

Empirical Assignment Brief

Assignment Overview

The goal of this assignment is to use the tools you have learned so far in your R assignments and apply them to an independent project on time series data of your choice. I will be providing a few sample datasets that are easy for you to use from which you can choose which one relates to a research question you find interesting.

- You cannot use exactly the same data as I use as an example in tutorials! I will primarily be using UK quarterly GDP growth as my Y_t variable and UK quarterly unemployment as my X_t .
- Remember that you can always subtract one time series from another if you are interested in the difference between two outcomes. For example, we considered the term spread, the difference between long and short run interest rates, in some of our R assignments as a predictor of GDP growth. You can also use this as an outcome, and look at the difference between profits in two different sectors as your Y_t or X_t or differences in outcomes for men and women as your Y_t or X_t , etc.
- You are welcome to seek out your own data and explore an independent research project if you wish to go above and beyond the assignment. You will, however, need to complete the same analysis tasks listed in the assignment. The grading scheme will be consistent for those using data I provide and for those who find their own.
- If you want to use this empirical work as the basis for your dissertation that would be an excellent use of your effort. You should be aware, however, that you cannot submit the exact same report for your dissertation as you submit for this module, and your dissertation would need to contain substantively additional content.

1

The first task is choosing an outcome variable that will be your Y_t for your analysis, and a primary X_t that will be the main explanatory variable you explore. Once you have chosen some data of interest, the first part of this

assignment will involve using the tools we learned in the first part of the module (up to our work with Dynamic Causal Effects) in order to explore what we can learn about your outcome Y_t as an Autoregressive process. You will complete the analytical tasks outlined below by adapting the code provided in R tutorials and write up an explanation of the task and the results. You will also use the tools of Volatility Analysis we will cover later in the course to test whether the volatility or variance of a time series is serially correlated.

The next step is to consider an additional explanatory variable, and estimate

the Dynamic Causal Effects of this explanatory variable on your outcome of interest. You will complete the analytical tasks outlined below by adapting the code provided in R tutorials and write up an explanation of the task and the results. Where we have learned a manual tool to complete a task, you should use this in your assignment. You are, however, free to use the automatic tools to check your work.

You will then test two variables for Cointegration, in a formal test of whether they move together and receive the same shocks. This can be the same as the variables you have used previously, but you can also choose different variables.

Finally, you will estimate a model testing for Volatility Clustering in your time series Y_t using the Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models.

Grading Criteria

Your assignment will be assigned a grade as follows:

- (Weight: 50%) Interpretation and Understanding of Econometric Tools Part of your grade will be based on whether you correctly use and interpret the tools of Time Series Econometrics that we learned. This means that you use the appropriate models for the given task, that you interpret results correctly, using the proper critical values for inference as well as interpreting null hypotheses correctly. This also depends on whether you explain why you use different tools, and the problems these are selected to deal with.

- (Weight: 25%) Programming and R Code Part of your grade will depend on correctly using R to implement the tasks you are assigned and whether your R code correctly implements the work that you describe in the

write-up of your assignment. Marks will be given for R code that is correct, and with comments to clarify you understand the tools you are using.

- (Weight: 25%) Economic Analysis and Discussion This part of your grade will depend on the economic analysis of your results and the depth of your

discussion. Marks will be given for the economic content of your analysis and your interpretation of the economic reasoning of your results.

Assignment Outputs to Submit

1 1.1

•

A PDF write-up of the results of your analysis, including graphs and tables. See the outline of the analysis tasks to complete below for details on exactly what tables & graphs you need to complete. Include your R code at the end of the document. You can write this document in Word or you can use LATEX. There are additional videos on the assignment page about exporting tables and using LATEX.

Word Count: Maximum 2,500 words, excluding R code, tables and figures.

Analysis Tasks to Complete Descriptive Analysis – Week 1 Exercises

Before running regressions, we will first examine our data and use some simple tools to look at the time series.

1.1.1 Data Description

First, write a very brief (just a few sentences) description of the outcome variable you are interested in analysing. Next write a brief description of your primary explanatory variable, and the rough research question.

1.1.2 Time Series Plots

Next, plot your Y_t time series., and give a few sentences of description. Do there appear to be significant outliers in this time series? Either exclude them if they are near the beginning or end of your time series, and if not, make note of this outlier and make sure to take it into account in your analysis.

3

1.2 1.2.1

•

•

Autoregression Analysis of a Time Series Estimate an Autoregression Model

First, run an AR(1) regression of your outcome variable. Then use the Bayes Information Criterion to select the appropriate lag length for your model, setting a maximum of four lags. Write down the four values of the $BIC(p)$ you calculate, and explain which model length you end up selecting. Now, estimate this model. (See Week 1 & 2 Exercises)

Next, test for violations of our key Time Series Assumptions: (See Week 3 Exercises)

- – Use the appropriate model to test for a unit root process. Does economic theory suggest that your time series should exhibit a roughly linear time trend? Justify your answer briefly, and explain what this means for the model you use for this test and the hypotheses you

test. Write a brief explanation of the result, and what this means for your time series. If you conclude your time series has a unit root, perform the necessary transformation, use the BIC again on your transformed time series and add this revised AR(p) model to your table.

- – Use the appropriate test for a break in your time series where you don't know the exact date of the break. Write a brief explanation of the result, and what this means for your time series. If you conclude your time series has a break and you identify the likely break date, make the necessary adjustment to your model and add this model to your table.

Report estimated coefficients from both the AR(1) and AR(p) models in a table, along with the coefficients from your modified model in the case that your time series either has a break or a unit root.

Is the coefficient on Y_{t-1} in your AR(1) significant? Write a brief explanation of whether it is statistically significant, and an additional brief interpretation of the economics of this result. How about the coefficient Y_{t-1} in your AR(p) model - is it similar? Discuss the implication of these results, and the persistence of shocks. If you correct for a trend or a break, discuss how your analysis of the non-transformed time series might be misleading.

-
-

1.2.2 Estimate an Autoregressive Distributed Lag Model - Week 2

Exercises

- Now we are going to introduce a second variable X_t . First, estimate an ADL(1,1) model. If you found your Y_t had a unit root above, then continue

4
to use the transformed data in this section.

- Repeat the exercise you conducted above using the BIC to select the length of lag, but now you will select a lag p to use for your ADL(p,p) model. For simplicity, consider again up to $p = 4$ and use the same lag length for Y_t and X_t .
- Use a Granger causality test to test whether the lags of your explanatory variable X_t are jointly significant predictors of Y_t . Report the test statistic in the text (no need to add it to a table).
- Produce a table with your coefficient estimates from the ADL(1,1) model as well as the ADL(p,p) model.
- Interpret the results from the above. Are the lags of X_t jointly

predictive of Y_t in a model where we also include lags of Y_t ? Discuss the economic significance of this result.

1.2.3 Check Out-Of-Sample Forecast Performance – Week 4 Exercises

- Using the Pseudo Out-Of-Sample forecasting method, with your ADL(1,1) model and with the final 25% of your sample as your excluded sample, and compare the within-sample SER (from the regression including none of your excluded observations) and the out-of-sample fit using your estimate of the Root Mean Squared Forecast Error.
- Compare the size of the SER to the size of your RMSFE. Which is larger? Does this suggest your forecast errors are larger, smaller, or the same as your within-sample errors? Is your model capable of predicting out-of-sample?

1.3 Dynamic Causal Effects – Week 5 Exercises

- Use GLS to estimate the dynamic multipliers for a distributed lag model regressing Y_t on X_t and lags. For simplicity, use a Distributed Lag model where $r = 3$, which means you will include X_t as well as the lag X_{t-1} and X_{t-2} , and an AR(1) error term, meaning that you model the error term just as in lecture using ϕ_1 . Now, estimate these dynamic multipliers using the Cochrane-Orcutt method (not the Iterated Cochrane-Orcutt!).
- Discuss the results above, beginning with a short discussion of whether it is reasonable to assume strict exogeneity or exogeneity and give an example of something that would mean we can only assume exogeneity but not strict exogeneity. For example, in lecture we considered crop prices as our outcome

5

1.4

Y_t and climate shocks as our X_t . If people potentially stockpile crops today based on anticipated climate shocks tomorrow then this would violate strict exogeneity. Give an example of the issues with assuming strict exogeneity in your setting. The important thing here is showing you understand the conditions, so you can use a slightly unrealistic example here, as long as you show you understand how to think of the exogeneity conditions in your context. Next, discuss the implications of the dynamic multipliers you estimate. Which dynamic multiplier of X_t is strongest? Does the effect increase, decrease, or stay the same over time?

Multiperiod Forecasts - Week 7 Exercises

Using both methods that we have learned for multiperiod forecasting, forecast the next ten periods of your time series past the end of your data using your ADL(p,p) model above. What is the forecasted value in ten periods time? How do the two forecasts differ?

1.5

-
-

1.6

-
-

Cointegration – Week 8 Exercises

Use the two-stage test for cointegration to test if your Y_t and X_t are cointegrated. You can use the same X_t as above, or you can choose a different X_t if you think they are more likely to be cointegrated and, therefore, a more interesting exercise. (Note: even if it isn't sensible to test for cointegration here, conduct the test anyways, and interpret the result accordingly, and explain why it isn't appropriate to test for cointegration of these time series)

Discuss the results from the above analysis. First of all, discuss whether it makes sense in this case to test for cointegration.

Volatility Clustering Analysis – Week 9 Exercises

Next, analyse whether the volatility of your time series is clustered, that is, whether your time series exhibits greater variance at some times than others. Estimate an ARMA(1,1)-GARCH(1,1) model on your data.

Report estimated coefficients from this model in a table. Are any of the coefficients significant? What does this mean about whether your time series display conditional heteroskedasticity? Give a very brief interpretation of the economics of this result: will a period of high volatility tend to be long-lived, or will it be brief?

6

1.7 Conclusion

Finally, write a brief paragraph summarizing your findings. Again, this should just be a brief summary of any of the results that give additional economic insights relating to your outcome variable Y_t or the relationship between Y_t and X_t .

7