PROJECT TEMPLATE ON DUMMY VARIABLES

(The attached PDF file has better formatting.)

Updated: January 26, 2006

Simple regression analysis uses constant regression coefficients. We use residual plots to test if the inflation rate and loss payment pattern are constant. If a regression coefficient is not constant, we use a dummy variable to differentiate two periods.

We test the constancy of both ß's: inflation rate and payment pattern. We use three data sets: one with a change in the inflation rate, one with a change in the payment pattern, and one with both changes.

Jacob: Does the textbook explain how we see this from the residual plots?

Rachel: This is not clear from the textbook. We guide you to the pages in the CAS paper explaining this procedure, and we provide more explanation in the postings. It takes time to get used to residual plots, and we discuss the issues on the discussion forum.

Jacob: What does the student project include?

Rachel: The student project has three parts:

- Textual explanation of why we perform each technique.
- Graphical analysis of residual plots
- Re-estimation of the regression equation with dummy variables.

Explanation: Suppose the ordinary least squares estimator for the inflation parameter is +15%, with a *p*-value of 2%, and the regression equation has an R^2 of 80%. You must decide whether the true inflation parameter is 15% each year or changes over the years.

- You explain what the *t* test shows. If the true inflation varies over the years (say from 10% to 20%), does this change the result of the *t* test? It doesn't, so the *t* value doesn't help us much.
- You explain what the R^2 shows. If the true inflation varies over the years (say from 10% to 20%), does this change the R^2? It has an effect on the R^2, but the effect may be small. Most of the R^2 relates to other variance, not changes in the inflation rate.

Residual plots: You form residual plots for each independent variable. For each data set, you explain what the residual plots imply. The sample data sets use clear changes in the β values. You must explain the shape of the residual plot for five scenarios:

- constant β (straight horizontal line at $\epsilon = 0$)
- discrete increase or decrease in β (two line segments: a V or an upside down V)
- smooth increase or decrease in β (smoothed V or upside down V)

Jacob: Are these the same type of residual plots we use for the heteroscedasticity project?

Rachel: These plots examine the *slope* of the line connecting the average residuals. You must differentiate between the slope of the residual plot and the value of the residuals. You will find it helpful to discuss these plots on the discussion forum. We give you several examples, and after discussing each one, you will see the effect of a change in the β parameter on the slope of the residual plot.

For each data set, you write the regression equation using an additional dummy variable. The textbook shows several ways to include a dummy variable; we presume only the slope parameter changes. You re-estimate the ordinary least squares estimators and the R^2.

You use three methods to test whether the dummy variable is significant. The best test is again a residual plot: the line should be a flat horizontal line at $\epsilon = 0$. You also examine the *t* statistic for the dummy variable and the *F* statistic for the restricted vs unrestricted regression equations.