MScFE 610 FINANCIAL ECONOMETRICS

Group Work Project # 2

See grading rubric here.

Scenario

The team is enjoying a lot of success. Other groups want to understand the concepts that were richly explained in the best-practices handbook. The handbook did an excellent job describing, illustrating, and attempting to resolve challenges of time series models. In this group work project, we shift focus to the actual models. Here are **some important challenges encountered when using time series models**:

1. **Feature extraction**. For those students not yet familiar with the ML technique of feature extraction, it means selecting certain features of the data or constructing new ones so that they retain the relevant information of the original data, but are more efficient and require less computational effort than the original dataset which could be very large in certain applications. If you select this challenge, make sure that you explain clearly how you did the "extraction" and convince the reader that you were able to retain key elements of the original dataset.
2. **Modeling non-stationarity and finding an equilibrium.** Note the two tasks: detecting non stationarity and modeling it. In particular, the concept of equilibrium must be researched.
3. **Handling Multicollinearity.** This is the most traditional of the 4 challenges assigned. So, make sure that you explain it well and illustrate it with a valid example.
4. **Detecting a regime change.** If you select this challenge, make sure that you can explain the issue, how regime switches can be detected, and how you would model your data afterwards.

Tasks

Your team will:

1. Write the model's equations and define the terms (Please do not add pictures of formulae which look sloppy and not professional. Use either LaTeX, Word or equivalent. If you decide to use Word that is fine, but then make sure that the

Equation Editor is used, subscripts and superscripts are properly positioned etc. This part must be accurate, easily readable and elegant.)

2. Write the code to import, structure, and model the data
3. Calibrate the parameters of the model
4. Interpret the parameters
5. Write a report in PDF explaining all of the above

Step 1

As a group of 3 your team **will select 3 of the 4 challenges (one must be either non-stationarity or regime changes) and complete all the tasks listed above.**

As a group of 2 or 1, instead, your team **will select 2 of the 4 challenges** (again, one must be chosen between non stationarity and finding an equilibrium model or regime switches)

Depending on which challenges you selected, you will find a dataset that illustrates:

1. Non-stationarity in the case of regression models (spurious regressions, cointegration, equilibrium models, etc.);
2. Undergoing a regime change;
3. Multicollinearity;
4. Features extraction.

Step 2

Individually:

1. Each team member is responsible for coming up a with specific example and associated dataset.
2. Each team member will also identify the reasons this particular dataset works well for the chosen model. For example, if the non-stationarity model were chosen, then suitable econometric tests to show that the series is non-stationary must be implemented.
3. In this GWP there should be no need to use simulated data as there are plenty of economic and financial time series that exhibit either non-stationarity or regime changes. For instance, when investigating regime changes one could look around the timelines of big world events (before and after)
4. **NOTE!** If you choose modeling non stationarity as your item, be sure that you **DO NOT just limit yourself to running unit root tests**.. The task here is mostly about

how to model time series in the presence of non-stationarity (i.e. cointegration, VAR model, ECM model.)

Step 3

As a group:

- Team members will select what they think is the best example (among the 3 or 2 examples supplied by each student) for each challenge they worked on. Team members will indicate why a particular dataset was chosen.
- Team members will write a 5-6 page paper that follows the same categories:
    - **Definition:** Technical definition using formulas or equations
    - **Description:** Written explanation (1–2 sentences)

  The remaining sections will all use the specific dataset chosen. For example, if you chose the daily return of Apple stock, then each of the following sections will include data, graphs, or statistical analyses using Apple's returns.
    - **Demonstration:** Prepare the data for the model; run the model; and show its output. Be sure to list any parameters that were calibrated by the model and interpret them.
    - **Diagram:** Include any exploratory plots or charts that show how processing works.
    - **Diagnosis:** Show diagnostic plots.
    - **Damage:** Identify problems that the model reveals about the data. Also, determine if any of the 6 or more challenges your group discussed can be used to assess the quality of how the model fits the data.
    - **Directions:** Even though we ran a model, it may not fit well. Should we try to fit the same model by manipulating data (e.g., removing outliers, shortening the time horizon, etc.)?
    - **Deployment:** Describe in detail how you would use the model.

Submission Requirements and Format

One team member submits the following on behalf of the entire group:

1) A **zipped folder** including:
    a) An executable Jupyter notebook* that addresses all the challenges

b) A duplicate version of the Jupyter notebook code and output in PDF or HTML format.
    i) In order to include the output of the code, you must RUN the code before downloading the PDF.

2) ****1 PDF document** with all sections EXCEPT ANY CODE. This PDF should just contain text, formulas, and graphs, but no Python or other code. In fact, this document should contain all the sections EXCEPT any parts or sections that have code. Please be sure that code only appears in the two files above.
    a) Use the available Report Template and fill out the required information in the first page.

    b) READ the **Academic Policy on the Use of AI**: **in-text citations and references are mandatory**, use **this link** to learn how to add them.

****Use Google Colab or GitHub to collaborate** in completing the executable Python program.*

*****The PDF file with the answers to the questions must be uploaded **separately** from the zipped folder that includes any other types of files. This allows Turnitin to generate a similarity report.*

Rubric

Your instructor will evaluate your group submission for GWP3 using the following rubric:

Quantitative Analysis (Open-Ended Questions) 40 Points	Technical and Non-Technical Reports 30 Points	Writing and Formatting 20 Points
The group is able to apply results, formulas, and their knowledge of theory to real-life finance scenarios by doing the following: ● Providing all the necessary information to support their arguments. ● Presenting arguments that reflect group discussion and research. ● Using authoritative references to support a position and provide updated information. ● Concluding with practical takeaways for more insightful financial decision-making.	Technical Reports contain 3 parts: 1) **code** for each question (be sure to explicitly state the question number), 2) the corresponding output of that **code**, and 3) **interpretations** and/or **recommended courses of action** that reasonably follow from those results. **Note**: Technical reports will include the technicalities of models, such as names, methods of estimation, parameter values, etc., and exclude generalities about the work done. It should NOT include names of Python code that were used.	A submission that looks professional should: ● Include the axes, labels, and scales in graphs. ● Be free of significant grammatical errors or typos. ● Be an organized, well-structured, and easy-to-read document. ● Include proper citations and a bibliography in MLA format.
	Non-Technical Reports contain 3 parts: 1) clear explanation of results; 2) the recommended course of action that follows; and 3) the identification of factors that impact each portfolio. **Note**: AVOID all references to model names, algorithms, and unnecessary details. Instead, focus on the investment decision.	

Revision Date: February 18, 2025