

## Midterm Exam – March 2, 2021

Instructions

- i. Answer all 10 questions.
- ii. You may type your work or write it all by hand, or a combination. If you write it all by hand, make sure to write neatly. I prefer you type your explanations, but if you type a lot slower than you can write neatly, don't worry about it. If you decide to type most of it, feel free to write the equations by hand—only type them if you are a master and can type them as fast as you can write them neatly.
- iii. Show and explain your work where relevant. A correct final answer, without explanation, is worth little. A good explanation and mostly correct work, with a wrong final answer due to a small mistake, is worth a lot.
- iv. Be neat and concise.
- v. List and explain any assumptions you make.
- vi. You may use any resources available other than speaking or otherwise communicating with anyone else.
- vii. Keep your camera, microphone, and speakers on for proctoring and for test related communications.
- viii. Upload an electronic copy of your work to canvas by 7:00 PM.

**Background**

You are interested in determinants of housing prices in Florida. You have the following three hypotheses:

- i) Building permits mean an anticipated increase in future supply. So when building permits increase, you expect a decrease in price shortly thereafter.
- ii) The cost of borrowing to build or buy a home depends on the real interest rate, Interest. This is the difference between the interest rate and expected inflation. So, when interest rates increase, you expect housing prices to fall shortly thereafter. are higher at time  $t-1$ , you expect housing prices to be lower at time  $t$ .
- iii) When inflation is expected to be higher, home prices go up accordingly. So, when inflation expectations increase, you expect an increase in home prices shortly thereafter.

You have monthly data on the median per square foot price of homes listed for sale in Florida since July of 2016. You plan to ignore data after December of 2019 for this purpose. Assume, for purposes of this question, that gathering more data from earlier periods is not possible. (In reality, it exists, it is just not readily available for free and on FRED.) You will need to keep in mind the limitations imposed on model complexity by the amount of data. In particular, you do not have enough data to estimate models with many lags for multiple variables while keeping a reasonable number of degrees of freedom.

Given the limited data, you have in mind the following causal model:

$$\ln List_t = \beta_0 + \beta_p \ln Permits_{t-1} + \beta_{Int} \ln Interest_{t-1} + \beta_{Inf} \ln Inflation_{t-1} + r_t.$$

A Stata log file, augmented with several graphs, accompanies this exam.

**Questions**

- 1) Express the model above in first differences. Under what conditions would you need to work with the differenced model instead of the original?
- 2) Suppose you think the residual,  $r_t$ , follows an AR(1) process with parameter  $\rho$ . Write the dynamically complete version of the original model **and** of the model in first differences.

The remaining questions refer to the Stata do file and (augmented) log file provided. The file has page numbers, and line numbers that start at 1 on each page.

- 3) What is the purpose of the commands on page 1 lines 25-32?
- 4) What is the purpose of the commands and results from page 2 line 4 through page 3 line 30, and what conclusion should be drawn from the results of these commands?
- 5) Four sets of models are estimates. What are the differences between the sets (**not** between the models in a given set)? Which set is better for the purpose at hand? Why?
- 6) There are three models within the set you chose. Each of those is estimated twice. What is the difference between the two sets of estimates? Does the difference matter? Why? Which is better? Why?
- 7) For the set you chose as best, interpret the F-test for the first model in that set. That is, if set X is best, interpret the F-test that follows one of the two estimates of Model X.1. Again, there are two versions. Use the better one. Your answer to 6 should have made it clear what the difference is, which is better, and why.
- 8) How do the three models in your chosen set relate to the model set out on the previous page and to questions 1 and 2?
- 9) What assumption must be defended to apply a causal interpretation to the results of this model, as opposed to a purely predictive one?
- 10) Within the set of models you chose as best, X, which model is best for predicting *List*? That is, X.1, X.2, or X.3? Why? Which is best for testing they hypotheses of interest? Why? If the two are different, why?