Replication Project Report

Jacob Hood

May 14th, 2021

The main differences in my replication are due to 1) my elimination of data from the year 1970 and 2) the less restrictive income range I used. In cleaning the data, I eliminated the top and bottom one-percent of incomes. As shown in my replication of Figure 1, the range of average weekly wages is larger than in the corresponding figure the authors create. As consequence, my final samples for both men and women are larger. The authors eliminated more income percentiles from their analyses, potentially the top and bottom two-percentiles. My final sample included 8,689,284 people compared to 7,221,747 in the original paper. For the original paper, the highest average weekly wage barely reached $900, while the highest weekly wage in my sample exceeded $1000 a week. Given the difference of over 1.5 million people, this is the one data decision I did not agree with, as it eliminated a considerable portion of incomes.

In the regression analysis, the most significant differences were in the values for the constants, however all differences were less than 1 unit. All regression coefficients generally deviated no more than .2 units from those calculated in the original paper. My regression analyses also showed identical relationships (in terms of positive/negative relationships) among variables. However, for the Public Sector variable, some relationships were inverted (for instance, among men in 1980 and 2010), although these values deviate less than .05 units from the original values. These differences can likely be attributed to my larger sample size and corresponding broader income range.

Another issue and source of potential deviation was the coding of potential years of work experience, and how years of schooling was calculated. To calculate years of schooling, the EDUC variable was used, which provides intervals of school grades. I chose the mean number of years to represent each new category when re-coding. However, it appears that Mandel and Semyonov used the top-code of each EDUC character to create this measure, leading to higher estimates of potential work experience in their study.

Future extensions should include additional racial/ethnic groups. For instance, we could include Asian and Hispanic men and women in our analysis in order to see how these patterns compare. Complementary theoretical work, like the racial triangulation theory, for instance, can provide additional frameworks for understanding how inequality affects non-black people of color. Looking at trends for both of these groups - who may not track as cleanly as whites and blacks - can help us build more robust theories of how racial and gender differences influence economic inequality.

A comparative component could also identify the importance of national context for trends in wage gaps over time. For instance, how do these trends compare to similar nations with similar, but still distinct, histories of racial and gender discrimination, such as England or Canada? This extension could allow for making cross-national claims about how racism and sexism affect economic inequalities.

A fuller consideration of economic discrimination and inequality could also consider the highly gendered divisions of household labor alongside more formal waged labor. For instance, an ethnographic study could compare men and women's experiences of inequality in formal labor markets to their experiences of inequality in the private sphere of home economics where labor is not always compensated or even recognized.