such variables are collinear with the subject-specific intercept.

The Hausman test can be used to decide between FEM and ECM.

Some specific problems with panel data model need to be kept in mind. The most serious problem is the problem of attrition, whereby for one reason or another, members of the panel drop out over time so that in the subsequent surveys (i.e. cross-sections) fewer original subjects remain in the panel. Also, over time subjects may refuse or be unwilling to answer some questions.