1. Discuss the steps of conducting Out-of-sample forecast evaluations. Make sure to note on its limitations (hint: what is it really telling us)

OOS stands for Out-of-sample forecasting. By using out-of-sample forecasting we can see how well the model would have forecasted an outcome if you did it at a specific point in time, using only the data available up to that point.

The first step in OOS is splitting the sample into 2 parts. The first part is the in sample data (nested sample) which estimates the model's parameters. The second part is the out of sample (non-nested sample) which is for testing the model's ability to forecast. For example, if your data spans 2000 to 2020, you might use 2000–2015 as the nested sample and 2016–2020 as the non-nested sample.

The second step is estimating the parameters. This involves fitting the model using the in sample. You calculate the coefficients which are the parameters that will be used for forecasting. For example, if it is a regression model, you would get the values of the intercept and slope coefficients.

The third step is forecasting using the out of sample data. We apply the model to the out of sample sample to generate predictions after the models in sample parameter were estimated. These predictions represent what the model would have forecasted based on the data available in the in sample. We can then forecast a specific amount of time.

The fourth step is determining the accuracy of the forecast. This is done by comparing the forecasts from the model to the actual values in the out of sample. Mean Square Forecasting Error (MSFE) is used to quantify the accuracy by averaging the squared differences between the predicted and actual values. A lower MSFE indicates a more accurate model, while a higher MSFE suggests greater forecasting error. Using the MSFE we calculate the R^2 Out of Sample (R^2 OOS) from Campbell 2018. This is done by comparing the MSFE of the model to a baseline, typically the mean forecast model, using the formula below. This provides a measure of the model's relative forecasting performance. It is useful when you think your base model is better than the mean model. If R^2 OOS >1 the the mean model beats your model. If R^2 OOS <1 then your model beats the mean. If R^2 OOS = 1 use the mean model to keep it simple

$$MSFE = \frac{\sum (y_t - \hat{y}_t)^2}{R}$$

$$R^2_{OOS} = \frac{MSFE_{model}}{MSFE_{mean}}$$

Finally, to compare different models, the Diebold-Mariano test can be used. This test evaluates whether there is a statistically significant difference in forecast errors between two competing models. Essentially, testing if one model is better than another.

While out of sample forecasting is good for evaluating a model's predictive performance, it has some limitations. One key limitation is its reliance on the assumption that the relationships observed in the nested sample will hold true in the non-nested sample. It assumes the patterns in the in sample will stay the same in the out of sample. This can be a problem if there are big changes, like during COVID when economies and markets shifted suddenly.Also, the results can depend on how the data is split, which might make the model seem better or worse than it really is. Out-of-sample forecasts show how well the model works with past data but may not always predict what will actually happen in real life. There is also limitations in testing which model is better. The Diebold-Mariano test helps compare models but only shows which one performs better. It doesn't mean that either model is good for predicting.

2. Discuss the limitations of the linear probability model.

The linear probability model is a way to estimate probabilities for limited dependent variables. Limited dependent variables is when the dependent variable is a categorical variable such a a dummy variable where 1 = success and 0 = failure. A linear probability model's fitted values represent the expected probability of success. For example, it can be applied in contexts like betting or elections. The linear probability model does have limitations, however.

A major problem with the linear probability model is that it can predict probabilities outside the range of 0 and 1, which doesn't make sense since probabilities must be between 0 and 1. This shows that the model isn't always a good fit for binary data. For example, in the exam I got a probability of winning to be roughly -1. This does not make sense in the context of problems because you can't have a negative probability for winning. You either do not win = 0 or win = 1. You can not lose so badly that the probability is negative.

Another problem with linear probability modeling is that it assumes the effect of the independent variables on the probability, or the marginal effect, is always the same. In reality, this relationship is often nonlinear. For example, small changes in a variable can have a bigger impact when the probability is around 0.5 compared to when it's near 1. For every observation, the marginal effect is different. This makes LPM less accurate. Some solutions for this issue include to estimate for everything and graphing it out, however this can be tedious. Or you could multiply the beta 1 hat by 0.25 to give an idea of what your estimated marginal effect would be if it was a 50/50 shot. Or you could also take the average marginal effect, which may also not be accurate for all observations.

Finally, while LPM is good for its simplicity for limited dependent variables. However, it is less flexible than logit or probit models. These models keep the probability between 0 and 1 which make more contextual sense and good for nonlinear relationships, making them better for real-world use.