In the context of ACFV, what has happened to returns to schooling in the 21st Century?

⛳ Aims

This assignment has several goals:

- Provide exposure to yet another high quality conditional quantile analysis paper in Angrist et al. (2006). Hereafter refered to as ACFV.
- Put practical skills to use by downloading, cleaning and implementing data from the yearly "census style" American Community Survey (ACS).
- Answer a (hopefully) interesting question by yourself, using a conditional quantile perspective and interpreting the results from your research.

📚 Background Information

ACFV is a very good (and complex) theoretical paper that looks at properties of Quantile Regression under misspecification. Within it, there is an empirical application to see how the returns to schooling coefficient changed during the last two decades of the past century.

They find that through the 80s returns to schooling increased across the entire income distribution, with slightly higher gains at the top of the distribution. In the 90s, average returns to schooling (calculated using OLS) didn't change much. However, using quantile regression they show that at the top of the income distribution returns to schooling had kept growing while at the bottom they had dropped!

In ACFV, the data used comes from the US census which takes place every ten years, much like the one currently going on here in the UK. However, since 2000 a new study, the ACS, has been collecting data on economic variables from individuals every year. Some of which, the census has stopped taking. Therefore for this project we will be using the ACS which is available yearly from the IPUMS website.

🤔 What should you do?

1. Have a look at ACFV, specially the empirical section (Section 4).
2. Download the code and data that the authors have made available and replicate their findings (Figure 2). Become familiar with what the code does and what the data looks like.

   > *There are some small details with their code that might need tweaking after 15 years.*
   > *In Angrist's Data Archive there is a document which walks you through the data that they used, you might want to check it out.*

3. Go to the IPUMS website. Register and download the relevant data to expand their analysis to more recent years. I leave to your discretion what years you select.
4. Perform some analysis on this "new" data to try and see what has happened to returns to schooling in the USA in the last two decades. I am leaving this fairly open for you to be creative, one good starting point is to use the ACFV code on your new data.

   > *If you do this, make sure that the variables are actually the same so you can compare your findings to theirs!*

5. Write a short report, **5 pages maximum**, on what you have found, hopefully answering the research question at the top of this document. You may also comment on any issues you found throughout the process relating to the coding and data work, as well as how you dealt with them.

   > *You are encouraged to be creative with this report!*

This assignment can be done individually or **in collaboration with one other student.** This includes in the report, which you should submit via email to me (d.laradeandres@bristol.ac.uk) by **Monday 3rd of May 2021.** Note that this will be purely formative work, an oportunity for you to use real life data and answer a relevant question.