Project 2

Consider the problem of modeling the price charged for motor transport service. In the early 1980s, several states removed regulatory constraints on the rate charged for intrastate trucking services. (Florida was the first state to embark on a deregulation policy on July 1, 1980.) Prior to this time, the state determined price schedules for motor transport service with review and approval by the Public Service Commission. Once approved, individual carriers were not allowed to deviate from these official rates. **We want to estimate a model of the supply price for predicting future prices.**

The data employed for this purpose (*n* = 134 observations) were obtained from a population of over 27,000 individual shipments in Florida made by major intrastate carriers before and after deregulation. The shipments of interest were made by one particular carrier whose trucks originated from either the city of Jacksonville or Miami. The dependent variable of interest is *y*, **the natural logarithm of the price (measured in 1980 dollars) charged** per ton-mile. The independent variables available for predicting *y* are listed and described below. These data are saved in the **TRUCKING file.**

1. DISTANCE: Miles travelled (in hundreds)
2. WEIGHT: Weight of product shipped (in 1000 pounds)
3. PCTLOAD: Percent of truck load capacity
4. ORIGIN: City of origin (JAX or MIA)
5. MARKET: Size of market destination (LARGE or SMALL)
6. DEREG: Deregulation in effect (YES or NO)
7. PRODUCT: Product classification (100, 150, or 200) – Value roughly corresponds to the value-to-weight ratios of the goods being shipped (more valuable goods are categorized in the higher classification)

Note the first three variables are quantitative in nature, while the last four variables are all qualitative in nature. Of course, these qualitative independent variables will require the creation of the appropriate number of dummy variables: 1 dummy variable for city of origin, 1 for market size, 1 for deregulation, and 2 for product classification.

Import the data from canvas and

1. Use the Variable Screening method you have learned in chapter six and select the three or four important variables (including at least one qualitative variable) for model building.

2. Using the knowledge in chapter 4 and 5, propose at least 3 different models. Build your models as follows:

   Model1: Complete second order model

   Model2: Obtain a model by dropping quadratic terms from model 1

   Model3: Drop some interaction terms from model 2

   Model4: Drop the qualitative variables from model 3.

   Create dummy variables for qualitative variables.

3. Conduct partial F test repeatedly and find a best model from the proposed models.

4. After the final model has been selected, Conduct the Global F test for the adequacy of the model.

5. Perform individual t test if you have doubt with any one of the variables.

6. Report the value of adjusted R-squared and standard error of the model.

7. Depending on your report in step six, explain in your word if the final model is good to use or it may need some improvements.

8. Provide all the R codes you have used.