Project Proposal

Susan Cossa, Da Yeon Lee, Clara Rosenberg, Cas Sweeney

2/17/2018

"Hedgehog Economists" Group Project Proposal

Title

The Relationship Between Income Inequality and Educational Achievement

Purpose and Hypotheses

We are interested in looking at income inequality and the effect that it has on educational achievements. There is evidence that higher income levels lead to higher educational achievements. There might be more than one cause of differences in educational achievements, but in our project, we are going to focus heavily on the relationship between educational achievements and income level. Through this project, we want to ask these following questions: Is there a relationship between household incomes and the academic achievement of students? Are students more likely to succeed in schools if they come from wealthier families? Is academic achievement associated with household income?

The Primary Hypothesis is that, the greater the level of income, the higher the educational achievements will be because of increased access to a variety of resources such as tutoring.

The Secondary Hypothesis is that, Race Majority people tend to be wealthier and have a greater access to educational resources; therefore children from White families will have higher Education Achievements.

Data

This study will use data from U.S. Census Data for Social, Economic, and Health Research.

Response Variable

The Responsible Variable will be Educational achievements, such as GPAs, SAT scores, etc.

Main explanatory variable(s)

Income Inequality

Primary Hypothesis

The greater the level of income, the higher the educational achievements will be due to increased access to a variety of resources such as tutoring.

The primary hypotheses will use a variable indicating income level. Income level will be measured in dollars.

The test of our hypothesis will be

$H_0 : \beta_1 = 0$

$H_a : \beta_1 \neq 0$

from the regression model:

$EducationalAchievement = \beta_1 * HouseholdIncome + \beta_0 + \beta'_x$

where β'_x is a vector of covariates included to control for potential confounding. They are described below.

Secondary Hypothesis

Race Majority people tend to be wealthier and have a greater access to educational resources; therefore children from White families will have higher Education Achievements.

The secondary hypothesis will use a variable indicating Income Level and Race variable.

IPUMS included 5 categories: White, Black/African American, American Indian/Alaskan Native, Asian, Multiple Race in the RACENEW variable.

We will try to test each racial group separately, but we may have to collapse into an indicator/binary variable of white vs. non-white race.

The test of our hypothesis will be

$H_0 : \beta_3 = 0$

$H_a : \beta_3 \neq 0$

from the regression model:

$EducationalAchievement = \beta_1 * HouseholdIncome + \beta_2 * WhiteRace + \beta_3 * (HouseholdIncome * WhiteRace) + \beta_0 + \beta'_x$

Potential Confounding Variables

Even though Income Level is the main explanatory variable, there are other variables that would be associated with both Income Level and Academic Achievements.

Race (White & Non White)

Household structure (Single Parent & Multiple Parents)

Sex (Male & Female)

Quality of the School System/Curriculum (1: Excellent, 2: Good, 3: Moderate, 4: Poor)

Annual fund that a school gets from a government (in Dollar unit)