FIXING A HOLE WHERE THE CO2 EMISSIONS GET IN

Assignment
Multivariate Econometrics
VU Econometrics
Fall 2022

"Climate change will continue for many decades to come. The scale of future climate change and its impacts will depend on the effectiveness of implementing our global agreements to cut greenhouse gas emissions, but also ensuring that we have the right adaptation strategies and policies in place to reduce the risks from current and projected climate extremes."

Hans Bruyninckx, EEA Executive Director

"We are the first generation to feel the effect of climate change and the last generation who can do something about it."

Barack Obama, Former US President

VU UNIVERSITY AMSTERDAM

The path that took 195 states to sign the Paris Agreement was constructed by scientific research that give answers to questions such as: What will be the economic effects of limiting the increase to 1.5°C? What will happen if the increase goes above 2°C? What are the least costly ways to reduce emissions? Which policy tools governments can use to ensure less greenhouse gas emissions? Using accurate information generated by scientific research is crucial for informed and effective mitigation policies. According to the findings of Burke et. al. (2015), if the economic systems behave in a way that they have been behaving until now, unmitigated warming is expected to reshape the global economy by reducing average global incomes roughly 23% by 2100 and widening global income inequality, relative to scenarios without climate change.

In this assignment you will first improve your understanding of time-series by working with simulated data. Then, use econometric methods to improve our understanding of the relation between climate change and economic production. You will use novel modelling approaches to uncover empirical evidence on the nature and co-evolution of climate/economic variables.

Your task is to investigate time series properties of climate variables such as temperatures, precipitation, radiation, and economic variables such as GDP per capita, crop production index, agricultural land in use, agricultural GDP; the relationships between them; whether policies, agreements, important technological developments have any effect on climate change and on the relation between climate and economic variables.

Detailed instructions

* This assignment is mandatory and to be solved by groups of (at least) three or (at most) four students.

* **The deadline is 9th of December 2022, 23:59.**

* You are free to use one of R, Python, Matlab. If there are existing packages to use, you are allowed to use them. But in that case, you need to add comments to the code to clarify the function of each operation in the code.

* The assignment should be **maximum 25 pages**, including tables, figures; excluding references.

* Please make sure that you do not just hand in computer outputs. 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.

* A pdf 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.

* The pdf document must have the group number, the names and student numbers of each member of the group.

* You must include a short report explaining the contribution of each group member to the writing process of the assignment.

The Data

This assignment comes with a dataset that you need to analyse. The data set is uploaded on Canvas. The data set is a time-series ranging between 1961 - 2016. There is data on two countries France and India, the two of the biggest agricultural producers in the world. Here are the details of the variables that the data set contains:

mean_pre: Annual average of monthly total rainfall (millimetres per month)

mean_rad: Annual average of monthly total all-sky downward Surface Solar Radiation (SSR) (watts per square metre)

mean_tmn: Minimum of averages of monthly temperatures (degrees Celsius)

mean_tmp: Annual average temperature (degrees Celsius)

mean_tmx: Maximum of averages of monthly temperatures (degrees Celsius)

NY.GDP.MKTP.KD: GDP (constant 2010 US$)

SP.POP.TOTL : Population, total

AG.LND.AGRI.K2: Agricultural land (sq. km)

AG.PRD.CROP.XD: Crop production index (2014-2016 = 100)

NV.AGR.TOTL.KD: Agricultural GDP (constant 2010 US$)

The Assignment

The assignment consists of four parts: In Part (1) you will simulate some data and analyze the simulated data. In Part (2) you need to use graphical means to analyse your data and you need to decide which initial transformation of the data you need to apply. In Part (3) you need to investigate the non-stationarity properties of your data series (See Davidson 2000, Chapter 14, lecture notes and relevant literature). Part (4) should be devoted to cointegration analysis and the estimation of the models (See Davidson 2000, Chapters 15-16, lecture slides and relevant literature). It is important to understand that this assignment requires you to have a clear intuition/model in mind in the sense that you are going to investigate the existence of possible long run relationships in your data set.

You need to choose one of the two countries. By using your intuition or knowledge on economic-climate relationships, choose 2 climate variables and 2 economic variables.

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.

1. Simulating and analyzing time-series data

(i) Simulate $T = 200$, 3×1 multivariate time-series data by using the DGP:

$$\mathbf{x}_t = \boldsymbol{\delta} + \boldsymbol{\Lambda}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\delta}_t$ is a 3×1 vector of time-invariant deterministic components, $\mathbf{x}_t$ is a 3×1 vector. The coefficient matrix $\boldsymbol{\Lambda}$ is a 3×3 matrix and $\boldsymbol{\varepsilon}_t$ is a 3×1 vector of innovations. Make your own assumptions about all model components to ensure **wide-sense stationarity** of $\{\mathbf{x}_t\}$.

Write all your assumptions about all model components in the report.

(ii) Show that $\{\mathbf{x}_t\}$ is a wide-sense stationary multivariate time-series process. Explain how you conducted this analysis and elaborate on your findings in the report.

(iii) Plot the data set you generated. Include the plot in the report.

(iv) Obtain the average for each 3 time series data you simulated ($\overline{\mathbf{x}}_t$). Obtain the means of the 3 time series processes [$E(\mathbf{x}_t)$]. Compare the averages with the means. Report your findings. Briefly comment on your findings.

(v) Obtain the sample variance for each 3 time series data you simulated. Obtain the variance of the 3 time series processes. Compare the sample variance of the data with the variance of the processes. Report your findings. Briefly comment on your findings.

(vi) Simulate $T = 200$, 3×1 multivariate time-series data by using the DGP:

$$\mathbf{x}_t = \mathbf{\Lambda}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where $\mathbf{x}_t$ is a 3×1 vector. The coefficient matrix $\mathbf{\Lambda}$ is a 3×3 matrix and $\boldsymbol{\varepsilon}_t$ is a 3×1 vector of innovations. Make your own assumptions about all model components to ensure **unit root non-stationarity** of $\{\mathbf{x}_t\}$.

Write all your assumptions about all model components in the report.

(vii) Show that $\{\mathbf{x}_t\}$ is a unit root nonstationary multivariate time-series process. Explain how you conducted this analysis and elaborate on your findings in the report.

(viii) Plot the data set you generated. Include the plot in the report. Comment on the behaviour of the data.

(ix) Simulate $T = 200$, univariate time-series data by using the DGP:

$$x_t = x_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim N(0, 1)$ is the innovations.

(x) Plot the data, include the plot in your report.

(xi) Test for a unit root in the simulated data by using the Dickey Fuller test.

Please write your detailed comments and interpretations of the results you obtain.

2. Graphical analysis of the data

Now it is time to work with the data. Start with a graphical analysis of the time series. By graphical analysis we mean plotting the data (levels, logs, first differences). Using these graphics, you should already be able to address, at least partly, issues such as

(i) 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)?

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

(iii) Decide whether you want to use logarithms of any of the variables.

Please write your detailed comments and interpretations of the results you obtain.

3. Analysis of the order of integration

Consider now the issue of testing formally for the presence of unit roots in your series. Try to carefully apply and design 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.

Please write your detailed comments and interpretations of the results you obtain.

4. Cointegration analysis

Once the integration of the order of the data has been determined, you need to carry out a cointegration analysis. Search in the climate science literature, and in the economics of climate literature for any scienfitic/intiutive reason to investigate long-run relations between the variables you have chosen to analyze.

(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 some of the techniques discussed during the course, in Davidson 2000 and in the related literature. Compute and compare the outcomes of various tests for no-cointegration. You may use

- The standard residual based cointegration tests using a static regression. (Engle and Granger approach, ADF type tests, Phillips-Ouliaris' test);
- 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 and in Davidson 2000, such as static least squares, DOLS, FMOLS and ECM. Report your estimation and inference results (coefficient estimates, test statistics, hypothesis tests, test results etc.) Discuss the similarities and differences between

your results of different methods. Discuss the differences between the assumptions of these methods.

(v) Adopt a systems approach and use Johansen's analysis (Chapter 16 of Davidson 2000) to estimate the cointegrated system. Report your estimation and inference results (coefficient estimates, test statistics, hypothesis tests, test results etc.) Discuss the similarities and differences between the results of Johansen's systems approach and the results of the single equation approach (. 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?

Please write your detailed comments and interpretations of the results you obtain.

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) simulating and analyzing simulated data is 20 points

   (ii) visual analysis of the data set is 10 points.

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

   (iv) cointegration analysis is 30 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.