Udit Lal  
2/24/2024  
ECON 4385

HW#5

1. (You do not need to report anything for this part in the main submission, but you need to provide your data file). The deliverable for this question is the data file you need to upload with your submission.
   1. Ok boss
2. What data series are included in the data set? Why is each series included? Are they transformed, and why?
   1. USREC: NBER based Recession Indicators for the United States from the Period following the Peak through the Trough, +1 or 0, Monthly, Not Seasonally Adjusted.   
      TCU: Capacity Utilization: Total Index, Percent of Capacity, Monthly, Seasonally Adjusted  
      UMCSENT: Consumer sentiment  
      UNRATE:
3. Run a linear probability model (LPM) regression that predicts recessions based on the slope of the yield curve. The code runs this model for you, but you need to report and discuss the results. In particular
   1. Report the coefficient estimates for the slopes of the yield curve.
      1. Does an increase in the 10Y-3month spread significantly increase the probability of recession?
         1. For every unit increase in the 10Y-3month spread the probability of a recession increases by 0.02 units. This is hardly statistically significant only to a significance level of 0.01
      2. Does an increase in the 10year 2year spread significantly increase the probability of recession? You can use one of the templates on the last page if you want some ideas about how to report coefficients. If you are making your own tables, include 3 decimal places. If you are exporting or pasting the results from R, you can use the default number of decimal places, as long as your results are properly formatted and aligned, and the entire table is legible and shows up on the page.
         1. For every unit increase in the 10Y-2YM spread, the probability of a recession increases by 0.039 units. This is statistically significant with a significance of 0.001.
   2. In 1-2 sentences, informally discuss why the LPM model is suboptimal for predicting recessions. You can use a graph, a mathematical argument, or report some of the summary statistics for summary(mydata$predlp) to support your argument. For example, do you get negative probabilities of recessions or probabilities that exceed 1?
      1. The Linear Probability Model or LPM has a very problematic shortcoming. Due to its linearity and the dependant variables being unbounded the model can produce results that are unreasonable or uninterpretable like having a greater than 100% probability or negative probability that a certain outcome will happen.
4. Run a logistical regression that predicts recessions based on the slope of the yield curve. The code runs this model for you, but you need to report and discuss the results. In particular
   1. Report the coefficient estimates for the slopes of the yield curve. Does an increase in the 10Y-3month spread significantly increase the probability of recession? Does an increase in the 10year 2year spread significantly increase the probability of recession? Are the results different from those you got in Question 2?
      1. 10Y2YM: 0.546  
         10Y3MM: 0.314  
         An increase in the 10y-2y spread affects the model more than the 10y-3m spread with the larger coefficient.
   2. In 1-2 sentences, informally discuss why this model may be suboptimal for predicting recessions. For example, are there any instances when the model performs worse than a naïve model? The naïve model performance is given by the 45-degree line. Look at the bottom left corner and the top right corner and check if you’re below the 45 degree line: that’s typically when models then to do bad (underpredict relative to a naïve model if the probability of recession is low/ high) You can (but you do not have to) use graphs to support your argument
5. Run a logistical regression that predicts recessions based on your entire data set: percentage changes in US housing starts, total capacity utilization, consumer sentiment, the unemployment rate, changes in employment, and two yield curve variables: the 10year-2year spread and the 10year-3month spread. You will need to add this part to your code. Feel free to use the posted in class templates. Make sure that you’re not accidentally including predlp and predp\_small in your model. Those are predicted probabilities. You only want to include the actual data.
   1. Report and interpret the coefficient estimates for this model. In particular, discuss  
      TCU: -0.002  
      UMCSENT: -0.076  
      UNRATE: -0.753  
      PAYEMS: -0.554  
      HOUST: -0.048  
      T10Y2YM: -0.389  
      T10Y3MM: 0.861
      1. What factors significantly increase the probability of recession?
         1. TCU, UMCSENT, UNRATE, PAYEMS, HOUST, T10Y2YM
      2. What factors significantly decrease the probability of recession?
         1. The 10-year minus the 3-month treasury spread.
      3. Do any coefficient signs look counterintuitive (positive and significant when you expected them to be negative, negative and significant when you expected them to be positive).
         1. I expected housing and the 10y-2y treasury to be positive as an increase in those variables logically would have the opposite effect.