Statistics Citadel
Jeffrey R. Russell
Homework 6

1. We are interested in the relationship between the average bid ask spread of stocks and characteristics of the stock. Use the file "Spread" for this problem. The spd variable in the data set is the average bid ask spread over the day expressed as a fraction of the price for a given stock. The explanatory or X variables are:

   Vol: the daily volatility of the asset.

   Trd: average number of trades per day.

   Num: Number of analysists that follow the stock.

   Turn: The fraction of outstanding shares that are traded each day.

   Size: The total dollar value of outstanding shares.

   a. Use a scatter plot to examine the relationship between Spd and size. Does the relationship appear to satisfy the least squares assumptions?

   b. Run a regression of Spd on size (Spd is your "Y" variable). Plot the residuals. How do they look?

   c. Take the log of Spd and the log of size. Now look at the scatter plot. Does this look better?

   d. Regress log(Spd) on log(size). Plot the residuals. How do they look?

   e. Run a regression of log(Spd) on log(Size), log(Trades), log(Numan), log(Turn) and log(sd). Interpret the slope coefficients.

   f. Test the null that all slope coefficients are zero in the regression above.

   g. Next, run a regression of log(Spd) on all the variables except the number of analysts. Compare both the adjusted and unadjusted R-squareds from this regression to the

one in part a. Specifically, look at the change in the R-squared and the change in the unadjusted R-squared when the variable num is dropped from the model.

h. Let's use the model in part g. to predict what the spread will look like on a stock tomorrow. Consider a stock that has log size of 10.5. Suppose that for a given stock we expect that for the following day the log of the turnover to be -1.1, the log of number of trades to be 7.6 and the log of the standard deviation to be -3.5. Predict what the spreads (not log(Spd)) will be for this stock tomorrow.

i. Find a 95% prediction interval for the forecast in part h.

j. Using the full model in part g, test the joint null hypothesis that both log of turnover and the number of analysts have coefficient zero.

k. Test the null hypothesis that log(Size) and log(Trd) have the same slope coefficient.