The Effects of Student Growth Data on School District Choice:
Evidence From a Survey Experiment

Pre-Analysis Plan

David M. Houston
Harvard University, Kennedy School of Government

Jeffrey R. Henig
Teachers College, Columbia University

September 13, 2018

We seek to identify the effects of providing district-level average student achievement data and/or average student growth data on subjects' hypothetical school district enrollment decisions. Compared to average achievement at a single point in time, average growth is arguably better able to capture schools' and school systems' contributions to student learning. Average growth also bears a weaker relationship to the racial and socio-economic composition of the student body, making it easier to identify highly effective schools that serve less advantaged students.

This study consists of an online survey experiment in which each subject is asked to imagine that s/he is a parent who is moving to a new city. When deciding where to live, one of their top priorities is to choose a school district for their elementary school-age child. The survey will provide basic demographic information about the five largest school districts in the metropolitan area. In addition to the demographic information, subjects will be randomly assigned to receive either 1) average achievement data, 2) average growth data, 3) both, or 4) neither. Based on these data, subjects will choose their preferred school district.

Subjects will be asked to complete this process for the metropolitan areas of the five largest cities in the United States: New York City, Los Angeles, Chicago, Houston, and Phoenix. At the end of the survey, subjects will answer a small battery of basic demographic questions.

We will conduct this experiment three times with different subjects. First, we will run a pilot experiment with 50 subjects. Second, we will run an initial experiment with up to 2,500 subjects. Third, we will run a replication experiment with up to 2,500 subjects.

Research Questions

1. Compared to receiving no information about student performance, what are the effects of receiving average achievement data, average growth data, or a combination of both average achievement and average growth data on the racial and socio-economic compositions of subjects' hypothetical school district choices?
2. Compared to receiving only average achievement data, what are the effects of receiving average growth data or a combination of both average achievement and average growth data on the racial and socio-economic compositions of subjects' hypothetical school district choices?

3. Do these effects vary by:
    a. The presence of children age 0-17 in the household?
    b. The presence of public school students age 0-17 in the household?
    c. Race (white and non-white)?
    d. Family income (under and over $75,000)?

Subject Pool

We will use Amazon's Mechanical Turk (MTurk) service to recruit respondents for the survey experiments. For the pilot experiment, we will recruit 50 subjects. For the initial and replication experiments, we will recruit up to 2,500 subjects each. We will limit potential subjects to those who 1) are age 18 or older, 2) live in the United States, 3) have an MTurk approval rating of ~~95%~~ 90% or above, and 4) have completed at least ~~500~~ 100 MTurk HITs in the past. (10/9/18)

We plan on conducting the pilot, initial, and replication experiments ~~in mid-September~~ on 10/3/18, ~~mid-October~~ ~~10/17/18~~ 10/24/18, and ~~mid-November~~ ~~11/14/18~~ 2/13/19, respectively. We will post each experiment on MTurk at 12:00pm ET and collect data for up to 14 days (or 2,500 subjects). Each experiment will take roughly 10 minutes to complete and will pay $2.00. (10/9/18) (10/23/18) (1/29/19)

Other Data Sources

For measures of school district achievement, growth, median household income, free and reduced price lunch eligibility, and racial composition, we will use the Stanford Education Data Archive (Version 2.1). In addition to district-level demographic data, the Stanford Education Data Archive contains district-level data from state standardized tests administered to students in grades 3-8 from 2009-2015. The student test score data have been converted to a common scale that allows district-to-district comparisons across the country. We use the empirical Bayes grade cohort scale estimates for the measures of achievement and growth. Achievement is measured such that a score of six represents a school district where the average student scores at about the same level as the average sixth grader in the national reference cohort. Growth is measured such that a score of 1.10 represents a school district in which the average student's test scores improve about 1.10 grade level equivalents in one year. To aid in the interpretability of these values for respondents, we will also report achievement and growth scores in terms of national percentiles.

To assess the representativeness of our subject pool, we will also collect nationwide measures of age, gender, race/ethnicity, income, political party identification, political ideology, the presence of children in the household, and the presence of public school students in the household from the 2016 American National Election Study.

Random Assignment

We will use the Qualtrics survey platform to administer the survey experiment. Using Qualtrics' "Randomizer" feature, roughly equal portions of the subject pool will be randomly assigned to the control group (receiving neither achievement nor growth data), the achievement group (receiving only achievement data), the growth group (receiving only growth data), or the

combination group (receiving both achievement and growth data).

Intervention

All subjects will read a description of the school district selection task:

This survey contains a thought experiment:

You will be asked to imagine that you are a parent who is about to move to a new city. When deciding where to live, one of your top priorities is to choose a school district for your elementary school-age child.

The survey will provide basic information about the five largest school districts in the area. The information comes from the Stanford University Education Data Archive. You will be asked to choose the district that best meets the needs of your family.

You will be asked to complete this thought experiment for five different cities.

Are you willing to participate in this thought experiment?

○ Yes
○ No

For each of the five metropolitan areas, subjects will be asked to make hypothetical school district choices based on the information provided. For example, subjects randomly assigned to the achievement group would see the following for the New York City metropolitan area:

Please imagine that you are a parent who is about to move to the **New York City, NY** metropolitan area. When deciding where to live, one of your top priorities is to choose a school district for your elementary school-age child. You have narrowed your search to five options. Which would you choose?

Please click the button to the LEFT of your choice.

Source: Stanford University Education Data Archive

Toms River Regional School District

○
- Average achievement in grade 3-8: The average student scores at about the 7th grade level (84th percentile nationwide)
- Median household income: $83,256
- Eligible for free/reduced price lunch: 25%
- White students: 80%
- Black students: 5%
- Hispanic students: 11%
- Asian students: 4%

Paterson School District

○
- Average achievement in grade 3-8: The average student scores at about the 4th grade level (9th percentile nationwide)
- Median household income: $31,892
- Eligible for free/reduced price lunch: 79%
- White students: 5%
- Black students: 28%
- Hispanic students: 62%
- Asian students: 4%

Jersey City Public Schools

○
- Average achievement in grade 3-8: The average student scores at about the 5th grade level (23rd percentile nationwide)
- Median household income: $41,773
- Eligible for free/reduced price lunch: 76%
- White students: 11%
- Black students: 35%
- Hispanic students: 38%
- Asian students: 15%

New York City Public Schools

○
- Average achievement in grade 3-8: The average student scores at about the 5th grade level (36th percentile nationwide)
- Median household income: $41,888
- Eligible for free/reduced price lunch: 76%
- White students: 15%
- Black students: 29%
- Hispanic students: 40%
- Asian students: 15%

Yonkers City School District

○
- Average achievement in grade 3-8: The average student scores at about the 4th grade level (10th percentile nationwide)
- Median household income: $48,095
- Eligible for free/reduced price lunch: 70%
- White students: 18%
- Black students: 23%
- Hispanic students: 54%
- Asian students: 5%

Based on subjects' experimental assignment, they will receive either 1) average achievement data, 2) average growth data, 3) both, or 4) neither. To prevent ordering effects, the sequence of school district options will be randomized for all subjects.

A copy of the complete survey instrument is available on the Open Science Framework at https://osf.io/3e8wv/.

Outcome Measures

For each metropolitan area, there will be three outcome measures: subjects' choices with respect to 1) median household income, 2) the percentage of students eligible for free and reduced priced lunch, and 3) the percentage of white students. The estimands for each metropolitan area will be the average differences in the racial and socio-economic compositions of the school district choices between experimental groups.

To estimate average treatment effects across all five metropolitan areas, we will reorganize the data into a long-form dataset. Each subject will appear in the dataset five times: once for each school district selection. The estimands for these data will also be the average differences in the racial and socio-economic compositions of the school district choices between experimental groups. When conducting these analyses, we will cluster standard errors at the subject level.

Covariate Measures

Before completing the survey, subjects will answer a small battery of demographic questions about their age, gender, race/ethnicity, income, political party identification, political ideology, the presence of children in the household, and the presence of public school students in the household.

Analyses

Experimental Balance Check: To test for balance between experimental groups, we will compare the demographic composition of the control group with the demographic compositions of each of the other randomly assigned groups. To accomplish this, we will ~~conduct a series of t-tests on each demographic variable~~ regress each demographic variable on indicators of achievement group status, growth group status, and combination group status (leaving the control group as the comparison). (10/9/18)

Representativeness Analysis: To assess the representativeness of the subject pool, we will compare the demographic composition of the whole sample to the demographic composition of the 2016 American National Election Study. Because we do not expect the subject pool to be demographically representative of the U.S. adult population, we will forgo statistical testing.

Missing Data: If subjects do not answer one or more demographic questions (all other questions are required), we will recode the missing value with "999" and control for an indicator of missingness in subsequent analyses.

Descriptive Analysis of School District Choice: We will graph the distributions of subjects' school district choices for each metropolitan area across the entire subject pool, generating descriptive information about the characteristics of the most and least popular districts.

Research Question 1: We will conduct a series of OLS regressions that take the following form: a linear regression of each outcome on indicators of achievement group status, growth group status, and combination group status (leaving the control group as the comparison). When analyzing data from all five metropolitan areas together, we will cluster standard errors at the subject level.

Research Question 2: We will conduct a series of OLS regressions that take the following form: a linear regression of each outcome on indicators of control group status, growth group status, and combination group status (leaving the achievement group as the comparison). When analyzing data from all five metropolitan areas together, we will cluster standard errors at the subject level.

Research Question 3: For the summary analyses of school district choice across all metropolitan areas, we will also conduct four additional linear regressions that will include an indicator variable and a treatment-by-covariate interaction term for each of the following:
1. The presence of children age 0-17 in the household
2. The presence of public school students age 0-17 in the household
3. Race (white and non-white)
4. Family income (under and over $75,000)

Because the analyses to answer RQ3 require a notable increase in the number of statistical tests and therefore an increase in the likelihood of false positives, their results should be viewed as speculative and secondary to the results of the RQ1 and RQ2 analyses.

We will run all RQ1 and RQ2 regressions twice: with and without a vector of all available demographic covariates (to increase statistical precision). For each estimate, we will report the results from both analyses.

To test for statistical significance, we will rely on two-tailed hypothesis tests. We will use an alpha level of 0.05.

Institutional Approval

This project has been approved by the Teachers College, Columbia University Institutional Review Board (IRB Protocol: 18-456).

This project has also been approved by the Harvard University Institutional Review Board (IRB Protocol: 18-1500). (10/3/18)

Pre-Registration

This study is registered in the AEA RCT Registry: AEARCTR-0003401. (10/9/18)