

Explanation of the dataset and summary of analysis

Mathematics 23C: Mathematics for Computation, Statistics and Data Analysis

Presented by Michel Nehme and Ramiz A. Razzak

For our project, we analyzed data from Raj Chetty's Equality of Opportunity project, which goes over different metrics regarding a child's economic situation based on their parent's income percentile. The data is formatted in a specific way: it provides incomes for children based on the type of school they went to and their parent's income percentile. There are 15 different educational categories, ranging from Ivy Plus to Never Attended college (before year 2013).

Our analysis is primarily focused around the following aims: analyzing the effects of education on earnings and seeing how a parent's place in the economic ladder affects their child's economic outcomes. These questions are fundamental, because they directly or indirectly provide insights on the reality of education as an equalizing force in the United States of America. The implications of our analysis are wide-ranging, which is why we decided to approach this topic.

One of the key insights of our analysis was that ivy-league and elite college incomes were significantly higher than incomes from all other categories, *even* those incomes of students from highly selective private and public colleges. This conclusion is important, because if we make the safe assumption that the quality of education between America's top public and private colleges versus those that are Ivy/Ivy + (Elite) is almost identical, we realize that the large differential in income between students from Ivy+ Schools and top public and private colleges is primarily motivated by signaling. There is a significant amount of economic literature that touches this issue; for instance, economist Bryan Caplan believes that the value of education in America today is vastly exaggerated, since much of the increases in wages witnessed by those with college degrees depend on the college they've attended as opposed to the raw human capital they've gained. In our project, we performed a permutation test to arrive at this conclusion (by invalidating our null hypothesis that there wouldn't be an income gap between Ivy + vs top public/private) and we supplemented this analysis by a classic chi-square test.

An interesting addition to this discussion was when we looked into the difference in incomes between students attending a four-year public with two-year for-profit (across parental percentiles). What we see here is that there isn't a statistically significant difference in their earnings. This seems to go against the idea that more years of education would lead to higher human capital accumulation which in turn would lead to higher income. Our hypothesis for why this is the case again returns to signaling: similarly to how Ivy + students earn more money than their counterparts at equally challenging schools, because four-year public colleges and two-year for-profits convey a similar signal (they aren't particularly prestigious, but highlight that the student has done beyond a high school degree) they lead to similar earnings. If this hypothesis is correct, it strengthens the conclusions we've made in the paragraph above.

Our final conclusion from this project is that there is a non-negligible correlation between parental income percentile and child earnings. By producing a linear correlation between these two variables and computing the strength of the correlation, we found a line of best fit and we computed the strength of the correlation, which turned out to be a little above 0.5. In most statistics textbooks, this is categorized as a moderate correlation. The take away here is with regards to economic mobility: the influence of parental income percentile on child income is present, but not extremely strong. This again is an important question in both economic literature and in politics, and our analysis seems to indicate that both those who believe that success is entirely related to personal effort and those who believe that the wealthy inevitably end up with wealthy children are overstating their case.