LSE Centre for Macroeconomics
Pre-Doctoral Research Assistant Application

January 2026

Please complete the following data analysis and computation tasks (target completion time: 3–4 hours). Submit (i) a short document (preferably in LaTeX) describing what you did and summarising your findings, with sufficient detail for a trained economist to replicate your results, and (ii) your code. The code should be readable and fully reproducible.

You may use any statistical software and/or programming language, but ensure all outputs (figures, tables, etc.) are clear and easy to interpret. If any instructions are ambiguous, make a reasonable assumption and document it. If you encounter issues with formulas or notation, apply reasonable corrections and note them. If you get stuck, explain how you would proceed and what resources you would consult with more time. You are not expected to complete every component perfectly; submit what you are able to complete. **Importantly, aim to provide complete answers while remaining concise and avoiding unnecessary detail.**

You may, and are encouraged, to use generative AI tools. If you do, document how and where they were used in your submission. Good luck!

THIS TASK WILL BE GRADED ANONYMOUSLY—DO NOT INCLUDE ANY IDENTIFYING INFORMATION ANYWHERE IN THE TEXT OR THE CODE. FAILURE TO COMPLY WILL RESULT IN FORFEITURE OF YOUR APPLICATION.

Task A: Empirical Analysis

We are interested in investigating the relationship between productivity growth and exports at the sectoral level in Europe.

Q1a. Part of the job of an RA is to download raw data from credible sources and format it into a usable dataset. In this task, you will measure total factor productivity (TFP) using the Cobb–Douglas production function:

$$Y_{cst} = A_{cst} K_{cst}^{\alpha} L_{cst}^{1-\alpha}, \tag{1}$$

where c indexes countries, s sectors, and t years. Y_{cst} is output, K_{cst} capital, L_{cst} labour, and A_{cst} the Solow residual (TFP). Set $\alpha = 0.33$.

Using annual data from Eurostat (EU National Accounts), collect the series needed for Germany, Spain, Portugal, Austria, and the Netherlands over 2000–2023:

- Y: nominal gross value added from `nama_10_a10` (Current price million euro);
- L: total hours worked from `nama_10_a10_e` (Thousand hours worked);
- K: capital stock from `nama_10_nfa_st` (Current replacement costs, million euro);

Compute A_{cst} and annual TFP growth.

Do this for the total economy (NACE `TOTAL`) and for sectors `A`, `C`, `F`, `J`, `K`, and `L`. For this subquestion, plot only `TOTAL`: produce a single figure showing country TFP growth and an output-weighted aggregate across countries in each year. Additionally, report the mean aggregate TFP growth over 2000–2023, either as an annotation on the figure or in a short summary table.

Q1b. In a short written response, explain the economic interpretation of α in the Cobb–Douglas framework. Describe how α can be measured in practice using standard data sources, and discuss whether imposing a common α across sectors is appropriate.

Q1c. You are provided with a recession indicator series. Incorporate this indicator into the *sectoral* TFP growth figure. Produce a single figure with one panel per sector (`A`, `C`, `F`, `J`, `K`, `L`). In each panel:

- plot the *sectoral* country-level annual TFP growth series;
- plot a *sectoral* output-share weighted aggregate TFP growth series across countries (weights based on each country's sectoral gross value added);
- indicate recession years on the same time axis;
- report the mean of the aggregate TFP growth series for the sector, either as a numeric annotation within the panel or in a separate summary table.

Q1d. Briefly interpret what you observe in the sectoral recession plot. Be complete and concise in your answer, and rely on economic theory.

Q2a. Using Eurostat trade statistics (`ext_tec09`), obtain annual exports by country and sector for the same set of countries and sectors as in Q1, restricting the sample to 2012 onward. Merge these data with your output series to construct *exports as a percentage of output* at the country–sector–year level. Provide a table of descriptive statistics for this variable.

Q2b. Using the exports-as-a-percentage-of-output measure from Q2a and the productivity series constructed in Q1, merge the datasets on country, year, and sector. Produce two figures summarising the relationship between log productivity and exports as a percentage of output:

- one for the total economy (`TOTAL`);
- one with a separate panel for each sector.

Briefly describe what the plots suggest.

Q3a. Using the merged sectoral dataset, estimate an econometric model of your choice to assess the relationship between exports and log productivity.

- Provide a brief rationale for your specification.
- Report the parameter estimates and interpret the results.
- Discuss potential issues with the regression (e.g. identification concerns).
- Propose creative approaches to address the main concern you identify.

Task B: Literature

Tell us about your favourite paper in economics. What do you like about it? How would you improve it? [Max. 200 words]

Task C: Computational Exercise

All questions in this task are mandatory, but each question tests for a different set of skills, and we don't necessarily expect you to have the background knowledge to be able to solve all of them. You should attempt all questions and not get discouraged if you can't solve one or more of them. If you can't solve a question fully and explicitly, solve as much as you can and walk us through what you expect the answer to be given your intuition. What forces are at play? Do they all work in the same direction, or do they work against each other? Demonstrating your understanding of this is very valuable.

A representative household lives in discrete time $t = 0, 1, 2, \dots$ and maximizes expected lifetime utility:

$$E_0 \sum_{t=0}^{\infty} \beta^t \, u(c_t),$$

where c_t is consumption and

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma},$$

with parameters β and σ. *We interpret each period t as a quarter in this problem.*

The household holds a stock of risk-free assets a_t that yield a constant gross return $R \geq 1$. Each period it also receives stochastic labor earnings y_t. Borrowing is not allowed:

$$a_{t+1} \geq 0.$$

Stochastic Earnings Process

Labor income y_t follows a finite-state Markov process. It can take N_y possible values

$$y_t \in \{y^1, y^2, \dots, y^{N_y}\}, \qquad 0 < y^1 < y^2 < \cdots < y^{N_y},$$

with transition probabilities summarized by an $N_y \times N_y$ Markov transition matrix

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1N_y} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2N_y} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{N_y1} & \pi_{N_y2} & \cdots & \pi_{N_yN_y} \end{pmatrix}, \qquad \pi_{ij} = P(y_{t+1} = y^j \mid y_t = y^i), \qquad \sum_{j=1}^{N_y} \pi_{ij} = 1.$$

In the numerical code, the income process is specified as follows:

$$y_t \in \{0.35,\ 0.7,\ 1.0,\ 1.3,\ 1.65\},$$

with corresponding transition matrix

$$\Pi = \begin{bmatrix} 0.5 & 0.3 & 0.1 & 0.05 & 0.05 \\ 0.2 & 0.5 & 0.2 & 0.05 & 0.05 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \\ 0.05 & 0.1 & 0.2 & 0.5 & 0.15 \\ 0.05 & 0.05 & 0.1 & 0.3 & 0.5 \end{bmatrix}.$$

Tasks

1. Dynamic Programming Setup and Intuition

Q1a. Define the state and choice variables. Write the household's per-period budget constraint and the Bellman equation.

Q1b. Derive the Euler equation using the first-order and envelope conditions. Clearly explain each step.

Q1c. Provide an economic interpretation of the Euler equation and the borrowing constraint $a_{t+1} \geq 0$. In particular, explain how income risk and the no-borrowing constraint affect saving behavior.

Q1d. Briefly explain what it means for the income process to be persistent (one or two sentences are sufficient).

2. Numerical Solution

You are provided with Python/Matlab code.

Q2a. Briefly describe the code and the numerical method used.

Q2b. Without rewriting the code, list a few concrete ways to improve efficiency or numerical performance.

3. Distribution of Consumption and Savings

Q3a. Using the policy functions computed in the provided code, solve for the stationary joint distribution of assets and income, and the implied distribution of consumption. Describe clearly how you construct the cross-sectional distributions.

Q3b. Plot the stationary distributions of savings a, income y, and consumption c.

Q3c. Briefly discuss at least one alternative approach to computing the stationary distribution, and provide the intuition for how it works. *You do not need to code this.*

4. Marginal Propensity to Consume (MPC) and Calibration

Extend your numerical code to compute the model-implied marginal propensity to consume.

Q4a. Extend the code to compute the marginal propensity to consume (MPC) associated with an unexpected one-unit increase in assets, and report the model-implied *average* MPC.

Q4b. A key part of theoretical work is ensuring that models are disciplined by empirical evidence. You do *not* need to estimate the MPC empirically. Instead, choose a plausible target for the average MPC from the literature and briefly justify your choice. Then calibrate the model by adjusting the relevant parameter(s) accordingly. Clearly state which parameter(s) you vary and provide a brief rationale.

Q4c. Report the calibrated parameter(s) value(s) and briefly discuss how changing these parameter(s) affect the MPC in your model.

5. Open Question: Incomplete Optimization / Behavioral Agents

This question is open-ended. We do *not* expect a fully worked-out solution. You may answer using theory, diagrams, pseudo-code, or any other clear format. Clear intuition and reasoning will be valued more than formal rigor.

Q5a. Suppose the household does not always perfectly optimize its consumption–savings decision (e.g., it occasionally follows a rule of thumb, makes noisy or approximate decisions, or adjusts assets infrequently). Propose and briefly discuss a few ways to model and solve a dynamic programming problem with *imperfect* or *boundedly rational* optimization.

Q5b. Choose one approach from *Q5a* and explain how you would adapt the solution method for *our* consumption–savings problem to incorporate it. Describe the main steps required to solve and simulate the model, in words and/or pseudo-code.