



How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR

By *Raj Chetty, John N. Friedman, Nathaniel Hilger, Diane Whitmore Schanzenbach, Emmanuel Saez, and Danny Yagan*

Introduction

Project STAR (Student/Teacher Achievement Ratio) is one of the most widely studied education experiments in the United States. It randomly assigned one cohort of 11,571 students and their teachers to different classrooms within their schools in grades K-3. Some students were assigned to small classes (15 students on average), while others were assigned to large classes (22 students on average). The experiment was implemented across 79 schools in Tennessee from 1985 to 1989.

Though numerous studies have used Project STAR to show that class size, teacher quality, and peers have significant impacts on test scores, it remains an open question whether these gains on standardized tests translate into improvements in adult outcomes such as earnings.

Because of a lack of data linking childhood education and adult outcomes, evidence on the important policy question of the long-term impacts of these achievements remains scarce. In our paper, “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR,” we fill this lack by linking the original STAR data from the 1980s to administrative data from tax returns

from the 1990s and 2000s. Thus, we are able to follow 95 percent of the STAR participants into adulthood. We use these data to analyze the impacts of STAR on outcomes ranging from college attendance and earnings to long-term impacts such as retirement savings, home ownership, and marriage.

We begin our analysis by documenting the strong correlation between kindergarten test scores and adult outcomes. We find a one percentile increase in end-of-kindergarten (KG) test scores is associated with a \$132 increase in wage earnings at age 27 in the raw data, and a \$94 increase after controlling for parental characteristics (such as income or education). Several other adult outcomes such as college attendance rates, quality of college attended, home ownership, and 401(k) savings are also all highly correlated with kindergarten test scores. These strong correlations motivate the main question of the paper: do interventions that raise test scores, such as smaller classes and better teachers, cause analogous improvements in adult outcomes?

To answer this question, we first analyze the data using these observable dimensions of classroom size and teacher quality. We find

Raj Chetty

Raj Chetty is a Professor in the Economics Department at Harvard University.

John N. Friedman

John N. Friedman is an Assistant Professor of Public Policy at the Harvard Kennedy School.

Nathaniel Hilger

Nathaniel Hilger is a Harvard University Economics Graduate Student.

Diane Whitmore Schanzenbach

Diane Whitmore Schanzenbach is an Associate Professor of Human Development and Social Policy at Northwestern University.

Emmanuel Saez

Emmanuel Saez is the E. Morris Cox Professor of Economics at UC Berkeley.

Danny Yagan

Danny Yagan is a Harvard University Economics Graduate Student.

A. Alfred Taubman Center for State and Local Government

The Taubman Center and its affiliated institutes and programs are the Kennedy School of Government's focal point for activities that address urban policy, state and local governance and intergovernmental relations.

Taubman Center Policy Briefs are short overviews of new and notable research on these issues by scholars affiliated with the Center. This policy brief is based on “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR,” forthcoming in the Quarterly Journal of Economics.

© 2011 by the President and Fellows of Harvard College. The contents reflect the views of the authors (who are responsible for the facts and accuracy of the material herein) and do not represent the official views or policies of the Taubman Center for State and Local Government.

that students assigned to small classes are more likely to be enrolled in college by age 20. Students in small classes also exhibit statistically significant improvements on a summary index of the other outcomes we examine (home ownership, 401(k) savings, mobility rates, percent college graduate in ZIP code, and marital status). We find students randomly assigned to a KG teacher with more than 10 years of experience earn an extra \$1,093 on average at age 27 relative to students with less experienced teachers.

Next, we make some analyses using the unobservable characteristics of a classroom such as the combined effects of teachers and peers in a classroom as well as class-level “shocks” such as noise outside the classroom. Because there is no variable in our data

We find that a kindergarten class of twenty students that is one standard deviation higher in quality generates, on average, over three-quarters of a million dollars more over their lifetimes (in present value earnings).

that, for instance, gives the decibel level of construction noise outside a classroom on a given day, we must estimate these class effects by exploiting the fact that each student was randomly assigned to a classroom. We use a statistical test to determine whether earnings vary across KG classes by more than we would have predicted by the random variation in student abilities. We find that this is the case, which tells us that kindergarten classroom assignment has significant impacts on earnings. In fact, we find that a kindergarten class of twenty students that is one standard deviation higher in quality generates, on average, over three-quarters of a million dollars more over

their lifetimes (in present value earnings).

Once we determined that kindergarten classroom assignment had significant impacts on earnings, we try to determine whether educational interventions that improve scores also generate earning gains. To analyze this question, we proxy for each student’s KG class quality by the average test scores of his or her classmates at the end of kindergarten. End-of-class peer test scores are a good measure of class quality because they capture both the variation in peer abilities and the effects of the teacher on students. Using this measure, we find that kindergarten class quality has significant impacts on both test scores and earnings. Students randomly assigned to a classroom that is one standard deviation higher in quality earn 3 percent more at age 27.

The findings in this paper have important implications for policy. Improving early childhood education in disadvantaged areas — .e.g. through federal tax credits or tax policy reforms — could potentially reduce inequality in the long run. Additionally, some of the long-term payoffs for better education will accrue to the government via increased tax payments and reduced dependence on welfare programs.

Experimental Design and Data

The STAR experiment was conducted at 79 schools across the state of Tennessee over four years. The program oversampled lower-income schools, and thus the STAR sample exhibits lower socioeconomic characteristics than the state of Tennessee and the U.S. population as a whole. In the 1985-86 school year, 6,323 kindergarten students in participating schools were randomly assigned to a small (target size 13-17 students) or regular-sized (20-25 students) class within their schools. Students were intended to remain in the same class type (small vs. large) through 3rd grade, at which point all students would return to regular classes for 4th grade and subsequent years.

The study design randomly assigned students not only to class type (small vs. large) but also to a classroom within each type (if there were multiple classrooms per type, as was the case in 50 of the 79 schools). Teachers were also randomly assigned to classrooms.

In each year, students were administered the grade-appropriate Stanford Achievement Test, a multiple choice test that measures performance in math and reading. We obtain these scores, as well as the data on students and their parents from federal tax forms such as 1040 individual income tax returns. Information from the 1040s is available from 1996-2008. We use third-party reports to obtain information such as wage earnings (form W-2) and college attendance (form 1098-T) for all individuals, including those who do not file 1040s. Data from these third-party reports are available since 1999. The analysis dataset combines selected variables from individual tax returns, third party reports, and information from the STAR database, with individual identifiers removed to protect confidentiality.

From STAR, along with the classroom and test score variables, we obtain the demographic variables of gender and race. We can also infer a “low income” variable, as STAR reports a student’s eligibility for a free or reduced priced lunch. From the tax and third party reports, we are able to get earnings, college attendance, college quality (by estimating average earnings of graduates of the college), marriage, retirement savings, home ownership, mobility, neighborhood quality (percent of college grads in the neighborhood), as well as parental characteristics such as income, marital status, home ownership, 401(k) savings, and age.

Impacts of Observable Classroom Characteristics

After documenting the strong correlations between kindergarten test scores and adult outcomes, we use regressions to analyze the

impacts of three features of classrooms that we can observe in our data: class size, teacher characteristics, and peer characteristics.

Class Size: We find that assigned to a small class are more likely to attend college (even when controlling for demographic characteristics). By using the earnings-based college quality measure described above, we investigate the quality of colleges these students attend. Interestingly, we find that the increased population of students attending college from the small classrooms are attending lower quality colleges. This makes sense, because we would expect that the marginal students who were induced to attend college because of reduced class size would likely apply to relatively low quality colleges. We predict that students assigned to small classes will be earning approximately \$109 more per year at age 28. Students in small classes

When we examine the impacts of teacher experience on score and earnings, we find that students who were randomly assigned to more experienced kindergarten teachers have earnings at age 27, as well as higher test scores.

score higher on our summary index (including measures home ownership, 401(k) savings, mobility rates, percent college graduate in ZIP code, and marital status), as well as individually on having a 401(k), which may imply higher job quality.

Observable teacher effects: When we examine the impacts of teacher experience on score and earnings, we find that students who were randomly assigned to more experienced KG teachers have higher earnings at age 27, as well as higher test scores. However, the impact of kindergarten teacher experience on earnings must be interpreted very carefully. Our results

show that placing a child in a kindergarten class taught by a more experienced teacher yields improved outcomes. This finding does not imply that increasing a given teacher's experience will improve student outcomes. The reason is that while teachers were randomly assigned to classrooms, experience was not randomly assigned to teachers. The difference in earnings of students with experienced teachers could be due to the intrinsic characteristics of experienced teachers rather than experience of teachers *per se*. For instance, teachers with more experience have selected to stay in the profession and may be more passionate or more skilled at teaching. The few other observable teacher characteristics in the STAR data (degrees, race, and progress on a career ladder) have no significant impact on scores or earnings.

Observable peer effects: Better classmates could create an environment more conducive to learning, leading to improvements in adult outcomes. However, we are unable to obtain good evidence of peer effects on either test scores or earnings. In particular, the randomized nature of classroom assignment implies that there is relatively little variation between classrooms in peer characteristics. Thus, the analysis is very imprecise. We cannot rule out either the case of no peer effect or the case in which peers are highly influential on adult outcomes.

Impacts of Unobservable Classroom Characteristics

Many unobserved aspects of teachers and peers could impact student achievement and adult outcomes. For instance, some teachers may generate greater enthusiasm among students or some peers might be particularly disruptive, or there might be class-level shocks such as noise outside the classroom. To test whether such unobservable aspects of class quality have long-term impacts, we use an analysis of variance approach. In particular, we test

for “class effects” on scores and earnings by exploiting random assignment to classrooms. These class effects include the effects of teachers, peers, and any class-level shocks.

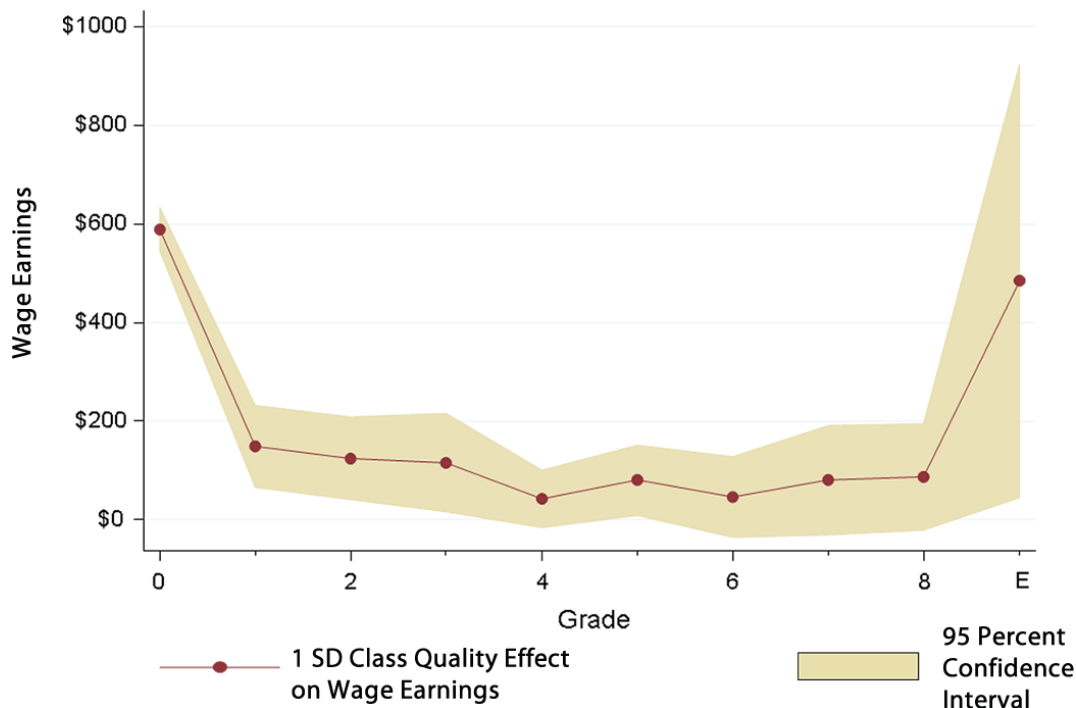
An analysis of variance approach uses a statistical model that tests if test scores and earnings are equal across classrooms. Since students are assigned to classrooms randomly, if there were no “class effects,” test scores and earnings would be the same across classrooms. However, we find that there are differences across classrooms, indicating the presence of class effects. For instance, assigning students to a classroom that is one standard deviation better than average quality in kindergarten generates an increase in earnings at ages 27 of 9.6 percent (\$1,520) per year (including demographic controls). Even when we control for all observable classroom characteristics (class size, teacher experience, and so on), we find that a one standard deviation in quality is associated with an earnings increase of \$1,424 per year. Further research on which factors contribute to a high “class quality” would be extremely valuable in light of these results.

Assigning students to a classroom that is one standard deviation better than average quality in kindergarten generates an increase in earnings at ages 27 of 9.6 percent (\$1,520) per year.

Fade-Out, Re-Emergence, and Non-Cognitive Skills

The impacts of early childhood class assignment on outcomes more than twenty years later may be particularly surprising because the impacts on test scores fade out rapidly. Previous studies on the impacts of class size on test scores become insignificant by grade 8 (Krueger and Whitmore 2001). Figure 1 demonstrates this “fade-out” and “re-

Figure 1: Fadeout of Class Effects of 1 Standard Deviation of Class Quality on Earnings



emergence” effect visually. Using regression analysis, we estimate the class quality effect from attending a better KG class on student test scores in each grade (K-8), and then calculate how these test scores translate into earnings. Figure 1 shows that the predicted earnings effect in KG is nearly \$600, but by 4th grade, we predict less than a \$50 per year gain in earnings. However, the final point in the figure (“E”), shows the actual observed earnings impact of a one standard deviation improvement in KG class quality. This number (\$483) is very close to the earnings gain predicted by KG test scores (\$588). Why do the impacts of early childhood education fade out on test scores but re-emerge in adulthood?

We find some suggestive evidence that part of the explanation may be non-cognitive skills. We find that KG class quality has significant impacts on non-cognitive measures in 4th and 8th grade such as effort, initiative, and lack of disruptive behavior. These non-cognitive measures are highly correlated with earnings even conditional on test scores but are not significant predictors of future standardized test

scores. These results suggest that high quality KG classrooms may build non-cognitive skills that have returns in the labor market but do not improve performance on standardized tests. While this evidence is far from conclusive, it highlights the value of further empirical research on non-cognitive skills.

Cost-Benefit Analysis and Policy Implications

We conclude by using our empirical estimates to provide rough calculations of the benefits of various policy interventions. These cost-benefit calculations rely on several strong assumptions, including the assumption that the percentage gain in earnings observed at age 27 remains constant over the lifecycle. We ignore non-monetary returns to education (such as reduced crime) as well as general equilibrium effects and discount earnings gains at a 3 percent annual rate back to age 6 (kindergarten age), the point of the intervention. Under these assumptions, we calculate the present-value earnings gains for a classroom of 20 students from three interventions: improvements in

classroom quality, reductions in class size, and improvements in teacher quality.

Based on our analysis, increasing class quality by one standard deviation of the distribution within schools raises earnings by \$1,520 (9.6 percent) at age 27. Under the preceding assumptions, this translates into a lifetime earnings gain of approximately \$39,100 for the average individual. For a classroom of twenty students, this implies a present-value benefit of \$782,000 for improving class quality for a single year by one (within-school) standard deviation. This large figure includes all potential benefits from an improved classroom environment, including better peers, teachers, and random shocks, and hence is useful primarily for understanding the stakes at play in early childhood education. It is less helpful from a policy perspective because one cannot implement interventions that directly improve classroom quality. This motivates the analysis of class size and better teachers, two factors that contribute to classroom quality.

Present value earnings gain from a 33 percent class-size reduction is \$9,460 per student and \$189,000 for the classroom.

We calculate the benefits of reducing class size by 33 percent in two ways. The first is estimate is imprecise and less useful, but in the second we estimate that a 1 percentile increase in class quality raises test scores by 0.66 percentiles and earnings by \$50.6, implying an earnings gain of \$76.7 per percentile increase in test scores. We make the strong assumption that the ratio of earnings gains to test score gains is the same for changes in class size as it is for improvements in class quality more generally. Under this assumption, a 33 percent class size reduction in grades K-3 (which raised test scores by 4.8 percentiles) is predicted to raise

earnings by $4.8 \times \$76.7 = \368 (2.3 percent) at age 27. This calculation implies a present value earnings gain from class size reduction of \$9,460 per student and \$189,000 for the classroom.

Previous studies have estimated that a one standard deviation increase in teacher quality raises test scores by approximately 0.2 standard deviations (5.4 percentiles). Under the strong assumption that the ratio of earnings gains to test score gains is the same for changes in teacher quality and class quality more broadly, this translates into an earnings gain of $5.4 \times \$76.7 = \416 (2.6 percent) at age 27 and hence a present-value earnings gain of \$10,700 per student. Hence, a one standard deviation improvement in teacher quality in a single year generates earnings gains of \$214,000 for a classroom of twenty students.

The magnitude of the estimated impacts of teachers on earnings suggests that good teachers can create great social value, perhaps several times larger than current teacher salaries. However, our results do not have direct implications for optimal teacher salaries or merit pay policies as we do not know whether higher salaries or merit pay would improve teacher quality.

Relative to efforts that seek to improve the quality of teachers, class size reductions have the important advantage of being more well-defined and straightforward to implement. Our findings on the importance of teacher quality caution that reductions in class size must be implemented carefully to generate improvements in outcomes. If schools are forced to reduce teacher and class quality along other dimensions when reducing class size, the net gains from class size reduction may be diminished (Jepsen and Rivkin 2009, Sims 2009).

Conclusion

The impacts of education have traditionally been measured by achievement on standardized tests. This paper has shown that many of the interventions that raise test scores also improve long-term outcomes. Students who were randomly assigned to higher quality classrooms in grades K-3 earn more, are more likely to attend college, save more for retirement, and live in better neighborhoods. Yet the same students do not do much better on standardized tests in later grades, due to fade-out. These results suggest that policy makers may wish

Students who were randomly assigned to higher quality classrooms in grades K-3 earn more, are more likely to attend college, save more for retirement, and live in better neighborhoods.

to rethink the objective of raising test scores and evaluating interventions via long-term test score gains. Researchers who had examined only the impacts of STAR on test scores would have incorrectly concluded that early childhood education does not have long-lasting impacts. While the quality of education is best judged by directly measuring its impacts on adult outcomes, our analysis suggests that contemporaneous (end-of-year) test scores are a reasonably good short-run measure of the quality of a classroom.

Finally, our analysis raises the possibility that differences in school quality perpetuate income inequality. In the U.S., higher income families have access to better public schools on average because of property-tax finance. Using the class quality impacts reported above, Chetty and Friedman (2011) estimate that the intergenerational correlation of income would fall by roughly 1/3 if all children attended schools of the same quality. It is

possible that federal tax credits or tax policy could help improve early childhood education in disadvantaged areas, leading to reduced inequality in the long run. This is an important area for further study.

REFERENCES

Chetty, Raj and John N. Friedman. "Does Local Tax Financing of Public Schools Perpetuate Inequality?" Forthcoming, *National Tax Association Proceedings* (2011).

Jepsen, Christopher and Steven Rivkin. "Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size," *Journal of Human Resources*, 44:1 (2009) 223-250.

Krueger, Alan B., and Diane M. Whitmore. "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR," *The Economic Journal*, 111 (2001), 1-28.

Sims, David. "Crowding Peter to Educate Paul: Lessons From a Class Size Reduction Externality," *Economics of Education Review* 28:4 (2009), 465-473.

RECENT POLICY BRIEFS

“Which Places are Growing? Seven Notable Trends from Newly Released Census Data,” by Edward Glaeser (Harvard University) March 2011.

“How Massachusetts Can Stop the Public-Sector Virus,” by Thomas A. Kochan (MIT Sloan School of Management) March 2011.

“Did Credit Market Policies Cause the Housing Bubble?” by Edward Glaeser (Harvard University), Joshua Gottlieb (Harvard University), and Joseph Gyourko (Wharton School) May 2010.

“Geography, Venture Capital, and Public Policy,” by Josh Lerner (Harvard Business School) March 2010.

“What Makes a City Entrepreneurial?” by Edward L. Glaeser (Harvard University) and William R. Kerr (Harvard Business School) February 2010.

“What Good Is Wealth Without Health? The Effect of Health on the Satisfaction Derived from Consumption,” by Amy Finklestein, Erzo F. P. Luttmer, and Matthew J. Notowidigdo (Harvard Kennedy School) May 2009.

“Improving Health Care Quality and Values: Local Challenges and Local Opportunities,” by Katherine Baicker (Harvard School of Public Health) and Amitabh Chandra (Harvard Kennedy School) May 2009.

“Urban Inequality,” by Edward Glaeser, Matt Resseger, and Kristina Tobio (Harvard University) March 2009.

“Inefficient Cities,” by Jose A. Gomez-Ibanez (Harvard Kennedy School) and Fernanda Ruiz Núñez (World Bank) March 2009.

“Silver Bullet or Trojan Horse? The Effects of Inclusionary Zoning on Local Housing Markets in Greater Boston,” by Jenny Schuetz, Rachel Meltzer, and Vicki Been (Furman Center for Housing Studies) March 2009.

“Searching the Family Tree for Suspects: Ethical and Implementation Issues in the Familial Searching of DNA Databases,” by David Lazar (Harvard Kennedy School) March 2008.

“The Greenness of Cities,” by Edward L. Glaeser

(Harvard University) and Matthew Kahn (UCLA) March 2008.

“The Seven Big Errors of PerformanceStat ,” by Robert D. Behn (Harvard Kennedy School) February 2008.

“Flypaper and Fungibility: Evidence from the Master Tobacco Settlement,” by Monica Singhal (Harvard Kennedy School) May 2007.

“The Rise of the Sunbelt,” by Edward L. Glaeser (Harvard University) and Kristina Tobio (Harvard Kennedy School) May 2007.

“Productivity Spillovers in Health Care” by Amitabh Chandra (Harvard Kennedy School) May 2007.

“High Performance in Emergency Preparedness and Response: Disaster Type Differences” by Herman B. “Dutch” Leonard and Arnold M. Howitt (Kennedy School of Government) May 2007.

“Racial Statistics and Race-Conscious Public Policy,” by Kim M. Williams (Kennedy School of Government) May 2007.

“Transparency Policies: Two Possible Futures” by Archon Fung, Mary Graham, David Weil, and Elena Fagotto (Transparency Policy Project) May 2007.

“Beyond Katrina: Improving Disaster Response Capabilities” by Arnold M. Howitt and Herman “Dutch” Leonard (Kennedy School of Government)

“Why are Smart Cities Getting Smarter?” by Edward L. Glaeser (Harvard University) and Christopher Berry (University of Chicago)

“The ‘Third Way’ of Education Reform?” by Brian Jacob (Kennedy School of Government)

“From Food to Finance: What Makes Disclosure Policies Work?” by Archon Fung, Mary Graham, David Weil, and Elena Fagotto (Kennedy School of Government)

The Fiscal Crisis of the States: Recession, Structural Spending Gap, or Political “Disconnect”? by Robert Behn and Elizabeth Keating (Kennedy School of Government)

“Smart Growth: Education, Skilled Workers and the Future of Cold-Weather Cities,” by Edward L. Glaeser (Harvard University)