Evaluation task

The goal of this exercise is to evaluate your ability to extract information from a data-set, present it in a clear and concise way and provide some interpretation. The task focuses on mobility patterns in the US across the 4 main regions in the NLSY 79. In each of the questions, please use combinations of plots and regressions as you see fit. We expect this to take about 5 hours, but to give you some flexibility, you have 24 hour to send us back your analysis. Using the data available in this folder (prepared as a CSV file), please do the following:

1) Summary statistics. Report the count of moves across US regions in the sample for each possible transitions. Report the count of moves between urban and non urban. Report the mean wages, employment and education attainment in each region and urban / non urban. Comment on differences you might find.

2) Event studies. Focusing on the movers only, report an event study of a window around the move for wages. This means that you should plot the mean wage of movers for years leading to the move, and years following the move (do plus and minus 2). The x-axis is time to the move and the y-axis is the mean wage. Each individual should be centered on the move at $t = 0$. Do this for moves between some regions and for moves between urban and non-urban.

3) Comparing movers to stayers. Compare the wage of movers to the wage of stayers in both region of origin and region of arrival. Do this while controlling for age, education and gender. Do this for different distance from the move. Document and interpret differences in levels and differences in trends if any.

Additional info:

About the data. relevant columns: i is individual id, educ is completed education, birth is birth year, region is the geographical area, urban is our measure of urban, wage is our measure of wages.

The data contains a lot of missing regions, urban and wages. Define employment as wage>0, and focus on individual for which you have information around the move.

What to turn in. Submit one report document with the results and your comments. We expect a length of about 5 to 10 pages including figures and tables. The text itself without figures should be 2 to 3 pages. We also ask that you submit the code that you used to generate the results. Please provide comments within the code so that we can evaluate how you approach the problem. You can use the platform of your choice.

Thank you!