**Economics 671**
**Problem Set #4:**

(**1**) Load the data set on the course website, cars2.txt. This data set adds another column as the final column in the data set, "German'," which is a binary variable, equal to one if the vehicle is made by a German manufacturer.

Now, go back to PS#2 and recall the description of the car price dataset. Then, estimate the regression:

$$Price_i = \beta_0 + \beta_1 NotReliable_i + \beta_2 RoadScore_i + \beta_3 MPG_i + \beta_4 Zeroto60_i + u_i.$$

(**1a**) This time, elaborate your MATLAB program so that it also reports $t - statistics$, (approximate) 95% confidence intervals for the estimated coefficients and approximate $p$-values. By "approximate" here, I mean that you can assume that $n - k$ is large enough so that normality applies. Comment on the statistical significance of the coefficient estimates. Note that the values you obtain should be very close to those reported in the STATA output.

(**1b**) Using the code you developed for (1a), estimate the following model:

$$\begin{aligned} Price_i &= \beta_0 + \beta_1 NotReliable_i + \beta_2 RoadScore_i + \beta_3 MPG_i + \\ &\quad \beta_4 Zeroto60_i + \beta_5 German_i + \beta_6 (German_i * MPG_i) + \epsilon_i. \end{aligned}$$

Answer the following questions about this regression:

- What has happened to the coefficient on $NotReliable$ relative to the regression in (1a)? What explanation can you offer for this?
- What is the interpretation of the coefficient on $German * MPG$?
- Calculate the average MPG in the sample. Now, consider the parameter $\Delta$:

  $$\Delta \equiv E(Price|\cdot, German) - E(Price|\cdot, NotGerman).$$

  In the above, "$\cdot$" denotes that all other covariates are set equal to sample mean values. Calculate a point estimate and standard error associated with $\Delta$ and comment on the result you obtain.
- Estimate, again, the model in (1b), but this time calculate White or "robust" standard errors, as we discussed in class. Briefly comment on how results change when allowing for heteroscedasticity. (You only need to report a table that calculates the coefficient estimates along with these standard errors and associated $t$-statistics.

(**1c**) Finally, for the model in (1b) (assuming homoscedasticity) test the null hypothesis that

$$H_0 : \beta_5 = \beta_6 = 0.$$

(For reference, note that STATA reports an F-statistic here equal to 6.04). What is the economic meaning of this test (that is, why might it be interesting)? Perform this test in several different ways:

1. First, using the Wald Test (page 33 of Regression #6 notes).
2. Second, using the difference between the restricted and unrestricted least squares estimates (page 15 of Regression #7 notes)
3. Third, using the residual sum of squares method (page 18 of Regression #7 notes)
4. Finally, using the method involving the difference in R-square values between the restricted and unrestricted models.