

Data Exercises

Chapter 9 Diagnostics

This set of questions looks at some diagnostic tests applicable to this model. We postpone the heteroscedasticity tests until we get to the questions for Chapter 11. As this data is not time-series the issue of serial correlation is not relevant.

1. This question is designed to help you look for outliers in your equations. Plot the standardized residuals of the regressions against the dependent variable for each of the three equations you estimated for the Chapter 6 work. You may need to run these regressions again to compare the three graphs of standardized residuals against the dependent variable.

To do this you need to click on the PLOTS option in the Linear Regression command box, paste *ZRESID in the Y box and *DEPENDNT in the X box.

2. Comment on the graphs you just produced.

3. Now go back and do these regressions again and this time click on the Statistics tab in the regression menu and check the box under Residuals which says 'casewise diagnostics'.

4. If you did not change the default value you will now get a table showing which residuals are more than three standard deviations from the mean. Compare the findings in these tables across the 3 equations.

5. Re-estimate the three equations with the outliers listed in the answer to Q.4 omitted.

6. Re-estimate the equations again and this time click on the save tab and check the box to save the unstandardized residuals which will now appear as a new variable at the end of the spreadsheet. Now go to the Analyze menu and choose Descriptive Statistics> Descriptives and choose the residual you have saved, click on the Options tab and check the boxes for Skewness and Kurtosis and continue. You can now put the Kurtosis and Skewness values into the formula for the Jarque-Bera test and test the residuals for normality.

7. Now estimate your three different functional forms (this time with the dummy variables for country of origin included) on sub-samples instead of the whole sample. Do four regressions using observations 1-100, 101-200, 201-300, 301-400. Note that these may not be equal sized, due to different numbers of missing observations in each. Compare the coefficients, t-ratios and R squareds for each functional form across these four groups.