credo
CENTER FOR RESEARCH ON EDUCATION OUTCOMES

City Studies
Technical Appendix
2019

Technical Appendix

Measuring Academic Growth

This project focuses on the academic progress (growth) of enrolled and tested students in the city under study. To estimate students' academic performance, we rely on scores students received on state standardized achievement tests. Achievement tests capture what a student knows at a point in time. These test results were fitted into a bell curve format that enabled us to see how students moved from year to year in terms of academic performance. Two successive test scores allow us to see how much progress a student makes over a one-year period; this is also known as a growth score or learning gain. Growth scores allow us to zero in on the contributions of schools separately from other things that affect point-in-time scores. The parsed effect of schools in turn gives us the chance to see how students' academic progress changes as the conditions of their education transform.

With the student-level data, each subject-grade-year group of scores has slightly different mid-point averages and distributions. For end-of-course assessments (EOCs) there are only subject-year groups because EOCs are not grade specific. This means a student takes this assessment after completing the course, no matter what grade they are in. In our study, scores for all these separate tests are transformed to a common scale. All test scores have been converted to standardized scores to fit a "bell curve", in order to allow for year-to-year computations of growth.[1]

When scores are standardized, every student is placed relative to their peers in the entire state. A student scoring in the 50th percentile in Idaho receives a standardized score of zero, while a standardized score of one would place a student in the 84th percentile. Students who maintain their relative place from year to year would have a growth score of zero, while students who make larger gains relative to their peers will have positive growth scores. Conversely, students who make smaller academic gains than their peers will have negative growth scores in that year.

Model and Methods

In this study, we run regressions to predict the academic growth of students by year.

Our baseline model predicts the learning gaps between the city of interest and the state average gains. To this end, we chose schools outside of the city that represent the state average. The average growth of all students in the state is indexed to 0.00. The representative schools reside outside the city of interest and have an average growth value lower than 0.005. In addition, the proportions of white students, black students, Hispanic students, and students eligible for free or reduced-price lunch in the representative schools as a whole are similar to the proportions for all students in the state. The representative schools serve as the proxy for the state average.

---

[1] For each subject-grade-year set of scores, scores are centered around a standardized midpoint of zero, which corresponds to the actual average score of the test before transformation. Then each score of the original test is recast as a measure of variation around that new score of zero, so that scores that fall below the original average score are expressed as negative numbers and those that are larger receive positive values.

This baseline models controls for differences in the students' race, gender, free or reduced-price lunch status, English Language Learner status, special education status, grade level, grade retention status, and prior academic achievement.

In addition to the baseline model, we explore additional interactions beyond a simple binary to indicate enrollment in schools residing in the city of interest. These included both "double" and "triple" interactions between the city variable and the type of schools or student characteristics. For example, to identify the impact of schools within the city on different sector enrollment, we estimate models that break the city variable into "city_charter" and "city_TPS." We can also identify the impact of schools within the city on different racial groups by breaking out the city variable into "city_black", "city_hispanic", etc. To further break down the impact of schools in the city by sector enrollment, the variables above can be split again. For example, student in a given city's charter schools are split further into different racial groups ("city_charter_black", "city_charter_hispanic", etc).

Presentation of Results

In this report, we present the effect size for the variables of interest in terms of standard deviations. The base measures for these outcomes are referred to in statistics as z-scores. A z-score of 0 indicates the student's achievement is average for his or her grade. Positive values represent higher performance while negative values represent lower performance. Likewise, a positive effect size value means a student or group of students has improved relative to the students in the state taking the same exam. This remains true regardless of the absolute level of achievement for those students. As with the z-scores, a negative effect size means the students have on average lost ground compared to their peers.

It is important to remember that a school can have a positive effect size for its students (students are improving) but still have below-average achievement. Students with consistently positive effect sizes will eventually close the achievement gap if given enough time; however, such growth might take longer to close a particular gap than students spend in school.

While it is fair to compare two effect sizes relationally (i.e., 0.08 is twice 0.04), this must be done with care as to the size of the lower value. It would be misleading to state one group grew twice as much as another if the values were extremely small such as 0.0001 and 0.0002.

Finally, it is important to consider if an effect size is significant or not. In statistical models, values which are not statistically significant should be considered as no different from zero. Two effect sizes, one equal to .001 and the other equal to .01, would both be treated as no effect if neither were statistically significant.

To assist the reader in interpreting the meaning of effect sizes, we include an estimate of the average number of days of learning required to achieve a particular effect size. This estimate was calculated by Dr. Eric Hanushek and Dr. Margaret Raymond based on the latest (2017) 4th and 8th grade test scores from the National Assessment of Educational Progress (NAEP). Using a standard 180-day school year, each one standard deviation (s.d.) change in effect size was equivalent to 590 days of learning in this

study. The values in Table 3 are updated from past reports using more recent NAEP scores, which show slower absolute annual academic progress than earlier administrations.[2]

In order to understand "days of learning," consider a student whose academic achievement is at the 50th percentile in one grade and also at the 50th percentile in the following grade the next year. The progress from one year to the next equals the average learning gains for a student between the two grades. That growth is fixed as 180 days of effective learning based on the typical 180-day school year.

We then translate the standard deviations of growth from our models based on that 180-day average year of learning, so that students with positive effect sizes have additional growth beyond the expected 180 days of annual academic progress while those with negative effect sizes have fewer days of academic progress in that same 180-day period of time.

---

[2] Hanushek, Eric A. P.E. Peterson, & L. Woessmann (2012). Achievement Growth: International and U.S. State Trends in Student Performance. *Education Next*, Vol. 12, 1–35.