DOWN BY THE WATER

Assignment
Multivariate Econometrics
VU MEOR
2024

"We know we'll run out of dead dinosaurs to mine for fuel and have to use sustainable energy eventually, so why not go renewable now and avoid increasing risk of climate catastrophe? Betting that science is wrong and oil companies are right is the dumbest experiment in history by far."

Elon Musk, Founder, The Boring Company

VU UNIVERSITY AMSTERDAM

Climate change stands as one of the most critical challenges of our time, with CO_2 emissions playing a central role in its acceleration. In this regard, determining the most effective approach for lowering CO_2 emissions and the collateral risk of catastrophic natural disasters is crucial. Hydropower, as one of the oldest and most widely used renewable energy sources, plays a crucial role in efforts to reduce these emissions. A recent publication, Perone (2024), uses a panel data approach to investigate the long-run relationship between disaggregated renewable energy production, including hydropower and CO_2 emissions per capita for a panel of 27 OECD countries from 1965 to 2020 and finds very interesting results.

In this assignment, you will investigate the dynamic relationship between hydropower consumption and CO_2 emissions along with some other variables, using time series techniques and compare your results with the panel approach of Perone (2024). The aim is to deepen our understanding of how hydropower can contribute to a more sustainable future and to compare the results of time series approaches with those obtained from panel data analyses.

Detailed instructions

1. This assignment is mandatory and to be solved by groups of (at least) three or (at most) four students.
2. **The deadline is 6^{th} of December 2024, 23:59.**
3. You are free to use R, Python or Matlab. You are allowed to use existing packages. You need to add comments to the code to clarify the function of each operation in the code.
4. You are allowed to use LLMs such as GPT to help you with coding. However, **you are not allowed to use GPT generated text in your assignment. If detected, these cases will be send to the examination board.** You must use your own words and sentences in the report you will submit.
5. The assignment should be **maximum 25 pages**, including tables, figures; excluding references.
6. In your report, you should explain how the results are obtained, why a specific choice (model, variables, long-run relationships, ...) is made and give an interpretation of the results.
7. A file must be uploaded in pdf format to Canvas before the deadline mentioned above. You must upload also all the R/Python/Matlab codes used to obtain the answers. The code should be clear and well commented.
8. The pdf document must have the group number, the names and student numbers of each member of the group.
9. You must include a short report explaining the contribution of each group member to the writing process of the assignment.

The Data

You will use the data that is made available by Perone (2024). The focus will be on Germany and France. You can find the description of the variables and data sources in Perone (2024). Here are the details of the variables that the data set contains:

Emission: The average carbon dioxide emissions per capita in metric tons

Pop_density: The number of people per square kilometer of land area

Agri_percent_land: Agricultural land area as a percentage of total land area in the country

Cattleperkm: The number of cattle per square kilometer

kWh_percapita: The average electric power consumption per capita expressed in KWh

Fertilizer: The sum of synthetic nitrogen, potassium, and phosphorous inputs, as well as organic nitrogen inputs, expressed in kilograms per capita

Hydropower: Hydropower electricity generation in terawatt hours

The Assignment

The assignment consists of seven parts. Below you will find a guideline to complete each part of the assignment. You should address at least all the items that are presented throughout the guideline. Discussions on issues/topics other than the ones mentioned in the items below for each part are also welcomed. For each part please write your detailed comments and interpretations of the results you obtain.

1. Literature review

Find other academic articles that investigate the relation between hydropower and CO_2 emissions. Use your findings in preparing a thorough literature review on the topic. Which econometric approaches are used? What are their findings? Do you find any issues with the approaches used? etc.

2. Graphical analysis of the data

The data set contains data on 7 variables and you need to perform time series analysis Germany and France. This means you need to analyze 14 time series in this section. Obtain the summary statistics for the data, then perform a graphical analysis by plotting the data. Using these plots and summary statistics, you should already be able to address, at least partly, issues such as

(i) Do you see any issues with the data? Missing values, strange changes etc. If yes, solve these issues.

(ii) Do you see any evidence in favor or against the assumption of covariance stationarity? Do you suspect them to be I(2), I(1), I(0)?

(iii) Do you see any evidence in favor or against for the presence of deterministic components such as a constant or a linear trend?

(iv) Decide whether you want to use logarithms of any of the variables?

3. Analysis of the order of integration

Consider now the issue of testing formally for the presence of unit roots in your series. Carefully design and apply a sequence of tests that enables you coherently address the issue of I(2) vs I(1) vs I(0). You may start with simple Dickey Fuller tests but there is no need to limit yourself to DF tests. On the contrary, you should look for some other tests.

(i) Discuss carefully your choice of the deterministic components.

(ii) Discuss the possible evidence of serial correlation in the residuals of your Dickey Fuller regression.

(iii) Taking into account the presence of possible serial correlation, consider various extensions such as Augmented Dickey Fuller test, Phillips-Perron tests.

(iv) Use some other tests that are robust to, for example, structural breaks. You need to find the test from the literature yourself.

(v) Present the results of the various tests and compare these results.

4. Cointegration analysis

Once the integration of the order of the data has been determined, you need to carry out a cointegration analysis.

(i) First briefly discuss the type of cointegrating relationship(s) you might expect by using some previous literature findings.

(ii) Assuming that your series are I(1), test for (no)-cointegration using the techniques discussed during the course and some techniques you will find in the related literature. Compute and compare the results of various tests for cointegration. You may use the standard residual based cointegration tests using a static regression (Engle and Granger approach, ADF type tests, Phillips-Ouliaris' test) and the Maximum Likelihood based tests (Johansen Trace and Maximum Eigenvalue tests).

(iii) Is there evidence in favor or against cointegration? Is there any evidence in favor of more than one cointegrating vector? If yes, what identification scheme one should adopt and why?

(iv) Assuming there is a single cointegrating vector, estimate the cointegrating regression by using various approaches discussed in the lectures, such as static least squares, DOLS, FMOLS and ECM. Report and discuss your estimation and inference results (coefficient estimates, test statistics etc.) Discuss the differences between the assumptions of these methods.

(v) Adopt a systems approach and use Johansen's analysis to estimate the cointegrated system. Report your estimation and inference results (coefficient estimates, test statistics, hypothesis tests, test results etc.) Compare your results with single equation approaches. Discuss the differences between the assumptions of these methods.

(vi) Are the results in accordance with your prior expectations? How should we interpret these results?

(vii) Are the long run relations you discover apparently constant through time? If not, do you think whether the interventions of policy makers affect these relations?

5. Causal analysis

Perform a Granger causality analysis. Use suitable Granger causality tests to investigate the Granger causal relations between variables. Interpret your results.

6. Comparison with Perone (2024)

Compare your results with Perone (2024). Comment on the differences. Explain which set of results policy makers should rely on when making policy decisions.

7. Ethics (Maximum 1 page)

Reflect on the ethical aspects of performing data analysis on the relationship between hydropower consumption and CO_2 emissions. In your discussion, address the following points:

(i) Discuss the ethical importance of ensuring the accuracy and reliability of the data you use. How can data quality issues, such as errors, omissions, or outdated information,

affect your analysis and conclusions? What steps should you take to verify and validate your data sources?

(ii) Propose ways to incorporate ethical considerations into your econometric modeling and policy recommendations regarding renewable energy adoption.

Scoring Guide

1. The main purpose of this assignment is to gain some understanding and the ability to analyse time series data. By doing this correctly and satisfactorily you will earn 80 of the 100 points. The division of these 80 points among the parts of the assignment is as follows:

(i) literature review and initial analysis of the data is 10 points.

(ii) analysis of the order of integration is 20 points.

(iii) cointegration analysis and causal analysis is 30 points.

(iv) comparison with the reference paper 15 points.

(v) discussion on ethics is 5 points.

2. The second aim of this assignment is to provide you an opportunity to gain more experience in writing well structured reports. The third aim of this assignment is to give you a chance to work within teams and to develop your communications skills. The structure and the quality of the report that is produced by an effective division of labor will be evaluated separately. This is worth 10/100.

3. You are free and encouraged to use econometric methods that are not discussed during the lectures of this course. This will give you the opportunity to develop your self learning skills. This is worth 10/100.