Time Series Modeling and Forecasting **Midterm Exam**

The file “fl and us monthly data.csv" contains nine monthly non-seasonally adjusted time series related to US and Florida economic performance. Variable definitions are:

- year
- month
- fl_amsparksarcadesemp_m: Florida employment in amusement parks and arcades.
- fl_nonfarmemp_m: Florida nonfarm employment.
- fl_laborforce_m Florida labor force (employed plus those seeking work).
- fl_unemprate_m: Florida unemployment rate.
- fl_bldpmt_m: Florida housing units authorized by building permit.
- us_civillianemptopopratio_m: The US employment to population ratio for all civilians.
- us_emptopopratio25to54_m: The US employment to population ratio for prime age workers (25-54).
- us_bldpmt_m: US housing units authorized by building permit.
- us_unemprate_m: US unemployment rate.

For all questions, use the logarithm of the variables.

Part A: Time Series Basics and Static Models

For part A, you need only be prepared to answer related questions during the in class exam administration, for which I will provide relevant output.

1) Estimate a static regression model relating employment in amusement parks and arcades to nonfarm employment in Florida without month indicators or a time trend. Then estimate the model with month indicators and a time trend. Explain the difference in the results.

2) Produce and examine the partial autocorrelograms for employment in amusement parks and arcades and for nonfarm employment. Also conduct Dickey-Fuller tests for non-stationarity for both variables. Interpret the partial autocorrelograms and Dickey-Fuller test results.

3) Estimate a static regression model relating the first difference of employment in amusement parks and arcades to the first difference in nonfarm employment in Florida including month indicators and a time trend. Explain the difference in the results you get and the results from #1.

4) Obtain and interpret the results of a Breusch-Godfrey test for autocorrelation for the model from #3.

5) Estimate the model from #3 using Newey-West standard errors and compare the results to #3.

6) How many lags did you use for the Newey-West errors in #5? Why? What are Newey-West errors for? How do they work?

Part B: Dynamic Models and Forecasting

For part B, you need to be prepared to answer related questions during the in class exam administration, for which I will provide relevant output. In addition, you should prepare and submit a report on your work according to the instructions below.

7) Estimate an autoregressive or autoregressive distributed lag model appropriate for making a one period ahead forecast of Florida employment in amusement parks and arcades. You will have to decide which, if any, variables other than lags of the dependent variable to include, and also how many lags to include for the dependent variable and for any other variables. Beware, you will probably not want to use most of the variables—you just have to figure out which, if any, can improve the forecast. Justify your choices. Your justification should include some or all of the following: solid reasoning, patterns visible in correlograms, joint hypothesis tests about groups of potential predictors, cross-validation of models, and out of sample fit (for which you need an appropriate test set different than the training and validation sets).

8) For the model you consider best, generate a variable equal to the predicted value and also the approximate upper and lower bounds of a 95% forecast interval. From these, generate two time series plots: one of the actual and predicted values (including the forecast value for January 2018) and one of the actual values and the forecast interval. Discuss the usefulness of the forecast. Do the movements in the forecast values correctly predict movements in the first difference? Does the forecast interval almost always bound the actual outcome? Is the forecast error typically small enough that the forecast might be useful for planning purposes (say for hotels that cater to amusement park visitors or an official wanting to predict tax collections)?

Report: Write up a forecast report justifying your model choice, presenting your results, and discussing its usefulness.

- Be concise but thorough.
- Write like you are preparing a report for a client that does not know the technical aspects of forecasting and does not have time to read a 20 page paper, but include enough detail to fully support your assertions to those with technical training that may see the report.
- Do not simply post screenshots of the Stata results window. Rather, provide neat professional looking tables of the results you obtain.
- Make sure any results you refer to in written answers to questions appear near the written answers so that your overall submission is easy to make sense of.
- Label and number tables and figures in a self-explanatory way.
- Use a 12 point serif font (for ease of reading).
- Use double spacing (for me to write comments easily when grading).
- Include your clean do file and log file as appendices (only for what is relevant to part B).
- Submit a copy through canvas by 5:45 pm Thursday March 15th.
- Bring a printed copy to class for me to grade.