Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion

Susan Dynarski*, Joshua Hyman†
and Diane Whitmore Schanzenbach‡

September 29, 2011

Abstract

This paper examines the effect of early childhood investments on college enrollment and degree completion. We use the random assignment in the Project STAR experiment to estimate the effect of smaller classes in primary school on college entry, college choice, and degree completion. We improve on existing work in this area with unusually detailed data on college enrollment spells and the previously unexplored outcome of college degree completion. We find that assignment to a small class increases the probability of attending college by 2.7 percentage points, with effects more than twice as large among blacks. Among those whose predicted probability of attending college is in the bottom quintile, smaller classes increase the college attendance rate by 11 percentage points. Smaller classes increase the likelihood of earning a college degree by 1.6 percentage points and shift students towards high-earning fields such as STEM (science, technology, engineering and medicine), business and economics. We confirm the standard finding that test score effects fade out by middle school, but show that test score effects at the time of the experiment are an excellent predictor of long-term improvements in postsecondary outcomes. We compare the costs and impacts of this intervention with other tools for increasing postsecondary attainment, such as Head Start and financial aid, and conclude (contrary to Heckman's assertion) that early investments are no more cost effective than later investments in boosting adult educational attainment.

*Associate Professor of Economics, Education and Public Policy, Gerald R. Ford School of Public Policy and School of Education, University of Michigan, dynarski@umich.edu.

†Doctoral Candidate, Department of Economics and Gerald R. Ford School of Public Policy, University of Michigan, jmhyman@umich.edu.

‡Associate Professor of Human Development and Social Policy, School of Education and Social Policy, Northwestern University, dws@northwestern.edu.

1 Introduction

Education is intended to pay off over a lifetime. Economists conceive of education as a form of "human capital," requiring costly investments in the present but promising a stream of returns in the future. Economists looking backward at a number of education interventions (e.g., Head Start, compulsory schooling) have identified causal links between these policies and long-term outcomes such as adult educational attainment, employment, earnings, health and civic engagement (Ludwig and Miller, 2007; Deming, 2009; Angrist and Krueger, 1991; Dee, 2004; Lleras-Muney, 2005). But decision-makers attempting to gauge the effectiveness of current education inputs, policies and practices in the present can't wait decades for these long-term effects to emerge. They therefore rely upon short-term outcomes – primarily standardized test scores – as their yardstick of success.

A critical question is the extent to which short-term improvements in test scores translate into long-term improvements in well-being. This is the question we address in this paper. Puzzling results from several evaluations make this a salient question. Three small-scale, intensive preschool experiments produced large effects on contemporaneous test scores that quickly faded (Schweinhart et al., 2005; Anderson, 2008). Non-experimental evaluations of Head Start, a preschool program for poor children, reveal a similar pattern, with test score effects gone by middle school. In each of these studies, treatment effects have re-emerged in adulthood as increased educational attainment, enhanced labor market attachment, and reduced crime (Deming, 2009; Garces et al., 2002; Ludwig and Miller, 2007). Further, several recent papers have shown large impacts of charter schools on test scores of disadvantaged children (Abdulkadiroglu et al., 2011; Angrist et al., 2010; Dobbie and Fryer, 2011). A critical question is whether these effects on test scores will persist in the form of long-term enhancements to human capital and well-being.

We examine the effect of smaller classes on educational attainment in adulthood, including college attendance, degree completion and field of study. We exploit random variation in class size in the early grades of elementary school created by the Tennessee Student/Teacher Achievement Ratio (STAR) Experiment. Participants in the STAR experiment are now in their thirties, an age at which it is plausible to measure completed education. Our postsecondary outcome data is obtained from the National Student Clearinghouse (NSC), a national database that covers approximately 90% of students enrolled in colleges in the U.S.

We find that attending a small class increases the rate of postsecondary attendance by 2.7 percentage points. Black students and students eligible for free lunch show the largest impacts, 5.8 and 4.4 percentage points, respectively. Attending a small class increases the chance of earning a postsecondary degree by 1.6 percentage points. These increases in degree receipt are driven exclusively by gains in degrees in high-earning fields such as science, technology, engineering and mathematics (STEM), business and economics.

Our results shed light on the relationship between the short- and long-term effects of educational interventions. We find that the short-term effect of a small class on test scores is an excellent predictor of its effect on adult educational attainment. In fact, the effect of small classes on college attendance is completely captured by their positive effect on contemporaneous test scores.[1] We further find that the relationship between scores and postsecondary attainment is the same in small and regular classes; that is, the scores of children in the small classes are no less (or more) predictive of adult educational attainment than those of children in the regular classes. We can, in fact, closely predict the effect of STAR on postsecondary attainment by combining information about the relationship between scores and attainment from an outside dataset (the CNLSY, Children of the National Longitudinal Survey of Youth) with the impact of STAR on contemporaneous scores.

Another contribution of this paper is to identify the effect of manipulating a single educational input on adult educational attainment. The early-childhood interventions for which researchers have identified lifetime effects (e.g., Head Start, Abecederian) are intensive and multi-pronged, including home visits, parental coaching and vaccinations in addition to time in a preschool classroom. We cannot distinguish *which dimensions* of these treatments generate short-term effects on test scores, and whether they differ from the dimensions that

[1]We show this by adding K-3 test scores to our identifying equation; the coefficient on the class size dummy drops to zero. The coefficient on an interaction of class size and test scores is also zero, indicating that scores are equally predictive of educational attainment in small and regular classes.

generate long-term effects on adult well-being. The effective dimensions of the treatment are also confounded in a recent paper (Chetty et al., forthcoming) that estimates (using the STAR data) very large effects of kindergarten classroom assignment on adult well-being. In that analysis, the "treatment" that produces significant variation in adult outcomes excludes random assignment to small vs. regular classes, but does include anything else that varies at the classroom level, such as teacher quality and peer quality. By contrast, the effects we measure in this paper, both short-term and long-term, can be attributed to a well-defined and replicable intervention: reduced class size.

2 The Tennessee STAR Experiment

The Tennessee Student/Teacher Achievement Ratio (STAR) Experiment randomly assigned class sizes to children in kindergarten through third grade.[2] The experiment was initiated in the 1985-86 school year, when participants were in kindergarten. A total of 79 schools in 42 school districts participated, with over-sampling of urban schools. An eventual 11,571 students were involved in the experiment. The sample is 60% white and the balance African American. About 60% of students were eligible for subsidized lunch during the experiment.

Children in the STAR experiment were assigned to either a small class (target size of 13-17 students) or regular class (22-25 students).[3] Students who entered a participating school after kindergarten were randomly assigned during those entry waves to a regular or small class. Teachers were also randomly assigned to small or regular classes. All randomization occurred within schools.

Documentation of initial random assignment in STAR is incomplete (Krueger, 1999). Krueger (1999) examines records from 18 STAR schools for which assignment records are available. He finds that, as of entry into STAR, 99.7% of students were enrolled in the experimental arm to which they were initially assigned. Krueger's approach, and that of the

[2] The experiment is described in detail in Word et al. (1990), Folger and Breda (1989), Finn and Achilles (1990), Krueger (1999) and Achilles (1999).

[3] A third arm of the experiment assigned children to a regular class with a teacher's aide. Previous research has shown no difference in outcomes between the regular-sized classes with and without an aide. We follow the previous literature in pooling students from both types of regular classes into a single control group.

subsequent literature, is to assume that the class type in which a student is first enrolled is the class type to which she was assigned. We follow that convention in our analysis.

Numerous papers have tested, and generally validated, the randomization in STAR (Krueger, 1999). There are no baseline outcome data (e.g., a pre-test) available for the STAR sample. On the handful of covariates available in the STAR data (free lunch eligibility, race, sex), the arms of the experiment appear balanced at baseline (see Table 1 for a replication of these results). Recent work by Chetty et al. (forthcoming) shows that the STAR entry waves were balanced at baseline on an expanded set of parental characteristics that they obtain by matching the STAR data to income tax returns for STAR subjects and their parents.

2.1 Previous Research on the Long-Term Effects of Small Classes

A substantial body of research has examined the effect of Project STAR on short- and medium-run outcomes; we do not comprehensively discuss this literature but instead summarize the pattern of findings. These papers show that students assigned to a small class experience contemporaneous test score gains of about a fifth of a standard deviation. These test score results fade after the experiment ends in third grade. There is, however, evidence of lasting effects on other dimensions. Krueger and Whitmore (2001) show that students assigned to small classes are more likely to take the ACT and SAT, required for admission to most four-year colleges. Schanzenbach (2007) shows that smaller classes reduce the rate of teen pregnancy by about a third.

A recent paper examined the effect of Project STAR on adult outcomes. Chetty et al. (forthcoming) match the STAR participants to their and their families' income tax returns, which include information on income, home ownership, and tuition paid to postsecondary institutions. They find that students assigned to small classes are more likely to be enrolled in college at age 20, but that this advantage erodes and becomes insignificant as students age. As we show later, this null finding is driven by measurement error in their college attendance variable, which is derived from data that colleges send to the Internal Revenue Service to verify eligibility for the Hope and Lifetime Learning tax credits and the tuition tax deduction. Chetty et al. (forthcoming) do show a large effect of kindergarten *classroom assignment* on several adult outcomes (e.g., income, home ownership and savings). This relationship, the focus of their paper, is not identified by random assignment to small vs. regular classes but rather by random variation *within* the arms of the experiment in all other classroom characteristics, including teacher quality and peer quality. The research and policy implications of that finding are therefore quite distinct from that of the present analysis, which identifies the effect of manipulating a single dimension of the education production function: class size.

3 Empirical Strategy

In this section we describe our empirical strategy and the data that we use to execute it.

3.1 Estimating Equation

The experimental nature of Project STAR motivates the use of a straightforward empirical specification. We compare outcomes of students assigned to small and regular classes by estimating the following equation using Ordinary Least Squares:

$$y_{isg} = \beta_0 + \beta_1 SMALL_{is} + \beta_2 X_{is} + \beta_{sg} + \epsilon_{isg} \tag{1}$$

where y_{isg} represents a postsecondary schooling outcome of student i, who entered the STAR experiment in school s and in grade g. X is a vector of covariates including sex, race and free lunch status, included to increase precision. β_{sg} is a set of school-by-entry-grade fixed effects. We include these because students who entered STAR schools after kindergarten were randomly assigned at that time to small or regular classes. The variable of interest is $SMALL_{is}$, an indicator set to one if student i was assigned to a small class upon entering the experiment. The omitted group to which small classes are compared is regular classes (with or without a teacher's aide).

We cluster standard errors by school, the most conservative approach. Standard errors are about ten percent smaller if we cluster at the level of school-by-wave.

3.2 Data

We use the original data from the STAR experiment, which includes information on the type of class in which a student is enrolled, basic demographics (race, poverty status, sex), school identifiers, and standardized test scores. These data also include the name and date of birth of the student, which we use to match to data on postsecondary attainment and completion, which we next describe.

3.2.1 Matching STAR to National Student Clearinghouse Data

Data on postsecondary outcomes for the STAR sample come from the National Student Clearinghouse (NSC). NSC is a non-profit organization that was founded to assist student loan companies in validating students' college enrollment. Borrowers can defer payments on most student loans while in college, which makes lenders quite interested in tracking enrollment. Colleges submit enrollment data to NSC several times each academic year, reporting whether a student is enrolled, at what school, and at what intensity (e.g., part-time or full-time). NSC also records degree completion and the field in which the degree is earned. States and school districts use NSC data to track the educational attainment of their high school graduates (Roderick et al., 2006). Recent academic papers making use of NSC data include Deming et al. (2011) and Bettinger et al. (2009).

With the permission of the Project STAR researchers and the state of Tennessee, we submitted the STAR sample to the NSC in 2006 and again in 2010. Since the STAR sample was scheduled to graduate high school in 1998, we capture college enrollment and degree completion for twelve years after on-time high-school graduation, when the STAR sample is about 30.[4]

The NSC matches individuals to its data using name and date of birth. If birth date is

[4] In 2006, when we first submitted the data, the NSC used social security number as well as name and date of birth in its matches. As of 2010, when we refreshed the match, NSC had ceased to use social security number for its matches.

missing, the NSC attempts to match on name alone. Some students in the STAR sample are missing identifying information used in the NSC match: 10% have incomplete name and nearly a quarter are missing name or birthdate. In our data, a student that attends college but fails to produce a match in the NSC database is indistinguishable from a student who did not attend college. If the absence of these identifiers is correlated with the treatment, then our estimates may be biased. To check on this, we regressed a dummy indicating missing identifiers against our main estimating equation. The results indicate that the probability of missing identifying information is uncorrelated with initial assignment.

3.2.2 Coverage Rate of NSC Data

Not all schools participate in NSC; the company estimates they currently capture about 92% of undergraduate enrollment nationwide. During the late 1990s, when the STAR subjects would have been graduating from high school, the NSC included colleges enrolling about 80% of undergraduates in Tennessee. We calculate this rate by dividing undergraduate enrollment at Tennessee colleges included in NSC as of 1998 by enrollment at all Tennessee colleges.[5] We obtain enrollment data from the Integrated Postsecondary Education Data System (IPEDS), a federally-generated database that lists every college, university and technical or vocational school that participates in the federal financial aid programs (about 6,700 institutions nationwide) (Statistics, 2010).

Since we miss about 20% of undergraduate enrollment using the NSC data, we expect that we will underestimate the college attendance rate of the STAR sample by about a fifth. Our matched NSC-STAR data indicate that 39.4% of the sample had attended college by age 30. According to the 2005 American Community Survey, among those born in Tennessee in the same years as the STAR cohort, 52.8% had ever attended college (Ruggles et al., 2010).[6] Our NSC estimate of college attendance is, as expected, about four-fifths of the magnitude of the ACS estimate.

[5] The list of all colleges participating in the NSC and the year that they joined was accessed on September 1, 2010 from http://www.studentclearinghouse.org/colleges/coreserv/docs/CoreParticipants.xls.

[6] We re-weight the Tennessee-born in the ACS data to match the racial composition of the STAR sample, which was disproportionately black

To determine whether the NSC systematically misses certain types of schools, we compare the schools that participate in NSC with those in IPEDS. Along all measures we examined (i.e., sector, racial composition, selectivity), the NSC colleges are similar to the universe of IPEDS colleges, with a single exception: NSC tends to exclude private, less-than-4-year colleges. These are primarily trade schools such as automotive, technology, business, nursing, culinary arts and beauty schools. If small classes tend to induce into such schools those students who would not otherwise attend college, we will underestimate the effect of small classes on college attendance; this bias will be the largest for those who tend to attend such colleges (e.g., low-income and nonwhite students).

4 Results

In this section, we examine the effect of assignment to a small class on a series of postsecondary outcomes: college entry, the timing of college entry, college choice, degree receipt and field of degree.

4.1 College Entry

In Table 2, we estimate the effect of assignment to a small class on the probability of college entry by age 30. The effect is close to three percentage points (Column 1, 2.8 percentage points), which is large relative to the control mean of 38.5% (control means are italicized). This estimate is statistically significant, with a standard error of about one percentage point. Including covariates does not alter the estimate, as is expected with random assignment. For the balance of the paper we report results that include covariates, since they are slightly more precise. Splitting the sample by race reveals that the effects are concentrated among Blacks (5.8 points, mean is 30.8%) and those eligible for free and reduced-price lunch (4.4 points, mean is 27.2%). The effects are twice as large for boys (3.2 points, mean is 32.4%) than girls (1.6 points, mean is 45.5%). [7]

[7] Breaking the effects down yet more finely shows that the effects are largest for Black females (7.2 points), with no effect on white females. The effects for Black and white males are indistinguishable (3 and 4 points,

Class size could plausibly affect the intensity with which a student enrolls in college in addition to the decision to enroll at all. The overall effect on the intensity of enrollment is theoretically ambiguous: students induced into college by smaller classes may be more likely to enroll part-time than other students, while the treatment could induce those who would have otherwise enrolled part-time to instead enroll full-time. In the control group, about three-quarters of college entrants (ever) attend college full-time, while a quarter never do (Table 2, second row). When we interact our college entry outcome with these two measures of enrollment intensity, we find that the effect on entry is evenly divided between part-time and full-time enrollment. While the standard errors preclude any firm conclusions, these results suggest that the marginal college student is more likely than the infra-marginal student to attend college exclusively on a part-time basis.

4.2 Timing of College Atendance

Class size could plausibly affect the timing of postsecondary attendance. Again, the net effect is theoretically ambiguous. Better-prepared students may advance through high school more rapidly, entering college sooner after graduation than their peers. They may also move through college and complete a degree more quickly. Students induced into college by smaller classes may, by contrast, delay college entry and move through college at a slower pace, slowed (for example) by remedial coursework.

We examine the effect of smaller classes on both the timing of college entry and persistence. We define "on-time enrollment" as entering college by the fall of 1999, or about 18 months after the STAR cohort is scheduled to have graduated high school. This variable captures the pace at which students complete high school, how quickly they enter college, and whether they attend at all. By this measure, 27.5% percent of the control group has enrolled on-time, or about three-quarters of the 38.5% who ever attend college. This rate is 2.4 percentage points higher in the treatment group, and the difference is statistically significant, with a standard error of 1.0 percentage point. These results suggest that students in smaller classes

respectively).

are no less likely to to start college on time than control students: 73% (=29.9/41.2) of the treated students who attend college do so on time, while among the controls the share of attendance that is on-time is 71% (=27.5/38.5).

We now look at the year-by-year evolution of the effect of class size on postsecondary attainment. For each year, we plot the share of students who have ever attended college, separately for the treatment and control group (Figure I, left panel). We also plot the treatment-control difference, along with its 95% confidence interval (Figure I, right panel). The fraction of the sample that has ever attended college (Figure I, left side) rises from under 5% in 1997 to over 20% in 1998 (when students are 18), peaking around 25% in 1999 (when students are 19). The rate rises slowly through age 30, when the share of the sample with any college experience reaches nearly 40%. The difference between the two groups (right panels) reaches about three points by age 19 and remains at that level through age 30. The top panel consists of raw means, the second includes school -by-wave fixed effects, while the last adds in the demographic controls we include in our preferred specification.[8]

When we turn to the shares of students who are currently enrolled in college (Figure II) we see that the treatment group is more likely to be enrolled in college at every point in time. Plausibly, smaller classes could have sped up college enrollment and completion, and the control group could eventually have caught up with the treatment group in its rate of college attendance. This is not what we see, however. The effect is always positive, and is largest right after high school, when the sample is 18 to 19 years old.

A recent paper finds smaller effects of STAR on college attendance than do we. Chetty et al. (forthcoming) find a statistically significant impact of assignment to a small class on the probability of attending college at age 20 (2 percentage points), but this drops to an insignificant 1.6 points by age 27. Chetty et al. impute college attendance from 1098-T forms, which colleges send to the Internal Revenue Service to confirm that tuition has been paid for a given student. IRS provided these forms for 1999 through 2007; our data capture

[8]To obtain the figures that include controls, we replace the small-class dummy in our identifying equation with a full set of its interactions with year dummies. The coefficients on these interactions and their confidence intervals are plotted on the right side. On the left side, we add these interactions to the year-specific control means.

enrollment from 1995 to 2010. These differences in scope of measurement drive the divergence in results. When censor the NSC data so that it excludes the same enrollment spells that are unobserved in the IRS data, our estimate drops from 2.7 to 1.6 percentage points - identical to that of Chetty et al.

4.3 College Choice

By boosting academic preparation, smaller classes in primary school may induce students to alter their choice of college. In particular, those who would have otherwise attended a two-year community college may instead choose to attend a four-year institution. Attending higher quality schools, which provide more inputs (including better peers) may be one mechanism through which students could increase their rate of degree completion.

In Table 3, we examine the effect of class size on college choice. We find little evidence that exposure to smaller classes shifts students toward higher-quality schools. The treatment effect is concentrated on attendance at two-year institutions. While 22 percent of the control group starts college at a two-year school, the rate is 2.5 percentage points higher in the treatment group (standard error is 0.9 percentage points). We find positive but imprecise effects on the probability of ever attending a four-year college or attending college out of state. We have also examined the effect of class size on the selectivity of the school attended (results not shown) but find no significant impacts.

4.4 Persistence and Degree Completion

While college entry has been on the rise in recent decades, the share of college students that graduates is flat or perhaps declining (Bound et al., 2009). About half of college entrants never earn a degree. A key concern is that marginal students attending college may drop out quickly, in which case the attendance effects estimated above would produce little in the way of social welfare.

We explore this issue by examining the effect of small classes on the number of semesters that students attend college, as well as on the probability that they complete a college degree.

Overall, the number of semesters attempted is quite low: the control group attempts an average of three semesters by age 30. This figure is weighed down by zeroes assigned to those who never attempt college: among those in the control group with any college experience, the average number of semesters attempted is eight.

The treatment group spends 0.22 more semesters in college than the control group (Figure III, top; Table 2). The estimate is marginally significant at a 90 percent level of confidence. The magnitude of the estimate is comparable to effects found in the Opening Doors demonstration, which gave short-term rewards to community college students for achieving certain enrollment and grade thresholds (Barrow et al., 2009).

Note that the NSC data measure semesters attended, and not credits earned. Besides semesters attended, our best available measure of persistence in college is degree completion. Among those in the control group, 15.1% earn a college degree. We find (Table 4) that those assigned to a small class are 1.6 percentage points more likely to earn a college degree. There is a positive treatment effect at every age, with the difference largest between age 22 and 23 (Figure V, top). The sharp peak in those earning a degree at those ages is driven by BA attainment (bottom); AA attainment is more evenly distributed along the age distribution. Figure IV plots the cumulative share of the sample that has earned any college degree (top), only an AA (middle), and a BA or higher (bottom).

Figure V shows that students assigned to small classes continue to outpace those in regular classes in earning degrees well into their late twenties. This pattern likely explains why Chetty et al. (forthcoming) do not find an effect of small classes on earnings, which they measure at age 27. Since members of the treatment group are still attending and completing college at this age, they have likely not yet spent enough time in the labor market for their increased education to offset experience lost while in college.

4.5 Field of Degree

A large literature has documented that earnings of college graduates differ considerably by field. In particular, those who study science, technology, engineering and medicine (STEM), as well as business and economics, enjoy higher returns than other college graduates (Arcidiacono, 2004; Hamermesh and Donald, 2008). In this section we examine whether class size affects the field in which a student completes a degree. We have coded degrees into two broad fields: STEM, business and economics; and all others.[9] Students can earn more than one degree (e.g., an A.A. and a B.A.); we code them as having a STEM, business or economics degree if any degree is in one of these fields.

Assignment to a small class shifts degrees earned toward STEM fields, business and economics. While 4.4 percent of the control group earns a degree in a STEM, business, or economics field, the rate is 5.7 in the treatment group (Table 4). This difference is statistically significant at the 5 percent level, with a standard error of 0.6 percentage points.[10] There is no difference in the rate at which students receive degrees in the remaining fields. These results are consistent with (at least) two scenarios. One is that marginal degree-completers tend to concentrate in STEM, business or economics. An alternative scenario is that the treatment induces inframarginal degree-completers to shift toward STEM, business and economics. Since we cannot identify the marginals, we cannot distinguish between these two scenarios.

5 Heterogeneity in Effects

Inequality in postsecondary education has increased in recent decades, with the gap in attendance between those born into lower-income and higher-income families expanding (Bound et al., 2009; Bailey and Dynarski, 2011). In this section, we examine how reduced class size affects inequality in postsecondary attainment. We examine whether class size reduction improves outcomes for those groups who historically have low levels of postsecondary attainment and degree completion: blacks, poor children and boys.

[9] We collapse fields into our two broad categories following schemes provided by the National Science Foundation. We use two variables provided by NSC: degree title (e.g., "associates" or "bachelor of science") and college major (e.g., "biology"). A small number of students who receive a degree are missing both degree title and college major; they are excluded from this analysis.

[10] When we separate STEM fields from business and economics fields, we find that the effects are driven equally by increases across both fields.

5.1 Effect of Class Size on Gaps in Educational Attainment

Smaller classes reduce inequality in rates of college entry across socioeconomic groups. In Table 2, we showed that assignment to a small class increased the probability of attending college by age 30 by 2.7 percentage points. Looking across the columns of Table 2, we see that the effect of class size on college attendance varies considerably (Figure VI depicts these effects graphically). We can also see that, in every case, the treatment effects are largest for the groups with the lowest control mean. The effect of assignment to a small class on black students is 5.8 percentage points, more than five times the effect on whites. The effect is larger for children eligible for free lunch (4.4 vs. 1.0 percentage points). The effects are twice as large for boys as for girls (3.2 vs. 1.6 percentage points).[11]

The pattern of effects just described will tend to decrease gaps in postsecondary attainment. Figure VII shows this graphically. On the top is depicted the gap in college attendance between blacks and whites in regular classes (left) and in small classes (right). The black-white gap is about half as large in small classes (7.7 percentage points) as it is in regular classes (12.4 percentage points). The income gap in college attendance in the control group is astoundingly large: 29.1 percentage points. It is slightly smaller in the treatment group (25.7 percentage points). Figure VIII shows the effect of small classes on the race gap in college attendance separately for males and females. We see that the drastic reduction in the race gap in college attendance is driven by females, for whom the race gap virtually disappears in small classes.

5.2 Heterogeneity in Treatment Effect or in Treatment Dosage?

One interpretation of these results is that the groups with the lowest control means are most sensitive to class size. An alternative interpretation, however, is that the groups that display the largest response are actually exposed to a more intense dosage of the treatment. All of our estimates so far have been of the effect of the intention to treat (ITT), which is attenuated

[11]The subgroup effects on semesters attempted are imprecisely estimated but suggest that effects are twice as large among blacks as among whites and twice as large among boys as among girls.

toward zero when there is crossover and noncompliance. One possibility, therefore, is that the groups that show the lowest ITT effects are those who actually experienced the smallest dosage of small classes. To check on this, we generate subgroup estimates of the effect of *assignment* to a small class on *years spent* in a small class. Specifically, we instrument for years spent in a small class using potential years in a small class, where potential years is the product of assignment to a small class with the number of years the student is potentially enrolled in a small class (e.g., four years for those who enter STAR schools in kindergarten, and one year for those who enter in third grade).[12]

We estimate the following equations:

$$YEARS_{is} = \delta_0 + \delta_1 Z_{is} + \delta_{sg} + \psi_{isg} \tag{2}$$

$$COLL_{isg} = \alpha_0 + \alpha_1 YEARS_{is} + \alpha_{sg} + \epsilon_{isg} \tag{3}$$

where $COLL_{isg}$ is a dummy for whether student i, who entered the STAR experiment in school s and in grade g ever enrolls in college. $YEARS$ is the number of years the student spends in a small class. Z is the potential number of years a student can be in a small class multiplied by an indicator for whether the student was assigned to a small class. School-by-entry-grade fixed effects are included in each equation.

We run these equations separately by subgroup. Table 5 reports the estimates of the first stage, reduced form (ITT) and second stage. The first column measures compliance, reporting the number of years actually spent in a small class for each year assigned to a small class. The compliance rate is consistently *smaller* for the most disadvantaged groups, for whom we have seen the largest effects of ITT. This is driven by the higher mobility rate of black and poor students. The 2SLS estimates (Columns 3 and 4) indicate that each year spent in a small class increases college attendance rates by one percentage point for the entire sample, but by 2.4 points for black students and 1.6 points for poor students. These results indicate that students who are black, poor and male benefit more from a year spent in a

[12] Abdulkadiroglu et al. (2011) and Hoxby and Murarka (2009) use a similar approach in instrumenting for years spent in a charter school with potential years spent in charter school, where potential years is a function of winning a charter lottery and the grade of application.

small class than do their peers.

5.3 Heterogeneity by Propensity to Attend College

The analysis of heterogenous effects in the previous sections has examined differences across identifiable subgroups in the effect of small classes on postsecondary outcomes. Examining each of these groups separately leads to low power, and also increases the probability that we will find some subgroup differences if we search across enough dimensions. An alternative approach is to collapse all of these observable differences into a single index that captures the probability that a student will attend college. We use the control group to estimate an equation that relates demographic characteristics to college attendance by age 30. We include in the equation all of the main effects and interactions of race, sex and free-lunch status. We also include school-by-entry-grade fixed effects.

From this regression we obtain coefficients we use to predict, for both the treatment and control group, the probability of attending college. We divide the sample into quintiles based on these propensities and run our estimating equation. While we have produced estimates for each quintile, we show only those for the first quintile and the (pooled) second through fifth quintiles, both for the sake of brevity and due to noisiness of the separate point estimates.

Students with the lowest propensity to attend college show the largest effects of class size on postsecondary outcomes (Table 6). The probability of college attendance in the bottom quintile is 15.2%, while the effect of ITT is an increase of 11.4 points. The effect in the other quintiles is near zero. In Column (5), we show that the difference in the effect size across these two groups is highly statistically significant (p-value=0.000). Looking down the rows of Table 6 we see a similar pattern for other postsecondary outcomes.

6 Do Short-Term Effects Predict Long-Term Effects?

We have shown that random assignment to small classes increases college entry and degree completion, and shifts students toward high-paying majors. Could these effects have been predicted, based on the short-term effects estimated in STAR? That is, are the effects measured at the time of the experiment predictive of the program's long-term effects? A back-of-the-envelope prediction would combine the experiment's effect on scores with information from some other data source on the relationship between scores and postsecondary attainment. We now make such an informed guess about the long-term effects of STAR, then check how well our guess compares with the actual effects.

Our guess requires information about the relationship between standardized scores in childhood and adult educational attainment, ideally for a cohort born around the same time as the STAR subjects.. The NLSY79 Mother-Child Supplement contains longitudinal data on the children of the women of the National Longitudinal Survey of Youth (1979 cohort). These children were born at roughly the same time as the STAR cohort. The children of the NLSY (CNLSY) were tested every other year, including between the ages of six and nine, the ages of the STAR subjects while the experiment was underway. Postsecondary attainment is also recorded in CNLSY. We estimate that, in CNLSY, a standard deviation increase in childhood test scores is associated with a 16 percentage point increase in the probability of attending college.[13]

We combine this information from CNLSY with the estimated effect of small classes on contemporaneous scores. Small classes in STAR increased the average of K-3 scores by 0.17 standard deviations. Under the assumption that the relationship between scores and attainment is the same for the STAR and NLSY79 children, a reasonable prediction of the effect of STAR on the probability of college attendance is 2.72 percentage points (=0.17*16). This back-of-the-envelope calculation is identical to the 2.7 point estimate we obtained in our regression analysis, indicating that the contemporaneous effect of STAR on scores is an excellent predictor of its effect on adult educational attainment.

Another way to approach this question is to examine whether the estimated effect of small classes on postsecondary attainment disappears when we control for K-3 test scores.

[13] We measure college attendance by 2006, when the children were 25 to 29 years old. We regress an indicator for college attendance against the scores from standardized tests administered when the subjects were between six and nine. We use the average of these scores, since respondents take multiple tests. Scores are normalized (within age) to mean zero and standard deviation one.

This is an informal test of whether class size affects postsecondary attainment through any channel other than test scores. This sort of informal test is often used when checking whether an instrument (e.g., assigned class size) affects the outcome of interest (e.g., postsecondary attainment) through any channel other then the endogenous regressor (e.g., test scores). We first estimate the following equation, which relates test scores and postsecondary outcomes:

$$Coll_{isg} = \alpha_0 + \alpha_2 TEST_{is} + \alpha_4 X_{is} + \alpha_{sg} + \epsilon_{isg} \quad (4)$$

Here, $Coll_{isg}$ is a dummy that equals one if student i who entered the STAR experiment in school s and grade g ever attended college. $TEST_{is}$ is the average of student i's kindergarten through third grade math and English test scores, normalized to mean zero and standard deviation of one. Results are in Table 7 (Column 2). In STAR, a one-standard deviation increase in K-3 scores is associated with a 17 percentage point increase in the probability of attending college.[14] This is very similar to the relationship estimated among the children of the NLSY.

We then add to this regression a dummy for assignment to a small class, as well as the interaction of this dummy with test scores. The latter variable allows the relationship between class scores and postsecondary attainment to differ between small and regular classes:

$$Coll_{isg} = \beta_0 + \beta_1 SMALL_{is} + \beta_2 TEST_{is} + \beta_3 SMALL * TEST_{is} + \beta_4 X_{is} + \beta_{sg} + \epsilon_{isg} \quad (5)$$

Results are in Column (3) of Table 7. The coefficient on scores does not change and the newly-introduced variables have coefficients of zero. The zero coefficient on the interaction term indicates that scores are equally predictive of postsecondary attainment for those in small and regular classes. The zero coefficient on the small class dummy indicates that there is no predictive power of assigned class size once we control for contemporaneous test scores (which are boosted by smaller classes). The pattern is similar if we replace college attendance with degree receipt (Columns 3-6).

[14] Results are unchanged if we exclude the school-by-wave fixed effects and demographics.

These findings indicate that short-term gains in cognitive test scores are indeed predictive of long-term benefits. What about medium-term gains - can they predict long-term effects? We estimate the equations just described, replacing contemporaneous scores with those obtained from tests administered three to five years after the experiment had ended (in grades six through eight).These scores are a strong predictor of postsecondary attainment: a standard deviation increase in (the average of) scores in grades six through eight is associated with a 23 percentage point increase in the college attendance rate (Column 2). In Column 3 we add to this regression the class-size dummy and its interaction with scores. The small-class dummy has a statistically-significant coefficient of 0.02, while the interaction has a coefficient of -0.014. The negative coefficient on the interaction indicates that smaller classes moderate the relationship between scores in grade 6-8 and college attendance, reducing the "penalty" to having a low score. Smaller classes also have a direct effect (two percentage points) that does not operate through scores. We conclude that scores recorded several years after the experiment do a significantly poorer job than contemporaneous scores in predicting the effect of the experiment on adult outcomes.

7 Do Early Interventions Pay Off More Than Late Ones?

A theory popularized by economist James Heckman and coauthors is that early interventions pay off more than late ones. Heckman theorizes that students are more plastic when young, and so as they age interventions are less effective in building their human capital. This theory is summarized by Figure IX, taken from Carneiro and Heckman (2003). In this figure, payoffs to interventions are portrayed as decreasing sharply with age of the subject, becoming cost-ineffective soon after preschool. In the past decade, we have accumulated a substantial body of evidence on the causal effect on postsecondary attainment of interventions administered during preschool, elementary, high school, and college. The present paper adds another piece of evidence to this growing collection.

In this section, we assess whether this body of evidence supports the theory depicted in Figure IX. We focus on the results of randomized trials when possible, turning to plausibly-identified quasi-experiments where no controlled experiment has been conducted. Deming and Dynarski (2010) provide a review of this literature, from which much of this information is drawn. We focus on evaluations of discrete, replicable interventions. We deliberately ignore some excellent papers that demonstrate that schools or teachers "matter" for postsecondary attainment if they do not identify the effect of a parameter of the education production function that can be manipulated by policymakers in the short-term (e.g., Deming et al. (2011), Chetty et al. (forthcoming)). We also exclude studies that lack cost estimates. We do not conduct a complete cost-benefit analysis of these programs. Our purpose is to estimate which programs are most cost-effective if the goal is to increase college attendance.

7.1 Preschool Interventions

Two small experiments have tested the effect of intensive preschool on long-term outcomes. Abecedarian produced a 22 percentage point increase in the share of children who eventually attended college. The Perry Preschool Program had no statistically significant effect on postsecondary outcomes. (Anderson, 2008). The subjects in these experiments were almost exclusively poor and black. The cost per student of these two programs were $90,000 and $15,700 respectively.[15] Head Start, a less intensive preschool program, increases college attendance by 6 percentage points (Deming, 2009), with larger effects for blacks and females (14 and 9 percentage points, respectively). While these effects are half of those of the preschool experiments, so too is the cost, at $8,000 per student. Head Start is also operating at scale, unlike the preschool experiments, so it is demonstrably replicable.

We can collapse all of these results into comparable costs by dividing the per-student cost of a program by the proportion of treated students induced into college by that program. For example, Head Start costs $8,000 per child and induces into college 6 of every 100 children treated (6 percent). The amount spent by Head Start to induce a single child into college is therefore $133,333 (=$8,000/0.06). For Abcederian, the figure is $410,000 (=$90,000/0.22).

[15] All costs in this section are in 2007 dollars and come from Deming and Dynarski (2010) unless otherwise indicated. The costs for the early childhood programs and STAR have been discounted back to age zero using a 3 percent discount rate. Costs of the high school and college interventions have not been discounted.

7.2 Elementary and Secondary School Interventions

The present paper shows that smaller classes in primary school increase college attendance by three percentage points, with the effect larger among blacks (five percentage points) and poor children (four percentage points). Among children with the lowest propensity to attend college, the effect is 11 percentage points. The cost of reduced class size is $12,000 per student, larger than that of Head Start but considerably maller than that of the preschool experiments. The amount spent in STAR to induce a single child into college is $400,000 (=$12,000/0.03). If the program could be focused on students with the lowest *ex ante* propensity to attend college (poor, nonwhite) then the cost drops considerably, to $109,000 per student induced into college.

Upward Bound provides at-risk high school students with increased instruction, tutoring and counseling. The program had no detectable effect on the full sample of treated students, but it did increase college attendance among students with low educational aspirations by 6 percentage points (Seftor et al., 2009). Upward Bound costs $5,620 per student. If the program can be targeted on students with low educational aspirations, the implied cost of inducing a single student into college is $87,500 (=$5,620/0.06).

7.3 Postsecondary Interventions

There are no experimental estimates of the effect of financial aid on college entry. Dynarski (2003) examines the effect of the elimination of the Social Security Student Benefit Program, which paid college scholarships to the dependents of deceased, disabled and retired Social Security beneficiaries. Eligible students were disproportionately black and low-income. The estimates from that paper indicate that about two-thirds of the treated students who attended college were inframarginal, while the other third was induced into the college by the $7,000 scholarship. These estimates imply that three students are paid a scholarship in order to induce one into college. The cost per student induced into college is therefore $21,000. [16]

[16] There are several other quasi-experimental estimates of the effect of aid programs on outcomes. It is difficult to assess the amount spent per induced student from the published data in many of these papers. We will continue to expand the list of programs we analyze in this section.

Another way of increasing college enrollment is by assisting students with the administrative requirements of enrolling in college. Bettinger et al. (2009) randomly assign families to a low-cost treatment that consists of helping them to complete the FAFSA, the lengthy and complicated form required to obtain financial aid for college. The cost per treated subject was $84. For every 100 subjects treated 7 were induced into college. The implied cost per student induced into college is $1,200.

7.4 Discussion

These results provide little support for Heckman's assertion that early investments are the most cost-effective, at least if the desired effect is college attendance. The programs producing the biggest effect per dollar spent are those aimed at teenagers and those in their twenties: the Social Security Student Aid Program ($21,000 per student induced into college) and the FAFSA application assistance program ($1,200 per student induced into college). Upward Bound, also aimed at teenagers, could be relatively cost-effective ($87,500 per student induced into college) if limited to students with low educational aspirations. However, since this program is open to all income-eligible students at participating high schools, this level of targeting is unlikely.

Small classes in primary school could also be relative cost-effective, if targeted on students with the lowest *ex ante* probability of going to college ($109,000). This level of targeting may be impossible in practice, since these students are likely scattered within and across schools. If class size reduction were limited to schools attended by poor students, the implied cost per student induced into college would be $300,000. This is cheaper than Abcederian ($410,000 but not as cheap as Head Start $133,333.

A fair conclusion from this analysis analysis is that there are cost-effective programs at every point in the educational pipeline, as well as programs that are ineffective or effective but relatively costly. A question unanswered by this analysis is whether the effects of these various programs would be additive, if implemented across the lifecycle of a targeted population (e.g., poor children). This will be the case only if each programs targets a different set of marginal students.

8 Conclusion

We measure the impact of class size reduction during early elementary school on postsecondary attainment. Assignment to a small class increases college attendance by 2.7 percentage points (7 percent of the control mean). Degree completion is increased by 1.6 percentage points (about 10 percent of the control mean). Gains in degree receipt are driven by increases in high-earning fields such as business, economics, and STEM fields. Effects are largest among black students and students from low-income families, indicating that class-size reductions during early childhood can help to close income and racial gaps in postsecondary attainment.

Our results shed light on the relationship between the short- and long-term effects of an educational intervention. We find that the short-term effect of a small class on test scores is an excellent predictor of its effect on adult educational attainment. In fact, the effect of small classes on college attendance is completely captured by their positive effect on contemporaneous test scores. We further find that the relationship between scores and postsecondary attainment is the same in small and regular classes; that is, the scores of children in the small classes are no less (or more) predictive of adult educational attainment than those of children in the regular classes. This is an important and policy-relevant finding, given the necessity to evaluate educational interventions based on contemporaneous outcomes.

A further contribution of this paper is to identify the effect of manipulating a single educational input on adult educational attainment. The early-childhood interventions for which researchers have identified lifetime effects (e.g., Head Start, Abecederian) are intensive and multi-pronged, including home visits, parental coaching and vaccinations in addition to time in a preschool classroom. We cannot distinguish *which dimensions* of these treatments generate short-term effects on test scores, and whether they differ from the dimensions that generate long-term effects on adult well-being. By contrast, the effects we measure in this paper, both short-term and long-term, can be attributed to a well-defined and replicable intervention: reduced class size.

References

Abdulkadiroglu, Atila, Joshua Angrist, Susan Dynarski, Thomas Kane, and Parag Pathak, "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots," *Quarterly Journal of Economics*, 2011, *126* (2), 699–748.

Achilles, Charles M., *Let's Put Kids First, Finally: Getting Class Size Right*, Thousand Oaks, CA: Corwin Press, 1999.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Angrist, Joshua D. and Alan B. Krueger, "Does Compulsory School Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 1991, *106* (4), 979–1014.

Angrist, Joshua, Susan Dynarski, Thomas Kane, Parag Pathak, and Christopher R. Walters, "Who Benefits from KIPP," NBER Working Paper 15740, National Bureau of Economic Research, Cambridge, MA February 2010.

Arcidiacono, Peter, "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 2004, *121*, 343–375.

Bailey, Martha J. and Susan M. Dynarski, "Gains and Gaps: A Historical Perspective on Inequality in College Entry and Completion," in Greg Duncan and Richard Murnane, eds., *Social Inequality and Educational Disadvantage*, Russel Sage, 2011.

Barrow, Lisa, Thomas Brock, Lashawn Richburg-Hayes, and Cecilia Elena Rouse, "Paying for Performance: The Education Impacts of a Community College Scholarship Program for Low-income Adults," Working Paper 2009-13, Federal Reserve of Chicago, Chicago 2009.

Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu, "The Role of Simplification and Information in College Decisions: Results From the H&R Block FAFSA Experiment," NBER Working Paper 15361, National Bureau of Economic Research, Cambridge, MA September 2009.

Bound, John, Michael Lovenheim, and Sarah E. Turner, "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources," NBER Working Paper 15566, National Bureau of Economic Research, Cambridge, MA December 2009.

Carneiro, Pedro and James Heckman, "Human Capital Policy," in James Heckman and Alan Krueger, eds., *Inequality in America: What Role for Human Capital Policies?*, MIT Press, 2003.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence From Project Star," *Quarterly Journal of Economics*, forthcoming.

Dee, Thomas S., "Are There Civic Returns to Education?," *Journal of Public Economics*, 2004, *88*, 1697–1720.

Deming, David, "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 2009, *1* (3), 111–134.

_ , Justine Hastings, Thomas Kane, and Douglas Staiger, "School Choice, School Quality and Postsecondary Attainment," NBER Working Paper 17438, National Bureau of Economic Research, Cambridge, MA September 2011.

Dobbie, Will and Roland G. Fryer, "Are High Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, July 2011, *3* (3), 158–187.

Dynarski, Susan M., "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion," *The American Economic Review*, March 2003, *93* (1), pp. 279–288.

Finn, J. D. and C. M. Achilles, "Answers and Questions about Class Size: A Statewide Experiment," *American Educational Research Journal*, 1990, *27*, 557–577.

Folger, J. and C. Breda, "Evidence from Project STAR about Class Size and Student Achievement," *Peabody Journal of Education*, 1989, *67*, 17–33.

Garces, Eliana, Duncan Thomas, and Janet Currie, "Longer-Term Effects of Head Start," *American Economic Review*, 2002, *92* (4), 999–1012.

Hamermesh, Daniel S. and Stephen G. Donald, "The Effect of College Curriculum on Earnings: An Affinity Identifier for Non-Ignorable Non-Response Bias," *Journal of Econometrics*, 2008, *144*, 479–491.

Hoxby, Caroline M. and Sonali Murarka, "Charter Schools in New York City: Who Enrolls and How They Affect Student Achievement," NBER Working Paper 14852, National Bureau of Economic Research, Cambridge, MA April 2009.

Krueger, Alan B., "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics*, 1999, *114*, 497–532.

_ **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," *Economic Journal*, 2001, *111*, 1–28.

Lleras-Muney, Adriana, "The Relationship Between Education and Adult Mortality in the United States," *Review of Economic Studies*, 2005, *72*, 189–221.

Ludwig, Jens and Douglas L. Miller, "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *The Quarterly Journal of Economics*, 2007, *122* (1), 159–208.

Roderick, Melissa, Jenny Nagaoka, and Elaine Allensworth, *From High School to the Future: A first look at Chicago Public School graduates' college enrollment, college* *preparation, and graduation from 4-year colleges*, 1313 E. 60th St., Chicago, IL: Consortium on Chicago School Research at the University of Chicago, 2006.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek, *Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]*, Minneapolis: University of Minnesota, 2010.

Schanzenbach, Diane Whitmore, *What Have Researchers Learned from Project STAR?*, Brookings Paper on Education Policy, 2007.

Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, William S. Barnett, Clive R. Belfield, and Milagros Nores, *Lifetime effects: The High/Scope Perry Preschool study through age 40*, Ypsilanti: High/Scope Press, 2005.

Seftor, Neil S., Arif Mamun, and Allen Schirm, "The Impacts of Regular Upward Bound on Postsecondary Outcomes 7-9 Years After Scheduled High School Graduation: Final Report," Technical Report, Mathematica Policy Research, Princeton, NJ 2009.

Statistics, National Center For Education, *Integrated Postsecondary Education Data System*, U.S. Department of Education, 2010.

Word, E., J. Johnston, H. Bain, and et al., *The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project: Technical Report 1985-990*, Tennessee State Department of Education, 1990.

Table 1. Means of Demographics and Outcome Variables by Class Size

	Regular Class	Small Class	Regression Adjusted Difference
Demographics			
White	0.620	0.660	-0.003 (0.005)
Female	0.471	0.473	-0.000 (0.011)
Free Lunch	0.557	0.521	-0.015 (0.011)
College attendance			
Ever attend	0.385	0.420	0.027 (0.011)
Ever attend full-time	0.278	0.300	0.013 (0.011)
Ever attend, but never full-time	0.108	0.120	0.014 (0.006)
Enrolled On-Time	0.274	0.308	0.024 (0.011)
Number of Semesters			
Attempted	3.07	3.39	0.219 (0.133)
Attempted, conditional on attending	7.98	8.08	0.132 (0.209)
Degree Receipt			
Any degree	0.151	0.174	0.016 (0.009)
Associates	0.027	0.034	0.007 (0.004)
Bachelors or higher	0.124	0.141	0.009 (0.008)
Degree Type			
STEM, business or economics field	0.044	0.060	0.013 (0.006)
All other fields	0.085	0.094	0.003 (0.006)
First Attended			
2-year	0.215	0.245	0.025 (0.009)
Public 4-year	0.127	0.132	0.005 (0.007)
Private 4-year	0.042	0.043	-0.003 (0.004)
Ever Attended			
Public 4-year	0.197	0.205	0.003 (0.010)
Private 4-year	0.088	0.100	0.008 (0.007)
Out of state	0.138	0.152	0.013 (0.009)
Sample Size	8,316	2,953	

Notes: Column (3) controls for school-by-wave fixed effects and demographics. Standard errors, in parentheses, are clustered by school.

Table 2. The Effect of Class Size on College Attendance and Persistence

	Total		White	Black	No Free Lunch	Free Lunch	Female	Male
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
College Attendance								
Ever attend	0.028	0.027	0.011	0.058	0.010	0.044	0.016	0.032
	(0.012)	(0.011)	(0.013)	(0.022)	(0.017)	(0.015)	(0.018)	(0.015)
	0.385		*0.432*	*0.308*	*0.563*	*0.272*	*0.455*	*0.324*
Ever attend full-time	0.014	0.013	-0.000	0.037	0.000	0.025	-0.006	0.024
	(0.011)	(0.011)	(0.013)	(0.021)	(0.016)	(0.014)	(0.018)	(0.013)
	0.278		*0.317*	*0.212*	*0.440*	*0.175*	*0.342*	*0.221*
Ever attend, but never	0.014	0.014	0.011	0.021	0.010	0.019	0.022	0.009
full-time	(0.006)	(0.006)	(0.009)	(0.008)	(0.012)	(0.008)	(0.009)	(0.010)
	0.108		*0.115*	*0.095*	*0.123*	*0.098*	*0.113*	*0.103*
Enrolled On-Time	0.025	0.024	0.017	0.036	0.025	0.024	0.024	0.021
	(0.012)	(0.011)	(0.013)	(0.021)	(0.017)	(0.014)	(0.018)	(0.013)
	0.275		*0.321*	*0.197*	*0.449*	*0.163*	*0.334*	*0.221*
Number of Semesters	0.23	0.22	0.15	0.29	0.23	0.22	0.13	0.25
Attempted	(0.14)	(0.13)	(0.15)	(0.28)	(0.22)	(0.17)	(0.22)	(0.14)
	3.07		*3.45*	*2.44*	*4.95*	*1.88*	*3.86*	*2.37*
Demographics	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	11,269	11,269	7,160	4,109	4,454	6,815	5,314	5,955

Notes: Linear probability model regressions used for college attendance dependent variables. The unit of observation is the student. All regressions control for school-by-entry-wave fixed effects. Demographics include race, sex and free lunch status. Standard errors, in parentheses, are clustered by school. Control means are in italics below standard errors.

Table 3. The Effect of Class Size on College Choice

	Total	
Dependent variable	(1)	(2)
College attendance	0.028	0.027
	(0.012)	(0.011)
	0.385	
First Attended:		
2-year	0.025	0.025
	(0.009)	(0.009)
	0.215	
Public 4-year	0.005	0.005
	(0.008)	(0.007)
	0.127	
Private 4-year	-0.002	-0.003
	(0.004)	(0.004)
	0.042	
Ever Attended:		
Public 4-year	0.003	0.003
	(0.010)	(0.010)
	0.197	
Private 4-year	0.009	0.009
	(0.007)	(0.007)
	0.088	
Out of state	0.013	0.013
	(0.009)	(0.009)
	0.138	
Demographics	No	Yes
Sample Size	11,269	11,269

Notes: Linear probability model regressions. The unit of observation is the student. All regressions control for school-by-entry-wave fixed effects. Demographics include race, sex and free lunch status. Standard errors, in parentheses, are clustered by school. Control means are in italics below standard errors.

Table 4. The Effect of Class Size on Degree Receipt

Dependent variable	Total (1)	Total (2)
Any Degree	0.016	0.016
	(0.009)	(0.009)
	0.151	
Highest Degree		
Associates	0.007	0.007
	(0.004)	(0.004)
	0.027	
Bachelors or higher	0.009	0.009
	(0.008)	(0.008)
	0.124	
Degree Type		
STEM, business or economics field	0.013	0.013
	(0.006)	(0.006)
	0.044	
All other fields	0.004	0.003
	(0.006)	(0.006)
	0.085	
Expected Earnings	339.10	329.15
	(147.35)	(146.00)
	27,590.17	
Demographics	No	Yes
Sample Size	11,269	11,269

Notes: Linear probability model regressions. Expected earnings use average earnings in the NLSY97 by educational attainment and college major. The unit of observation is the student. All regressions control for school-by-entry-wave fixed effects. Demographics include race, sex and free lunch status. Standard errors, in parentheses, are clustered by school. Control means are in italics below standard errors.

Table 5. The Effect of Class Size on College Attendance Using Potential Years Instrument

	First Stage	Reduced Form	Two-Stage-Least-Squares	
	(1)	(2)	(3)	(4)
Everyone	0.643	0.006	0.010	0.009
(n=11,269)	(0.016)	(0.003)	(0.005)	(0.005)
			0.385	
Black	0.590	0.017	0.029	0.024
(n=4,109)	(0.019)	(0.006)	(0.010)	(0.010)
			0.308	
White	0.668	0.001	0.002	0.004
(n=7,160)	(0.019)	(0.004)	(0.006)	(0.006)
			0.432	
Free Lunch	0.628	0.010	0.017	0.016
(n=6,815)	(0.015)	(0.005)	(0.007)	(0.007)
			0.272	
Non-Free Lunch	0.665	0.002	0.003	0.003
(n=4,454)	(0.024)	(0.005)	(0.008)	(0.008)
			0.563	
Female	0.651	0.003	0.005	0.004
(n=5,314)	(0.018)	(0.005)	(0.007)	(0.008)
			0.455	
Male	0.637	0.007	0.011	0.012
(n=5,955)	(0.017)	(0.005)	(0.007)	(0.007)
			0.324	
Demographics	No	No	No	Yes

Notes: This tables reports regressions using years spent in a small class. The instrument is potential years in a small class interacted with the small class dummy. Potential years calculated as four minus the entry grade, where K=0. Includes school-by-wave fixed effects. Standard errors clustered by school. Control means are in italics below standard errors.

Table 6. The Effect of Class Size on College Attendance and Persistence, By Probability of Attending College

	Total		Quintile of Ex-Ante Probability of Attending College		P-value: 1st vs. 2nd-5th
			1st	2nd-5th	
Dependent variable	(1)	(2)	(3)	(4)	(5)
College Attendance					
Ever attend	0.028	0.027	0.114	0.006	0.000
	(0.012)	(0.011)	(0.024)	(0.012)	
	0.385		*0.152*	*0.446*	
Ever attend full-time	0.014	0.013	0.059	0.002	0.014
	(0.011)	(0.011)	(0.020)	(0.012)	
	0.278		*0.091*	*0.326*	
Ever attend, but never full-time	0.014	0.014	0.056	0.004	0.008
	(0.006)	(0.006)	(0.016)	(0.008)	
	0.108		*0.062*	*0.119*	
Enrolled On-Time	0.025	0.024	0.052	0.018	0.084
	(0.012)	(0.011)	(0.017)	(0.013)	
	0.275		*0.090*	*0.322*	
Number of Semesters	0.23	0.22	0.54	0.15	0.090
Attempted	(0.14)	(0.13)	(0.17)	(0.16)	
	3.08		*0.96*	*3.63*	
Demographics	No	Yes	Yes	Yes	
Sample Size	11,269	11,269	2,268	9,001	

Notes: Linear probability model regressions used for college attendance dependent variables. The unit of observation is the student. All regressions control for school-by-entry-wave fixed effects. Demographics include race, sex and free lunch status. Standard errors, in parentheses, are clustered by school. Control means are in italics below standard errors.

Table 7. The Effect of Class Size on College Attendance and Degree Receipt, Conditional on Test Scores

	College Enrollment			Degree Receipt		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean K-3 Test Score						
Test score	0.188	0.169	0.169	0.113	0.099	0.096
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Small class * test score	-0.008		-0.008	-0.001		0.000
	(0.010)		(0.010)	(0.008)		(0.008)
Small class	0.000		0.002	-0.001		0.001
	(0.009)		(0.009)	(0.009)		(0.009)
Mean 6-8 Test Score						
Test score	0.247	0.229	0.230	0.156	0.141	0.141
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Small class * test score	-0.016		-0.014	0.007		0.009
	(0.008)		(0.008)	(0.008)		(0.008)
Small class	0.020		0.020	0.010		0.010
	(0.010)		(0.010)	(0.008)		(0.008)
Demographics	No	Yes	Yes	No	Yes	Yes
Control Mean	0.385	0.385	0.385	0.151	0.151	0.151
Sample Size	11,269	11,269	11,269	11,269	11,269	11,269

Notes: Linear probability model regressions. The unit of observation is the student. All regressions control for school-by-entry-wave fixed effects. Demographics include race, sex and free lunch status. Missing test-score indicators included for students with no test scores in grade range. Standard errors, in parentheses, are clustered by school.

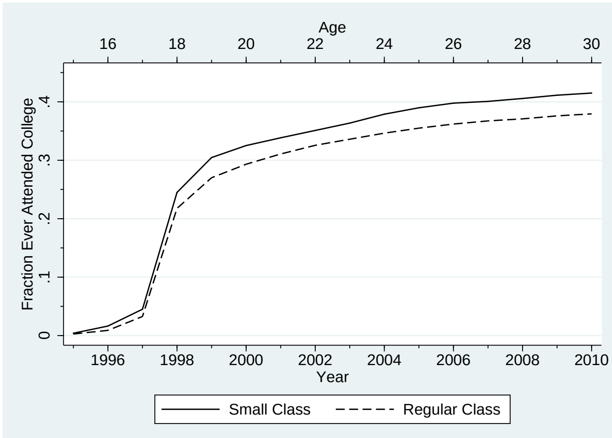
Table 8. The Effect of Class Size on College Attendance, by Years Enrolled

	Baseline - All Years of Enrollment		Exclude Pre-1999 Enrollment		Exclude Post-2007 Enrollment		Include 1999-2007 Enrollment Only	
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ever attend	0.028	0.027	0.019	0.018	0.024	0.023	0.017	0.016
	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
	0.385		*0.369*		*0.372*		*0.357*	
Demographics	No	Yes	No	Yes	No	Yes	No	Yes
Sample Size	11,269	11,269	11,269	11,269	11,269	11,269	11,269	11,269

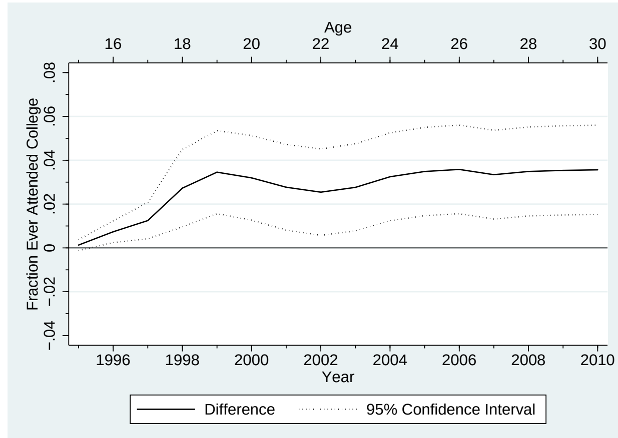
Notes: Linear probability model regressions. The unit of observation is the student. All regressions control for school-by-entry-wave fixed effects. Demographics include race, sex and free lunch status. Standard errors, in parentheses, are clustered by school. Control means are in italics below standard errors.

Figure I: Fraction Ever Attended College Over Time, By Class Size

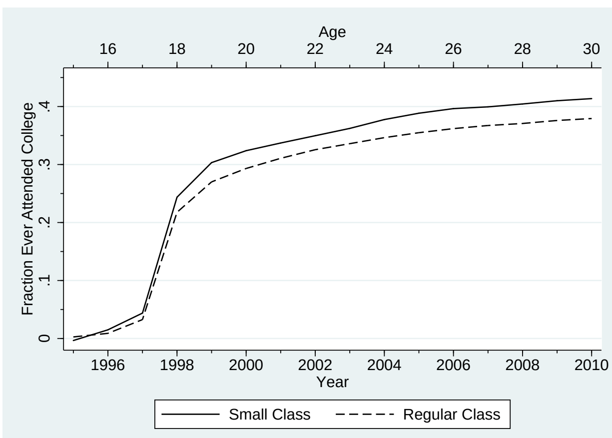
(a) Raw Means



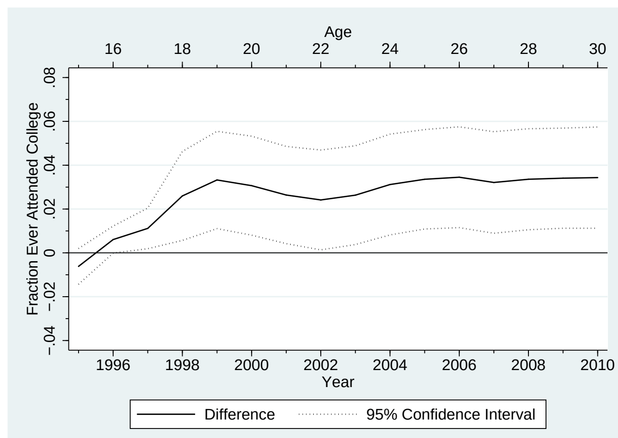
(b) Difference Between Small and Regular



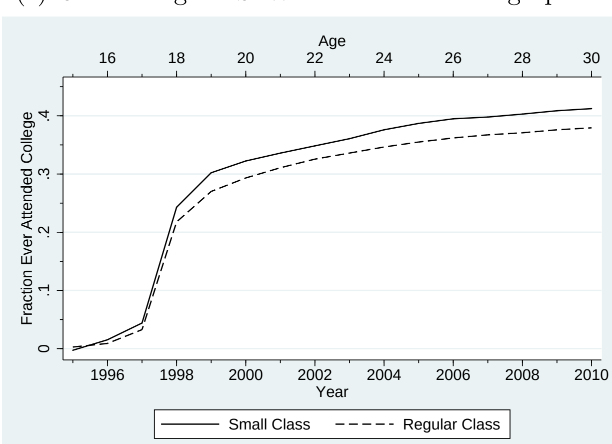
(c) Controlling for School-By-Wave Effects



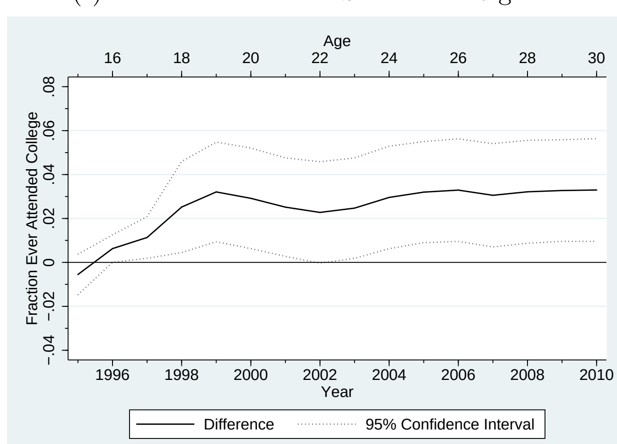
(d) Difference Between Small and Regular



(e) Controlling for SxW Effects and Demographics



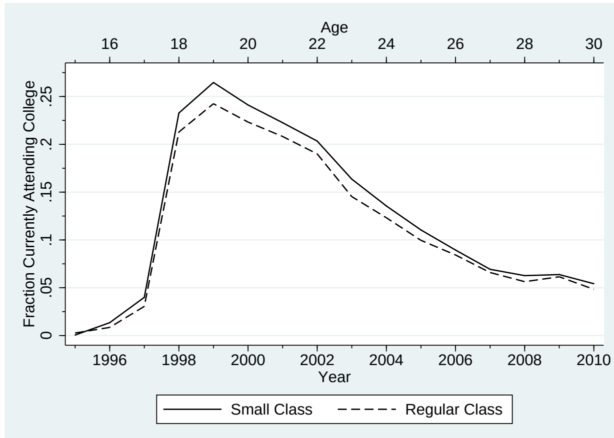
(f) Difference Between Small and Regular



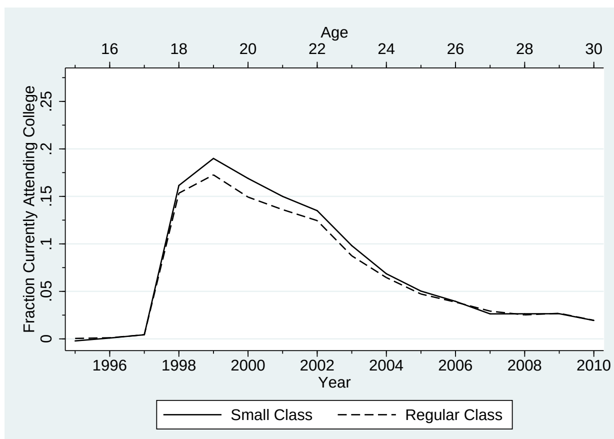
Notes: Figures (a), (c) and (e) plot the mean fraction ever attended college by year for students who were in small vs. regular size classes. Figure (a) plots raw means, (b) controls for school-by-wave fixed effects, and (c) controls for both school-by-wave fixed effects and demographics, including race, sex and free lunch status. Figures (b), (d) and (f) plot the difference and its 95% confidence interval by year for figures (a), (c) and (e), respectively. Standard errors are clustered by school.

Figure II: Fraction Currently Enrolled in College Over Time, By Class Size and Enrollment Status

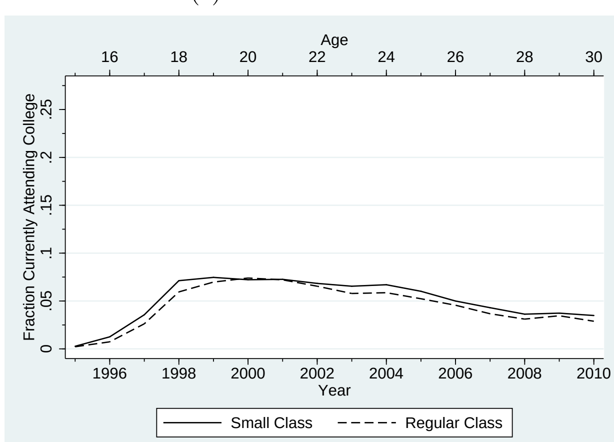
(a) Any Enrollment Status



(b) Full-Time Status



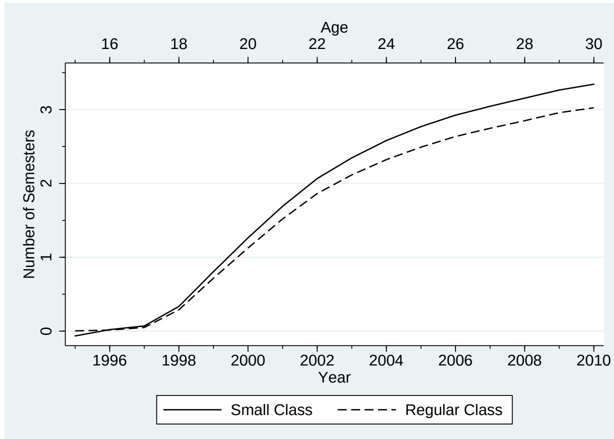
(c) Part-Time Status



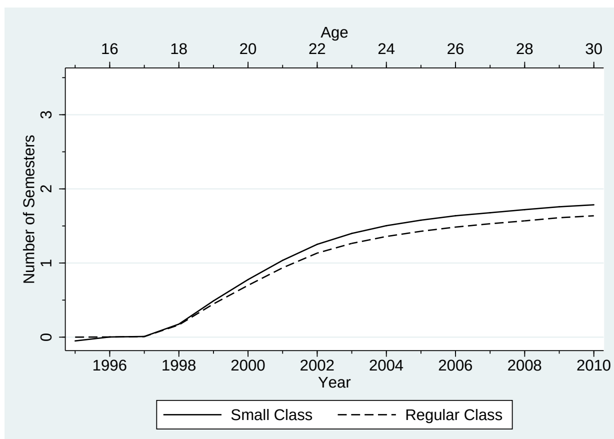
Notes: Figures plot the fraction currently attending college by year for STAR students who were in small vs. regular size classes. All figures control for both school-by-wave fixed effects and demographics, including race, sex and free lunch status.

Figure III: Cumulative Number of Semesters Attended Over Time, By Class Size and Enrollment Status

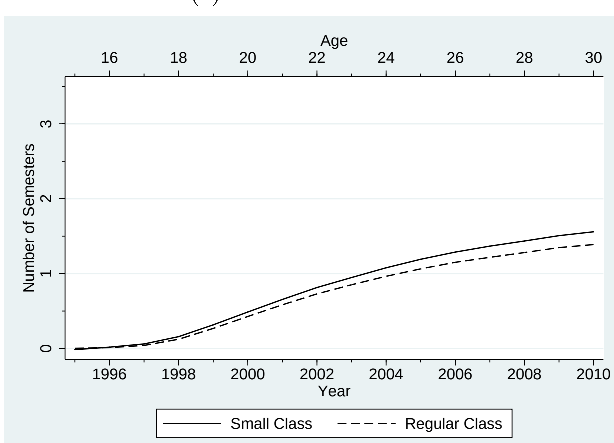
(a) Any Enrollment Status



(b) Full-Time Status



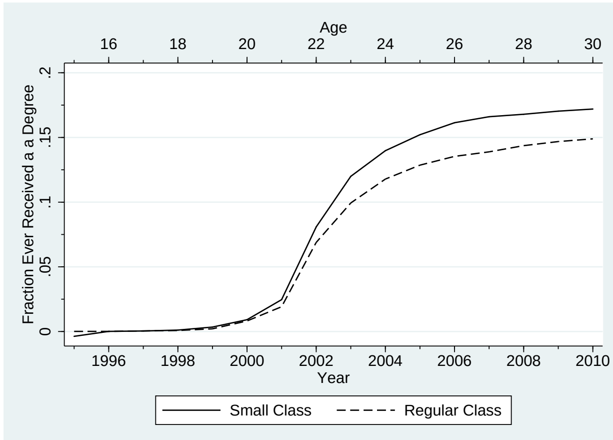
(c) Part-Time Status



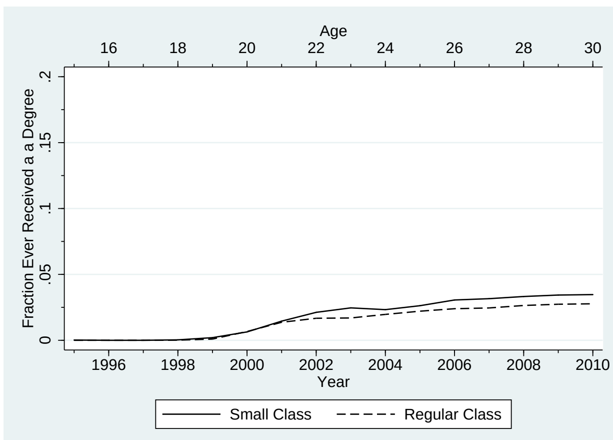
Notes: Figures plot the average cumulative number of semesters attended by year for STAR students who were in small vs. regular size classes. All figures control for both school-by-wave fixed effects and demographics, including race, sex and free lunch status.

Figure IV: Fraction Ever Received A Postsecondary Degree, By Class Size and Degree Type

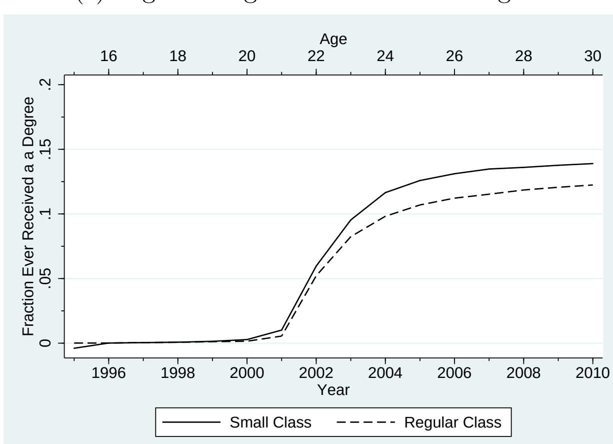
(a) Any Degree (Associates and Higher)



(b) Highest Degree - Associates



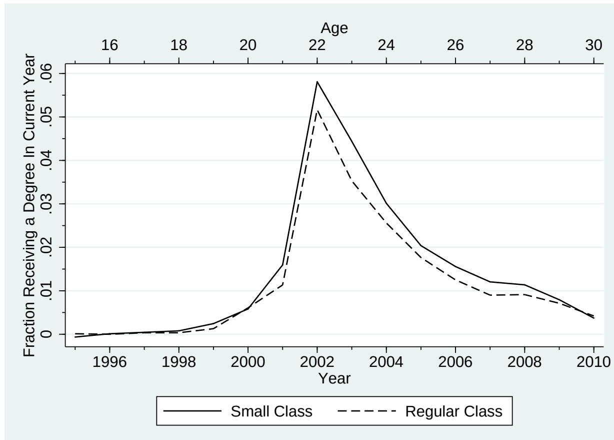
(c) Highest Degree - Bachelors or Higher



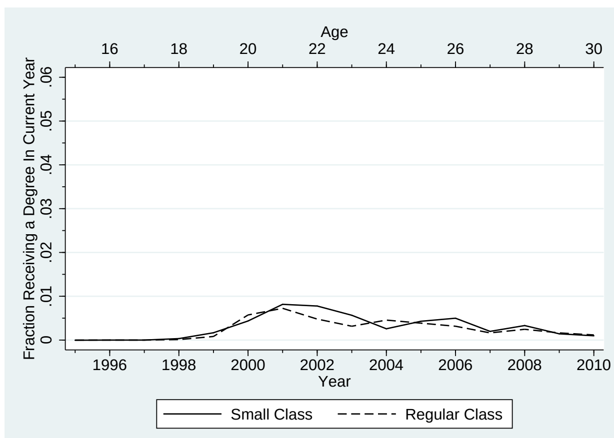
Notes: Figure IVa plots the fraction of students ever receiving any postsecondary degree. Figure IVb plots the fraction of students with an Associates as their highest postecondary degree received. Figure IVc plots the fraction of students with a Bachelors or higher as their highest postsecondary degree received. All figures control for both school-by-wave fixed effects and demographics, including race, sex and free lunch status.

Figure V: Fraction Receiving a Postecondary Degree in Current Year, By Class Size and School Type

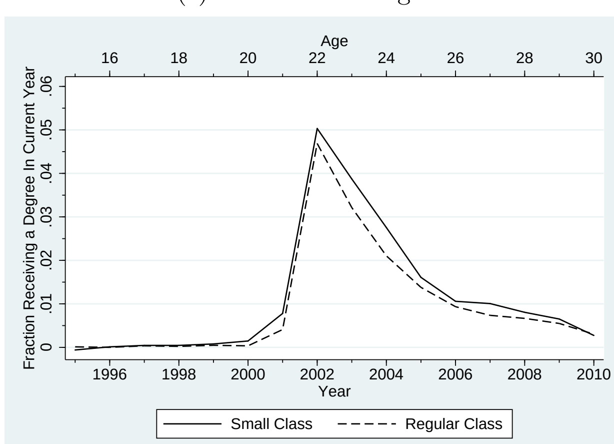
(a) Any Degree



(b) Associates

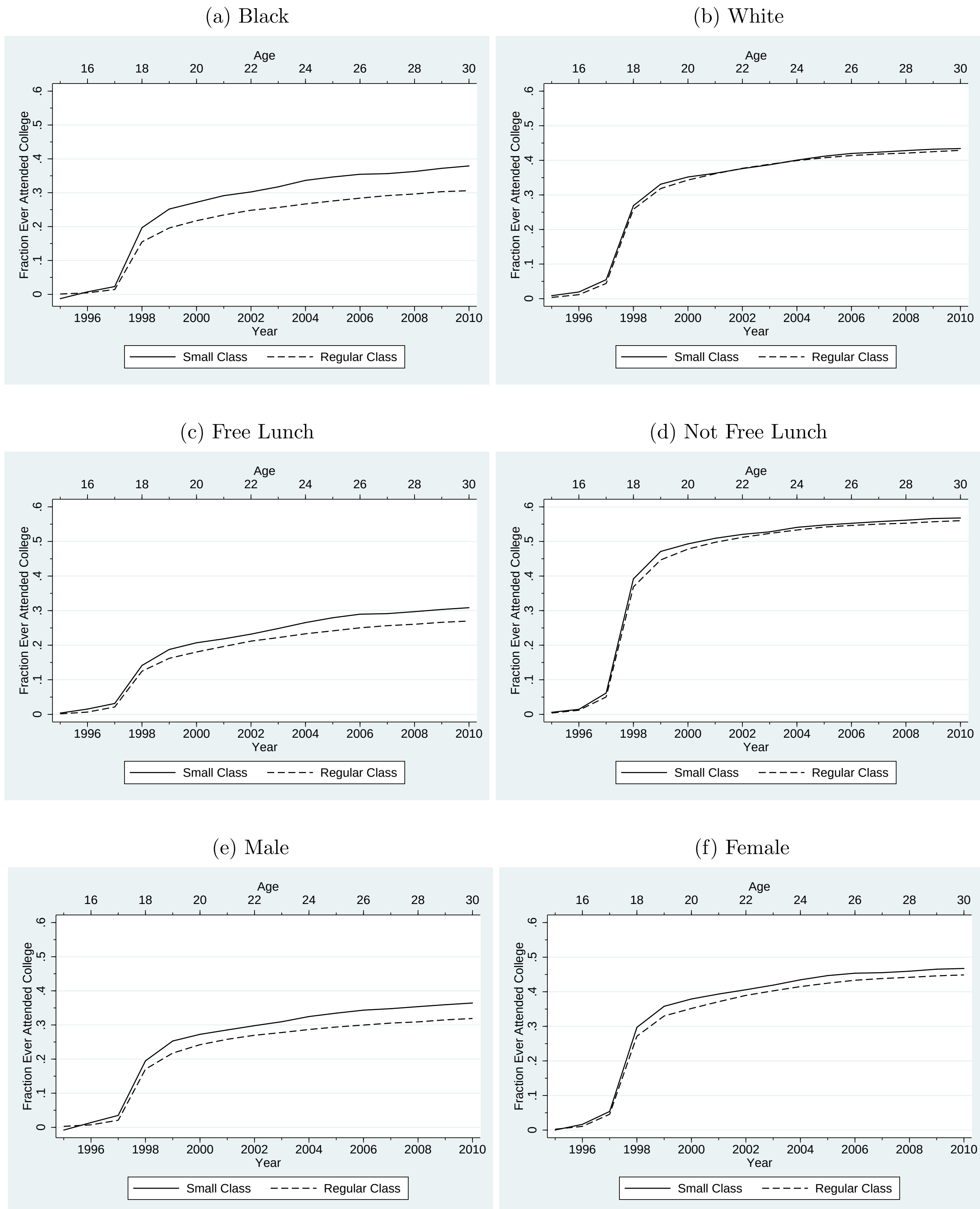


(c) Bachelors or Higher



Notes: Figures plot the fraction receiving a postsecondary degree in the current year for students who were in small vs. regular sized classes. All figures control for both school-by-wave fixed effects and demographics, including race, sex and free lunch status.

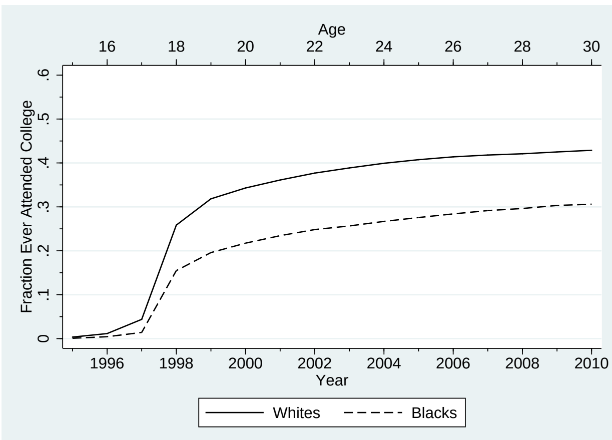
Figure VI: Fraction Ever Attended College Over Time, By Class Size and Subgroup



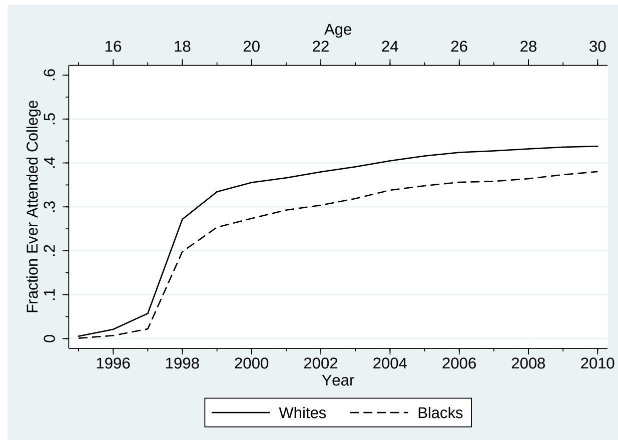
Notes: Figures plot the fraction ever attended college by year for STAR students who were in small vs. regular size classes. All figures control for both school-by-wave fixed effects and demographics, including race, sex and free lunch status.

Figure VII: Impacts on Inequality - The Effects of Class Size on the Race and Income Gap in College Attendance

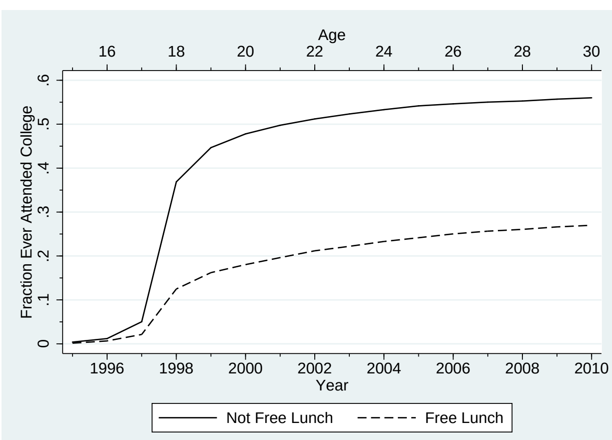
(a) Regular Class



(b) Small Class



(c) Regular Class

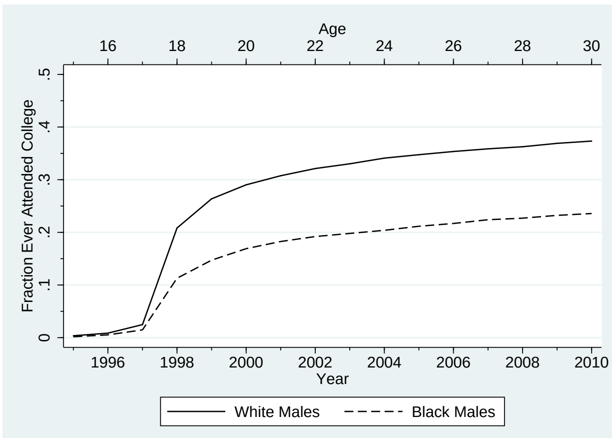


(d) Small Class

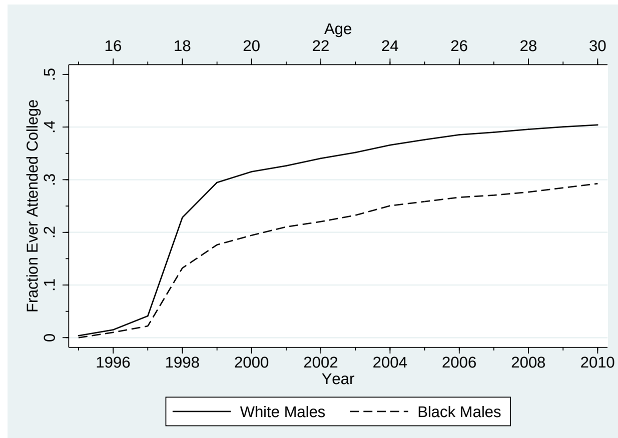
Notes: Figures (a) and (c) plot the fraction ever attended college by year for STAR students who were in regular size classes, and figures (b) and (d) for STAR students who were in small size classes. Figures (a) and (b) compare college attendance by race, and figures (c) and (d) compare college attendance by free lunch status.

Figure VIII: Impacts on Inequality - The Effects of Class Size on The Race Gap, by Sex

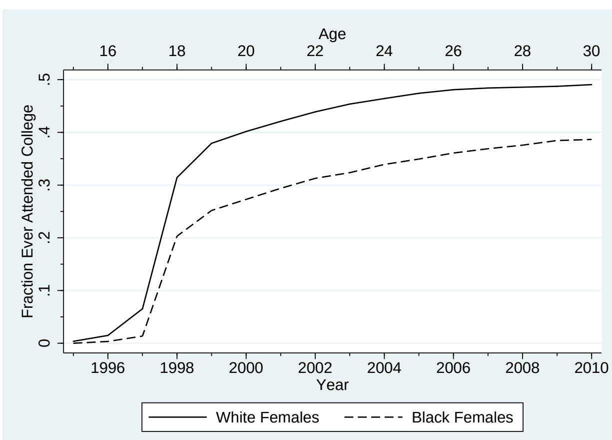
(a) Regular Class



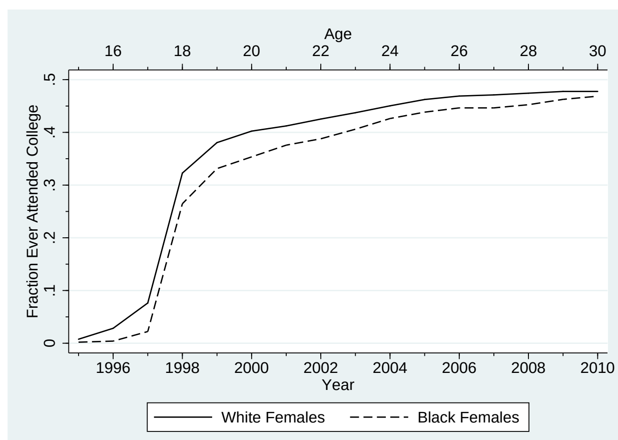
(b) Small Class



(c) Regular Class

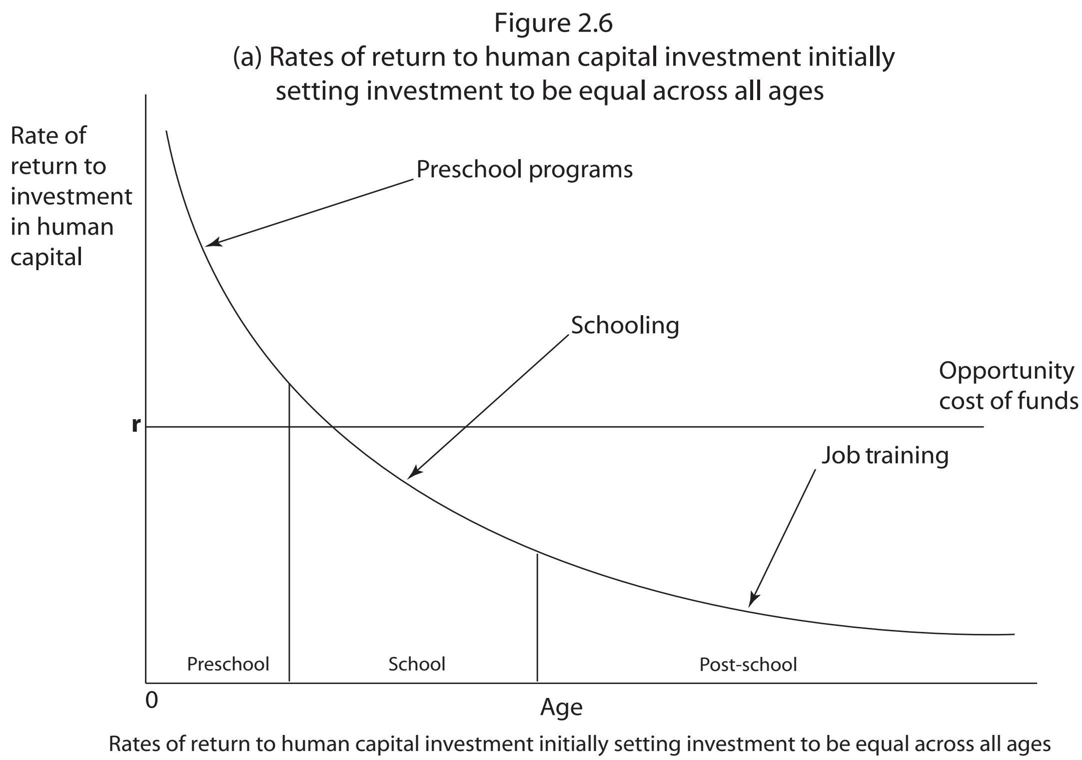


(d) Small Class



Notes: Figures (a) and (c) plot the fraction ever attended college by year for STAR students who were in regular size classes, and figures (b) and (d) for STAR students who were in small size classes. Figures (a) and (b) compare college attendance by race for male students, and figures (c) and (d) compare college attendance by race for female students.

Figure IX: Illustration of Return to Educational Interventions Decreasing with Age



Notes: This picture taken from Carneiro and Heckman (2003).