FIE401 - Third assignment
Does financial literacy cause participation in the stock market?

Maximilian Rohrer

Overview

Individuals have become increasingly active in financial markets. This increase in the market participation by individuals has been accompanied or even promoted by the advent of new financial products and services. However, some of these products are complex and difficult to grasp, especially for financially unsophisticated investors.

At the same time, market liberalization and structural reforms to social security and pensions have caused an ongoing shift in decision-making responsibility away from the government and employers and towards private individuals. Thus, individuals have to take more responsibility for their own financial well-being. This process is not universally seen as only positive since not all individual might have the skill and knowledge to do so.

This started a long string of research on the stock market participation of private individuals. In this exercise, you investigate one of the first questions asked in this literature: are individuals that have more finance related knowledge (a concept often referred to as financial literacy) more likely to participate in the stock market?

Formalities

This assignment will be handed out on 27.10.2022 at 17:00 and has to be submitted no later than the 3.11.2022 at 14:00. Submit your commented coding file (file extension: ".R") and a pdf file (file extension: ".pdf") including the numerical results and answers to questions posted.

Please comment your code shortly so that a reader can reconstruct your thinking. You do not need to explain the used functions. You do not need to describe your coding in the pdf file. Please keep your answers very brief.

After submission, the assignments will be randomly redistributed. Read your peer's work carefully and compare your solution to theirs. There are many different approaches to code the same exercise, so let's learn from each other. You do not need to provide comments and there is no deadline. From experience, I would encourage you to look at the work of your peers. They will use different functions and different approaches.

In order to make it easy for us to allocate individual assignment submission as well as peer review to groups, please follow following file name structure. For the assignment submission: "Assignment3_group_XX.R/pdf". "XX" indicates the group number that submitted the assignment.

Please do not include any personal information such as name, student number, etc. in the submitted files.

Work together in groups of four.

Submit your assignment even if you do not finish all tasks. For a pass, 50% has to be correct. I would stress that discussion, interpretation, and presentation of results is equally important to implementing code.

We touched on almost all functions you have to use to solve this assignment. If you are unsure how to use a function, either use R's own documentation (type ?command) or use www.stackoverflow.com. We encourage

you to code this assignment yourselves, and not to use purpose-made solutions or packages as provided on the internet.

Data provided for this assignment

The `Finlit_new.csv` file provides a dataset allowing you to explore the aformentioned question. In specific, you may find following variables.

- Main dependent variable
  - dstocks_mut: dummy for stock market participation
- Main independent variable
  - indexlit1_new1: advanced literacy index
- Controls group 1
  - age: age in number of years
  - edu2: dummy for pre-vocational education
  - edu3: dummy for intermediate vocational education
  - edu4: dummy for secondary pre-university education
  - edu5: dummy for higher vocational education
  - edu6: dummy for university education
  - male: dummy for gender
  - partner: married or not
  - numkids: number of kids
  - retired: retired or not
  - dum_selfempl: dummy for self-employment
  - lincome: log(houselhold income)
  - tot_non_equity_wealth_cat: wealth quartile
  - indexlit2_new1: basic literacy index
- Controls group 2
  - b2: economics education
  - b3: daily use of economics
- Potential instruments
  - f10: sibling's financial situation
  - f15: parent's financial knowledge
- Included variables not relevant for current analysis
  - edu1: dummy for primary education
  - b1: self-assessed literacy

Tasks

You estimate following model using a liner probability model:

$$dstocks\_mut = \alpha_0 + \beta_1 * indexlit1\_new1 + \beta_X * Controls + \varepsilon$$

`Control` indicates all control variables from control group 1. You might also want to include controls from group 2 in later regressions and compare.

In specific, execute following tasks:

- Task 1: Construct summary statistics to familiarize yourself with the data. If necessary make some adjustments to the data.
- Task 2: Use a `stargazer` table to report following three estimated models
  - Naive OLS regression

  - First stage of the IV regression
  - Second stage of the IV regression estimated in one go, so that standard errors are correct
  - Interpret your results. What does financial literacy mean for stock market participation, statistically and economically? Is there a difference between OLS and IV?

- Task 3: Assess the instruments. You replicated tables from Van Rooij, M., Lusardi, A., & Alessie, R. (2011): Financial literacy and stock market participation. Journal of Financial Economics, 101(2), 449-472. You can find the paper on Canvas. In particular, you replicated columns in Table 7, 8A and 8B. Carefully read chapter 5 (start page 461 at the second paragraph beginning with "In Table 7,. . . ") of the paper and answer following questions:
  - What reasons do the authors name for using an Instrumental Variable approach?
  - How do the authors abstract the estimates obtained using OLS and IV?
  - How do the authors discuss relevancy? (F-statistic)
  - After reading the authors explanations, are you convinced that both instruments are valid? (Exogeneity)

- PDF report
  - Make sure that your pdf starts with a **short and consise** abstract presenting your analysis (as you would in your Master thesis).
  - In writing, discuss your regression results briefly (5-10 sentences). Make sure that you discuss statistic as well as economic significance. Regarding the latter, argue whether or not the detected effects are meaningful.
  - The report has to be self-contained. In other words, it needs to contain all tables and any other piece of analysis that you reference in writing.
  - Answer the questions listed in Task 3.

I would strongly advise you to make a proper regression table as you will do so in the final exam. Please see the guidelines published with the first assignment. In the final exam, the presentation of empirical results (such as tables) will be graded.