

ECON 2300: INTRODUCTORY ECONOMETRICS

Research Project 2

Due: 4pm, 7 November 2022

Submission of your report

Your report must be single-spaced and in 12 Font size. You should give your answer to each of the following questions following a similar format of the solutions to the tutorial problem sets. When you are required to use R, you must show your R command and R outputs (screenshots or figures generated from R). You will lose **2 points** whenever you fail to provide R commands and outputs. For each question, when you are asked to discuss or interpret, your answer has to be brief and compact. You will lose **2 points** if your answer is needlessly wordy. You must upload your assignment on the course webpage (Blackboard) in PDF format. (Do not submit a hard copy.)

This project has two research questions. You are required to investigate both of them.

Problem 1: money, Growth, and Inflation (30 marks)

Background

To examine the quantity theory of money, Brumm (2005) ["Money Growth, Output Growth, and Inflation: A Reexamination of the Modern Quantity Theory's Linchpin Prediction," *Southern Economic Journal*, 71(3), 661–667] specifies the inflation equation

$$\text{inflat} = \beta_1 + \beta_2 \text{money} + \beta_3 \text{output} + u$$

where `inflat` is the growth rate of the general price level, `money` is the growth rate of the money supply, and `output` is the growth rate of national output. Economic theory suggests that $\beta_2 = 1$ and $\beta_3 = -1$. The dataset `brumm.dta` consists of 1995 data on 76 countries.

Research tasks

1. It is argued that `output` may be endogenous. Four instrumental variables are proposed, `initial` = initial level of real *GDP*, `school` = a measure of the population's educational attainment, `inv` = average investment share of *GDP*, and `poprate` = average population growth rate.
 - (a) Give an intuitive explanation as to why `output` can be endogenous (3 marks)
 - (b) Explain why the proposed IVs can be valid. (6 marks, i.e., 2 marks for understanding of valid IVs and 1 mark for convincing story for each IV)
2. Using the four IVs, obtain TSLS estimates of the inflation equation (4 marks), and test the economic theory using the IV estimates (4 marks).
3. Determine whether the IVs are strong or not (3 marks) and test if they are exogeneous (3 marks).
4. Present a short research note (less than a half page) of your findings. You are allowed to use your previous findings here again or to estimate again the model using a different set of IVs (7 marks).

Problem 2: Demand for Democracy (70 marks)

Background:

Do citizens demand more democracy and political freedom as their incomes grow? That is, is democracy a normal good? To investigate this issue, you will explore the dataset `Income.Democracy.dta` which contains a panel data set from 195 countries for the years 1960, 1965, ..., 2000. A detailed description is given in `Income.Democracy.Description.pdf`.¹ The dataset contains an index of political freedom/democracy for each country in each year, together with data on the country's income and various demographic controls. (The income and demographic controls are lagged five years relative to the democracy index to allow time for democracy to adjust to changes in these variables.)

Research tasks:

1. Is the data set a balanced panel? Explain. (5 marks)
2. The index of political freedom/democracy is labeled `dem_ind`.
 - (a) What is the value of `dem_ind` for the United States in 2000? What is the average of `dem_ind` for the United States over all years in the data set? (4 marks) Repeat this exercise for Libya (2 marks).
 - (b) List five countries with an average value of `dem_ind` greater than 0.95; less than 0.10; and between 0.3 and 0.7. (5 marks)
3. The logarithm of per capita income is labeled `log_gdppc`.
 - (a) Regress `dem_ind` on `log_gdppc` using standard errors that are clustered by country (3 marks).
 - (b) How large is the estimated coefficient on `log_gdppc`? Is the coefficient statistically significant? (2 marks)
 - (c) If per capita income in a country increases by 20%, by approximately how much is `dem_ind` predicted to increase? What is a 95% confidence interval for the prediction? Is the predicted increase in `dem_ind` large or small? Explain what you mean by large or small. (5 marks)
 - (d) Why is it important to use clustered standard errors for the regression? Do the results change if you do not use clustered standard errors? (4 marks)
4.
 - (a) Suggest a variable that varies across countries but plausibly varies little or not at all over time and that could cause omitted variable bias in the regression in Question 3 above. (5 marks)
 - (b) Estimate the regression in Q3, allowing for country fixed effects. How do your answers to Q3(b) and Q3(c) change? (5 marks)
 - (c) Exclude the data for Azerbaijan and rerun the regression. Do the results change? Why or why not? (5 marks)
 - (d) Suggest a variable that varies over time but plausibly varies little or not at all across countries and that could cause omitted variable bias in the regression in Q3. (5 marks)
 - (e) Estimate the regression in Q3, allowing for time and country fixed effects. How do your answers to Q3(b) and Q3(c) change? (5 marks)
 - (f) There are additional demographic controls in the data set. Should these variables be included in the regression? If so, how do the results change when they are included? (5 marks)
5. Based on your analysis, what conclusions do you draw about the effects of income on democracy? (10 marks)

¹These data were provided by Daron Acemoglu of M.I.T. and were used in his paper with Simon Johnson, James Robinson, and Pierre Yared, "Income and Democracy," *American Economic Review*, 2008, 98:3, 808–842.