Project 4: Socioeconomic Status, Gender, and Scientific Literacy

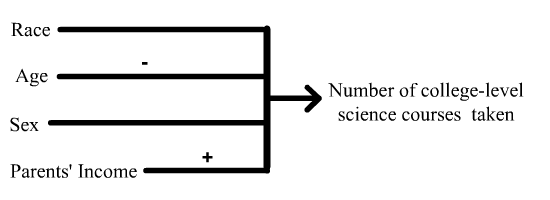
INST314-0102

Suneel Rahman

We would like to examine whether socioeconomic factors such as race, age, sex, and parents’ income influence the level of science education American adults have received. To do this, we will be examining a target population of American adults who participated in the General Social Survey. This is an important issue because being scientifically literate allows adults to make well-informed decisions both at the polls and in everyday life. These decisions influence the world around us. For example, when individuals do not understand the effect of greenhouse gases on the atmosphere, they are more likely to live their lives in such a way that is more harmful to the environment. This lack of scientific education may disproportionately affect individuals from certain backgrounds, such as those who grew up poor, women, or older Americans.

Our research question is as follows:

Do race, age, sex, and parents’ income correlate with the number of collegiate-level science courses they have taken? We believe that Black and Latino Americans will have taken fewer college science courses than their white counterparts. We believe men will have taken more college science courses than women. We believe younger Americans will have taken more science courses than older Americans. We believe those who come from wealthier families will have taken more college science courses than those who grew up in low-income families.



*H1: Nonwhite Americans will have taken fewer college science courses than whites.*

*H2: Younger Americans will have taken more college science courses than older Americans.*

*H3: Male Americans will have taken more college science courses than female Americans.*

*H4: As parents’ income rises, so does the number of college science courses one has taken.*

We will be using data from the General Social Survey (GSS). Our four independent variables are race, age, sex, and parents’ income. Race, the respondent’s race, is a nominal variable that have been recoded; black and other are 0 while white is 1. Sex is whether the respondent is male or female; Men have been recoded to 0 while women are 1. After recoding, there are 2100 1s and 767 0s. Parents’ income, the parent’s income in US dollars at the time they were 16, is ordinal and has been recoded sequentially in ones starting from 1. NAs were recoded as 0s. The median category for parents’ income was 19, which corresponds to an income between 50 and 60 thousand US dollars per year. Age has been recoded so the few respondents who are 89 or older are just counted as 89, leaving age as a ratio variable with a minimum of 18, a maximum of 89, a mean of 31.85, and a median of 32. Our dependent variable is the number of college science courses an individual has taken. It is a ratio variable and has been left as such except all ‘NA’s were recoded as 0s.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table 1: Independent Variables** | | | | | | |
|  | Freq | Min | Max | Med. | Mean | SD |
| Age | 2867 | 18 | 89 | 32 | 31.85 | 2.1 |
| Parent’s Income | 2867 | 0 | 26 | 19 | 17.37 | 3.1 |

|  |  |  |
| --- | --- | --- |
| **Table 2: Categorical**  **Independent Variables** | | |
|  | Freq. | Percent |
| Sex |  |  |
| Male (0) | 1276 | 44.5 |
| Female (1) | 1591 | 55.5 |
| Total | 2867 | 100% |
| Race |  |  |
| Nonwhite (0) | 767 | 26.8 |
| White (1) | 2100 | 73.2 |
| Total | 2867 | 100% |

Our regression model is the following:

*y = b0 + b1X1 + … + bnXn + ε*

Number of science courses taken = *b0 + b1*(race*) - b2*(age) *- b3*(sex) *+ b4*(parents’ income) *+ ε*

We performed an OLS multiple regression test with a = .05. None of the OLS regression assumptions were violated to an extreme or unworkable degree. We found the following:

Number of science courses taken = -.212 + .362(race) - .017(age) - .051(sex) + .117(parents’ income) + e

Our entire model was statistically significant. Parents’ income was the strongest predictor of number of college classes taken. Despite having a much smaller coefficient than race, it has 27 levels while race only had 2. A “bigger” difference is possible in income than race; there will on average be a smaller difference between an average nonwhite American and a white American than an American whose parents were poor and one whose parents were rich. Race was also a strong predictor as whites had taken more college science courses on average than their nonwhite counterparts. Men took more science courses on average than women. While significant, the average difference between the amount of science courses younger Americans had taken and the amount older Americans had taken was relatively very small. Our regression results confirmed all four of our alternate hypotheses, although the effect size of age was much smaller than we anticipated. Our R2 = .39, indicating our model has medium predictive strength. A lot of the other 61% is likely explained simply by whether an individual is interested in science.

Our findings imply that there is inequity in access to college-level science education. We can use this to identify marginalized groups that could use more equitable treatment in collegiate science.

The main weakness of our study is that anybody who did not go to college, which is already disproportionately nonwhites and people from low-income families, automatically gets a zero in the dependent variable. This may be remedied by comparing only individuals who went to college.