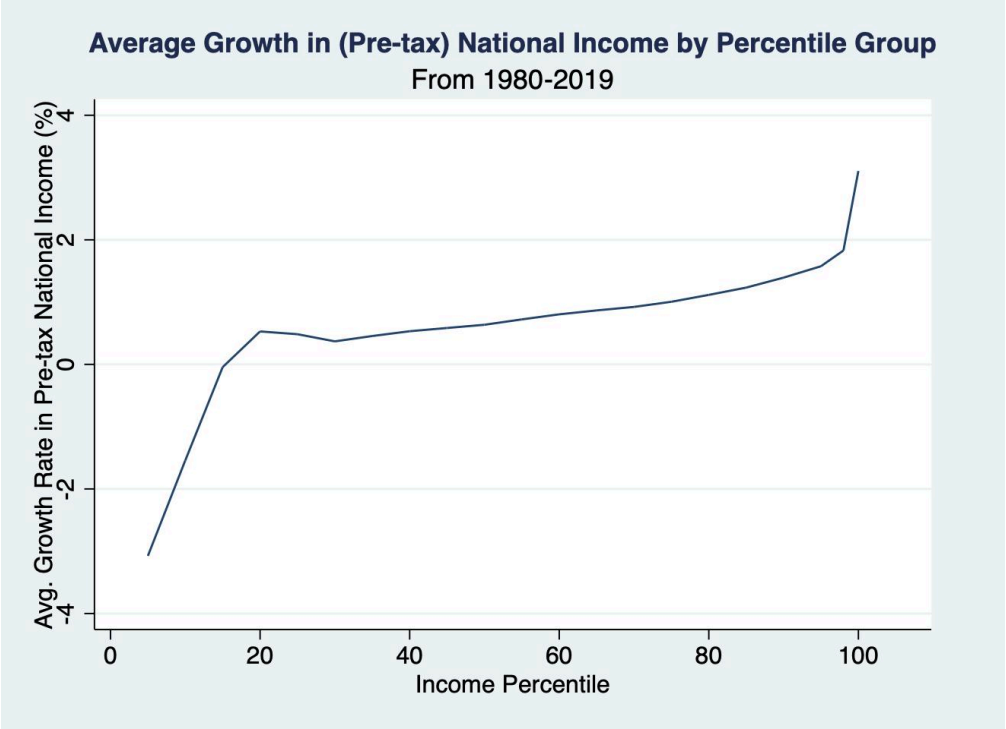
Garn Leangphibul

Task 1A - Saez & Zucman: Replicating Percentile Growth Chart

In this task, I replicated the percentile growth chart created in figure 4. of Saez and Zucman's paper, "The Rise of Income and Wealth Inequality in America" using data from WID. This chart models the average growth rate in national income during a time period, segmented by income percentile group:



The x-axis is the income percentile. Belonging in an x income percentile means that the individual earns equal or more than x % of the population. The y-axis is the average growth rate in national pre-tax income. The formula I used to calculate growth rate was:

Growth rate = (Year 2 Pre-tax income - Year 1 Pre-tax income) / Year 1 Pre-tax income

I then used the mean function on STATA to compute the average income growth rate for each income percentile.

Saez and Zucman prefer using national income over personal income as national income allows for better international comparisons of inequality, while also avoiding double counting some forms of income. Saez and Zucman also examine both pre-tax and post-tax income in the paper. Pre-tax income simply refers to all income flows

(including the pension system), prior to the tax/transfer system. I propose to use pre-tax income for the sake of easier comparison with Saez and Zucman's results, as they discussed the trends in average income growth amongst income percentiles using pre-tax income terms. In future work however, adopting post-tax income would be greatly beneficial in examining how effective the tax/transfer system has been on income growth inequality.

I decided to adopt a similar time period as Saez and Zucman, by using data between 1980-2019 (Saez and Zucman uses 1980-2018), with the same reasoning of allowing for closer comparisons in any trends found in average growth in income. Saez and Zucman chose to examine this time period as the US experienced significant growth in income inequality after 1980. This allowed for generating income percentile growth curve for two different time periods: before 1980 and after 1980, thereby allowing exploration of how income inequality has evolved over time (* but for our assignment, we were instructed to only look at post-1980).

I also adopted the same approach of using equal-split adults, which assumes that income of married spouses is fully shared or split equally between the spouses for the same reason. A merit stated in the paper is that this approach improves comparability, especially internationally by avoiding discrepant definitions across countries, such as for households and tax units.

A difference that occurred was that Saez and Zucman were able to use more granular data, at a 1 income percentile interval. However, due to limitations of the WID STATA package and consequently having to manually download the WID data, it was only feasible to work with data at 5 income percentile intervals. As Saez and Zucman's data were more granular, their results are able to better represent income growth inequality amongst extremely sub-specific groups (e.g. 10th percentile vs 11th percentile). Another caveat was the unavailability of pre-tax income data using the equal-split approach for the upper percentiles (above 95th percentile). As a result, I had to use individual income to supplement the missing data.

Despite these differences however, both the replicated chart and the paper's chart show a similar trend. This is exemplified in the negative average income growth observed below around the 15th percentile, followed by a positive relationship between income percentile and average income growth. A steep curve is also observed in both graphs for the top 1 income percentile – both graphs observe that between 1980-2018/2019, top income earners in the US experience average income growth significantly higher than everyone else in the economy. The only difference was that in our results, there was a short period of negative correlation between income percentile and average

income growth for the 20-30th income percentile. However, as Saez and Zucman's graph also show a very weak relationship (flat curve) at the same income percentile, this difference may be explained by difference in data granularity/income interval levels (as discussed earlier in our caveats).

Despite some limitations that prevented obtaining identical data and a minor difference in results (for the 20-30th income percentile), the percentile growth chart was able to be replicated in a similar manner and yielded very similar trends.

Comments:

- Data limitation: given data, what is missing or could have explained this better