UNIVERSITY OF CALIFORNIA, SAN DIEGO

# Essays on the Impacts of Education Policies

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Dallas Dwight Dotter

Committee in charge:

Professor Julian R. Betts, Chair
Professor Julie B. Cullen
Professor Gordon B. Dahl
Professor Alan J. Daly
Professor Joshua S. Graff Zivin

2013

The dissertation of Dallas Dwight Dotter is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

DEDICATION

To my wife, Shauna, for her unconditional support during times both good and bad.

EPIGRAPH

*Without a chance of failure,*

*a success is not as sweet.*

—Benjamin C. Horne

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

## ACKNOWLEDGEMENTS

This dissertation was supported by fellowship funding from the Spencer Foundation and National Academy of Education.

I am extremely grateful to Julian Betts for years of patient guidance and advising throughout my graduate studies.

I thank Julie Cullen, Gordon Dahl, Bridget Long, Leah Nelson, and Prashant Bharadwaj for insightful comments and suggestions. I also thank participants of the UCSD applied microeconomics seminars for useful feedback.

Thanks also go to Karen Bachofer and Andrew Zau of SanDERA for help with administrative data and institutional details.

Any errors within this dissertation are the author's alone.

Chapters 1, 2 and 3 are in preparation for publication as separate academic journal articles.

VITA

| | |
|---|---|
| 2007 | B.S. in Economics *magna cum laude*, University of North Texas |
| 2008-2013 | Graduate Teaching Assistant, University of California, San Diego |
| 2009-2011 | Research Assistant, Prof. Julian R. Betts, UCSD |
| 2009-present | Research Associate, San Diego Education Research Alliance |
| 2012 | Instructor, Department of Economics, University of California, San Diego |
| 2013 | Ph.D. in Economics, University of California, San Diego |
| 2013 | Researcher, Mathematica Policy Research |

ABSTRACT OF THE DISSERTATION

**Essays on the Impacts of Education Policies**

by

Dallas Dwight Dotter

Doctor of Philosophy in Economics

University of California, San Diego, 2013

Professor Julian R. Betts, Chair

This dissertation evaluates the impacts of education policies affecting students early in their educational careers. The first chapter studies the impacts of universally free school breakfast programs on student achievement. I study the staggered implementation of an in-classroom breakfast program in San Diego elementary schools, which provides meals to all students during class time, to determine the impacts of universally free school breakfasts on student attendance rates, classroom behavior and academic performance. Introducing universally free breakfasts increases math and reading test score gains by roughly 15 and 10 percent of a standard deviation on average, respectively. The results suggest that offering

universally free breakfast increases participation–perhaps by reducing associated social stigmas–and that the resulting positive impacts on academic achievement are at least partly driven by year round benefits rather than only consumption at the time of testing. Chapter two investigates the long-term effects of school entry age on student achievement and educational attainment. Using finely detailed longitudinal data on students in San Diego, I exploit the discontinuity in student ages resulting from school-entry cutoff dates to investigate the relationships between relative age, classroom environment, differentiated curricula and long-run educational outcomes. I find that achievement premiums from entering school relatively older persist well through the high school years. Older entrants are also more likely to enroll in four-year postsecondary institutions after high school, have higher levels of postsecondary persistence, but not higher levels of attainment. These differences can be partially explained by large and lasting differences in classroom behavior during elementary grades and access to higher level course placements beginning in grade eight. Finally, chapter three studies the impacts of high ability tracking on overall educational attainment for students just on the margin of inclusion in "‘gifted and talented"’ classrooms. Using a regression discontinuity design, I find GATE participation for these students increases rates of postsecondary degree attainment by roughly 8 percentage points. The results suggest that expanding high ability tracking programs to include students of slightly lower levels of ability may benefit overall educational attainment through more advanced curricula and higher quality peers.

# Chapter 1

# Breakfast at the Desk: The Impact of Universal Breakfast Programs on Academic Performance

## 1.1 Introduction

Medical and dietetic studies report that between 12 and 30 percent of school-aged children skip breakfast on any given weekday (NHANES survey data, 1999-2006; Nicklas et al., 1993; Sampson et al., 1995; Siega-Riz et al., 1998). This percentage is decreasing with income so that roughly 67 percent of middle-income African-Americans and low-income whites or Hispanics do not eat breakfast daily (ADA survey data, 2010). Furthermore, there is suggestive evidence that eating a breakfast which contains sufficiently balanced nutrients has a beneficial impact on both student health – in terms of nutrient intake (Bhattacharya et al., 2006), height-to-weight ratio and early physical development (Hofferth and Curtin, 2005; Millimet et al., 2010) – and cognitive skills such as focused attention and memory recall (Wesnes et al., 2003). In light of this information, many schools have attempted to increase participation in School Breakfast Programs (SBPs) by offering universally free breakfasts to all students before school. Despite these efforts, however, there is little research that has rigorously estimated the causal effects of universal school breakfast programs on academic outcomes of students. This can be attributed to at least two factors. Firstly, there are few data sources that provide both detailed information on participation in school breakfast programs and academic outcomes at the student level. Many data sets, such as the Child Development Supplement of the Panel Study of Income Dynamics (CDS-PSID) and the National Health and Nutrition Examination Survey (NHANES) contain excellent data on the breakfast and nutrient intake of children as well as measures of individual health, but do not provide comparable measures of academic performance. Secondly, most studies suffer from a lack of exogenous variation in SBP participation and cannot overcome between and within program selection issues when attempting to estimate causal impacts[1].

[1]The direction of selection bias is ambiguous. For example, there can exist both negative and positive within-school selection into SBP since participation is voluntary. To qualify for free or reduced-price meals, a student's household income must be below a predetermined threshold usually determined at the state level. Due to greater financial need or lesser social stigmas surrounding meal subsidy programs, one might expect eligible students from families at the lower end of the income

This paper exploits the staggered implementation of a new, "Breakfast in the Classroom" program (BIC) across San Diego[2] elementary schools which provides universally free breakfasts to all students during class time. Once BIC is integrated into the daily classroom experience, the proportion of students eating school breakfasts is reported to increase from 25 percent to roughly 92 percent. Additionally, uneaten meals are returned after breakfast to record the number of meals consumed[3]. Thus, this paper overcomes the main limitations in the literature on school breakfasts by 1) using near-complete treatment to avoid within-school selection biases typical of traditional SBPs, 2) using finely detailed longitudinal data with comparable treatments, comparable outcome variables and fixed effects, and 3) exploiting the plausibly exogenous variation in the timing of program implementations across schools, allowing for control of school-specific trends in outcomes over time that may be unrelated to school breakfasts.

There are at least a few ways by which BIC can affect academic performance. First, a nutritionally balanced meal provided early in the day can improve student health when the meal received otherwise is inferior. Some research suggests greater levels of health and nutrient intake can benefit cognitive functions, improving academic performance and learning efficacy (Kleinman et al., 2002; Gleason and Suitor, 2003; Florence et al., 2008). Second, in addition to improved health and nutrition, the provision of free meals in the morning may incentivize reduced tardiness and increased attendance rates among students (Murphy et al., 1998; Kleinman et al., 2002). Third, school-provided meals are likely to reduce student household food expenditure (Long, 1991). To the extent that positive income shocks are asso-

distribution to be more likely to participate in SBPs. To the extent that income levels are positively associated with academic outcomes, this introduces negative selection when examining the effect of SBPs on academic performance. On the other hand, most SBPs require students to arrive at school earlier than classes begin to receive breakfast. To the extent that parental abilities to enroll children in the program and get them to school early enough to participate may be correlated with higher levels of parental involvement in their child's academics, there will be some positive selection into SBP participation conditional on program eligibility.

[2]Chicago Public Schools, Dallas ISD, Little Rock School District, Memphis City Schools, Orange County Public Schools (Florida) and Prince George's County Public Schools (Maryland) are among other school districts using similar BIC programs in public elementary schools.

[3]A meal is considered consumed if any one part of the meal is eaten by the student.

ciated with increases in student performance (Dahl and Lochner, 2010), BIC may improve student outcomes through a household income effect. Fourth, because BIC is administered after class begins, reallocating instruction time for the program means BIC can potentially have a negative impact on students' academic outcomes. Fifth, BIC eliminates the need for students previously participating in the SBP to arrive earlier than the school day begins, essentially delaying the school day starting time for these students. To the extent that later school starting times translate into more sleep or time for morning preparations, they too can affect student outcomes. Finally, BIC integrates students across income strata within a classroom. Under the traditional SBP, breakfast participators are usually eating in the cafeteria while non-participators are free to socialize on the school grounds before class begins. This segregation of SBP participating students and students from households with higher income levels is eliminated when breakfast is moved into the classroom and uniformly distributed.

This paper estimates the net results of the aforementioned composite effects stemming from the BIC program on student outcomes. The net effects of BIC are separately identified for both students in schools where universally free breakfasts were already provided on a voluntary, before-school basis and for students in schools where eligibility for free meals was determined by household income levels before BIC. Additionally, data on the levels of pre- and post-BIC breakfast take-up levels are used to estimate the effects of BIC on student outcomes as a function of the increase in students eating school breakfasts. Identification of these effects are driven by the varied timing of implementation across schools. This timing appears to be quasi-random with respect to the student outcomes of interest. Anticipatory effects of future BIC implementation on contemporaneous variables are tested and the null hypothesis of no effect fails to be rejected.

BIC increases student gains in math and reading by roughly 15 and 10 percent of a standard deviation, respectively, in schools that did not previously offer universally free breakfasts. Within these schools, the impact magnitudes are increasing with the percentage increase of students consuming breakfast after the introduction of BIC and are mostly driven by students with below-average achievement levels. There are no significant impacts on achievement in schools that already provided universally free breakfasts. These results suggest that fewer children who would benefit from school breakfasts participate when free meal provisions are associated with income levels–consistent with the idea that social stigmas result in a less than optimal level of take-up. Moreover, increases in gains are at least as large in additional years of the program, suggesting there are benefits over time from exposure to free breakfasts rather than an initial boost from consumption during testing periods.

The impact magnitudes on test score gains are striking. For reference, 15 percent standard deviation gains in math are at least as large as gains realized from increasing teacher quality by one standard deviation (Rockoff, 2004; Aaronson, Barrow and Sander, 2007; Dobbie, 2011) and gains from decreasing classroom size by ten students (Rivken, Hanushek and Kain, 2005). Given the relatively low average cost of about two hundred dollars per newly participating student, per year, these results suggest universally free breakfasts are a relatively inexpensive way to drive student gains in schools with below average income levels. This holds policy implications for mitigating the persistent achievement gaps across socioeconomic backgrounds.

The paper proceeds as follows. Section 1.2 presents a brief review of some of the previous literature pertaining to public school meal programs. Section 1.3 describes the details of the school breakfast programs available to students of public schools in San Diego. Section 1.4 describes the data used for this study and Section 1.5 discusses the research design and identification strategies employed. Section 1.6 presents the results as well as robustness checks. Finally, Section 1.7 concludes with policy implications and some remarks pertaining to future research on the comparative efficacy of school breakfast programs.

## 1.2 Review of the School Breakfast Program Literature

For the most part, the previous literature investigating the effects of school breakfast programs fall into one of two categories: those concerned with the effects on student health and those looking at the effects on academic or cognitive performance. A majority of the literature on school breakfast programs has focused on the impact on student health measures, such as weight and the intake of various nutrients relative to baseline measures. Gleason and Suitor (2003) conclude that school breakfast and lunch programs increase the intake of nutrients among students, but also increase the intake of dietary fat. Hofferth and Curtin (2005) observe that students who participate do tend to have higher weights and find no effect of SBP after accounting for positive selection into free or reduced price lunch program participation.[4]

The economic literature on SBPs has also focused primarily on the impact on students' weight and nutrition. Using a difference-in-differences strategy, Bhattacharya, Currie and Haider (2006) find that the availability of a SBP has no effect on the total number of calories consumed or the likelihood of eating a breakfast. They do, however, find that access to a SBP substantially improves the nutritional quality of both the student's diet as well as the diets of others in the student's household. Millimet, Tchernis and Husain (2010) use ECLS-K data to estimate the long-run effect of school breakfast and lunch program participation on childhood weight 3 years later. They find that children who gained more weight prior to Kindergarten, and therefore with steeper weight trajectories, are more likely to participate in the program. When accounting for even low levels of positive selection into SBPs, Millimet et al. find a negative effect of program participation on child weight.

The body of literature examining the effect of SBPs on academic performance is considerably smaller and more limited in scope, but necessarily overlaps

[4]This is an important consideration, since eligibility is usually determined by the same income threshold for both voluntary breakfast and lunch assistance programs.

with various literature on the effects of meal quality and timing on outcomes that may influence academic performance. Schoenthaler et al. (2000) performed a randomized, double-blind, placebo-controlled trial in which a sample of elementary school children in the Phoenix area were given daily vitamin-mineral supplements for four months. The treatment group showed a significant 47% decrease in the rate of rule violations, as measured by school disciplinary records. Although there was no further investigation into which vitamins and minerals in particular explain the results, they are suggestive of an effect nutrition levels may have on individual behavior. Wesnes et al. (2003) performed randomized lab experiments with children between the ages of 9 and 16 years to determine the effect of breakfast on cognitive performance. Subjects were randomly given either one of three breakfast meals or no meal and took computerized attention and memory tests before and after meal treatments. Responses to a questionnaire on mood and alertness were also collected. They find that skipping breakfast impairs attention and short-term memory. Moreover, the impairment increases over time during the tests.[5] Florence et al. (2008) look at the effect of nutrition and health on school performance. Using survey data of 5200 5th grade students in Nova Scotia on lifestyle and health measures and dietary intake, they link individuals to results of a standardized literary assessment. Controlling for socioeconomic factors, they find that students with decreased levels of nutrition in their diet performed worse on the standardized assessment. Unfortunately, as with most of the literature in this area, they are unable to deal with the endogeneity of nutrition (e.g. diet quality being correlated with determinants of academic performance, such as overall parental involvement).[6]

Several studies look specifically at student outcomes when the treatments are SBPs. Murphy et al. (1998) test for a relationship between SBP participation and academic functioning in school-aged children. They collect data from three

[5] Disclosing financial interests, it was noted that brand name cereals made up two of the three meal treatments and that Cereal Partners UK sponsored this study.

[6] Additionally, though perhaps not externally valid for U.S. students, Afridi (2007) shows evidence that school meal programs in some Indian villages significantly boost attendance rates among young girls and enrollment rates among young girls from disadvantaged socioeconomic groups.

public schools, one in Philadelphia and two in Baltimore, before and after the implementation of Universally Free Breakfast Programs (UFBP), which make free breakfasts available to all students. Although their results are correlational and do not account for self-selection, they find that breakfast participation nearly doubled after the UFBP implementation. Compared to other students, the subsample who increased breakfast program participation had significant increases in math grades and attendance rates. They argue that the increase of participation under UFBP among students who were already eligible under the old program suggests that the UFBP alleviates the social stigma associated with subsidized meal programs.

Kleinman et al. (2002) investigate how the introduction of UFBPs in schools affects students' nutrient intake, psychosocial functioning and academic performance. Using academic records and 24-hour dietary recall data for 96 inner city school children, they find no statistically significant changes in the mean levels of any outcomes after the introduction of UFBP. However, when restricting the sample to the students which did experience improved nutritional intake post-UFBP, they do find a significant increase in students' math GPA and behavioral scores, as well as a decrease in school-day absences. It should be noted that, as with other SBPs, students still self-select into UFBPs and therefore the significant effects they find for treatment on the treated are possibly confounded by non-random selection. Moreover, pre- and post-treatment outcome data were collected 6 months before and after treatment began. Thus the 24-hour recall data used to determine nutrient intake may not be very representative of individuals' average diet over the study period.

Figlio and Winicki (2005) show that many schools under threat of No Child Left Behind accountability sanctions will increase the caloric content of provided meals in an attempt to boost student test scores. Interestingly, they find that of the schools which do, there is a significant increase in the performance of tested students, though anecdotal reports suggested students at these schools were also given pre-test snacks rich in glucose. More recently, Frisvold (2012) uses a regression discontinuity (RD) design to estimate the effect of SBP availability on student achievement. He exploits variation across states in the minimum percentage of students eligible for free or reduced price meals within a school that dictates a SBP be mandated. Frisvold finds that SBP availability increases achievement in reading, math and science by about 8, 10 and 14 percent, respectively.

During the course of this study, Imberman and Kugler (2012) was released as a working paper. They use data from a large urban school district to estimate the impact of an in-class breakfast program. Although the empirical methods differ slightly from this paper, they find similar results with respect to standardized testing outcomes[7].

## 1.3 Overview of School Breakfast Programs in San Diego

The School Breakfast Program (SBP) was established in 1966 under the federal Child Nutrition Act. Initially a pilot program providing federal grants to schools that serve breakfast to "nutritionally needy" students, the program was made permanent in 1975 and moved towards a per-meal reimbursement system. Today, the SBP provides free or low-cost, nutritionally balanced[8] meals to children in public and nonprofit private schools.

The traditional SBP in San Diego offers free breakfasts to eligible students at school, before the normal school day begins. To qualify, a student's household income must be below a predetermined threshold. Table 1.1 presents the federal poverty guidelines by household size which are used to determine eligibility for free or reduced price school meals. Students whose household income falls at or below 130 percent of the federal poverty guideline are eligible for free meals. Students

[7] Empirical results are discussed in Section 1.6.

[8] School meals provided under the federal school breakfast program must have less than 30% of total caloric content from fat, less than 10% from saturated fats and contain at least 1/4th of the USDA daily recommended dietary allowances for protein, iron, calcium and vitamins A and C. Minimum portion sizes are determined by age and grade groups. There are also minimum number of offerings required from each of the main food groups as defined by the USDA.

with a household income between 130 and 185 percent of the guideline are eligible for reduced-price meals. As shown in Table 1.2, however, all students in San Diego meeting federal free or reduced price eligibility receive free breakfasts.

A portion of elementary schools in San Diego held Provision 2 status prior to the years of implementation for the particular breakfast program studied here. Under Provision 2, a school provides universally free meals to all students, regardless of household income level. Universally Free Breakfasts (UFB) are offered at these schools on a voluntary basis before the school day officially begins. Under Provision 2 status, schools are only required to update free meal eligibility paperwork for students every 4 years, rather than annually. Thus, doing so reduces administration overhead and is thought to increase meal participation by eliminating the free meal application process. It may also increase participation by reducing social stigmas surrounding meal subsidy programs when participation signals eligibility by income.

The new Breakfast in the Classroom (BIC) program is motivated by a desire to increase breakfast participation among students who qualify for free meals. Before BIC, an average of 33 percent of San Diego students ate breakfast at school while over 60 percent qualified for free meals. Roughly 25 percent of students participated under the traditional SBP, while about 65 percent participated under Provision 2 when breakfast was universally free. The BIC program attempts to increase participation by serving free breakfasts to all students in participating schools during classroom time. Meals are centrally provided by the school's food services department. All school meals, BIC or otherwise, are designed to meet USDA guidelines for daily nutritional and caloric intake. Under BIC, students eat breakfast at their desk during approximately the first 15 minutes of class. Although students and parents have the option of declining the meal, classrooms report high meal participation levels from the program. Including meals not utilized due to student absences, roughly 92 percent of total number of possible meals are consumed on average. Figure 1.1 plots the annual average percentage of students enrolled who consume school breakfasts for each school, by the percentage of students that qualify for free meals. Each point represents a school location and are separated by 4 types: 1) BIC schools that previously provided UFB under Provision 2 status, 2) BIC schools that did not previously offer UFB, 3) non-BIC schools that do provide UFB and 4) non-BIC schools that do not offer UFB.

The BIC program has been implemented in a staggered fashion across eligible elementary schools in San Diego. Figure 1.2 presents a map of the San Diego area highlighting the attendance zone boundaries of BIC-eligible schools and the years each BIC school implemented the program. To ensure estimates capture the impacts on academic outcomes, schools are considered to be treated by BIC if the program begins at least 30 days before annual standardized testing begins. One elementary school began BIC as a pilot in Fall of 2006. The first wave consisted of five more schools implementing the program during the 2007-2008 school year, followed by four more waves of 5 to 18 schools per year staggered across the academic years through December 2011.

The only requirement for an elementary school in San Diego Unified School District (SDUSD) to be eligible for the BIC program, is that more than 70 percent of its students meet the eligibility criteria for federal meal assistance programs. Provided the BIC program is not targeted towards specific schools based on other determinants of academic outcomes, the staggered implementation and nearly complete treatment rate avoid the within-school treatment selection biases prevalent in the traditional SBP model. However, since school-choice options are available to students, between-school treatment selection may still exist. This and other identification concerns are addressed in the discussion of average treatment effects in Section 1.5.

One of the criticisms of moving the breakfast program into classroom time is that it will diminish the amount of instruction per classroom-day. To mitigate this impact, students are assigned various tasks required to distribute the prepared meals and dispose of any resulting waste. San Diego schools report an average of 15 minutes per school day towards the in-class breakfasts. Another criticism is that the program requires additional funds over the traditional SBP. Indeed, as illustrated in Figure 1.3, schools that have implemented BIC experienced an average increase of 183 percent in the proportion of students consuming school breakfasts. The BIC program does not necessarily cost more at the school or district level, however. All types of SBPs receive funding from multiple sources. As part of the Child Nutrition Act, federal funds are set aside annually to reimburse schools for meals provided to students on a per-meal basis. The amounts vary by the level of eligibility of each student and are outlined in Table 1.2. The state of California also reimburses public schools 26 cents for each meal provided to students qualifying for free or reduced price (FRP) meals. Finally, Title I funds are allocated to schools according to the proportion of students qualifying for FRP meals. In San Diego there are three tiers of Title I allocations. Schools with more than 85, 60, or 40 percent FRP qualifying students receive $434, $274, or $168 of Title I funds per student, respectively. Clearly, implementing BIC does not affect Title I allocations across the district, but does affect the amount of federal and state meal reimbursement funds received. School districts bear the financial burden for newly participating students who do not qualify for FRP meals but they also receive additional funds through increased participation among qualifying students after implementing BIC. Schools may actually decrease net breakfast expenditures through BIC when the number of newly participating qualified students is relatively larger than the number of students who do not qualify. This is likely to be the case in San Diego, where BIC implementation is restricted to schools where at least 70 percent of students qualify for FRP meals and previous participation among FRP-eligible students was relatively low.

## 1.4 Data

This study combines student level administrative data from elementary schools in San Diego Unified School District (SDUSD) with information on whether and when each school began the BIC program. The administrative records provide a longitudinal panel which tracks students' educational careers for as long as they attend any school within SDUSD. These data contain students' attendance records, report cards, standardized testing scores and personal characteristics such as gender, ethnicity, English-learner status and postal zip code of residence. The federal Child Nutrition Act prohibits anyone not directly associated with the administration of school meal programs from observing individual level data on school meal eligibility or participation. However, the proportion of students eligible for FRP meals and the number of meals served in each school is observed. The test scores used as outcomes are student results from the California Standards Test (CST) in English language arts (ELA) and mathematics. These exams are administered annually to all elementary students in grades 2 through 6. In the statistical analyses that follow, student scores are normalized using California statewide means and variances for each grade level and year. The CST replaced SAT9 testing in California during the 2001-2002 school year. Therefore, data from 2001-2002 through 2010-2011 school years are used when estimating the impacts on academic achievement and attendance. Because the behavioral measures used by teachers to evaluate students' classroom behavior changed in 2007, data from 2007 to 2011 are used to estimate the impact of BIC on classroom behavior. There are four key behavior measures, each rated on a scale between 1 and 3 in 0.25 increments: student interest level, student's level of respect shown in class, student's preparedness for class and student's assignment completion rate.

Table 1.3 presents summary statistics comparing schools selected for BIC in the first implementation year used and those not yet selected. Table 1.4 presents statistics of the same variables, but compares schools with and without Provision 2 status the year before BIC implementation began. As shown in Table 1.5, there are 65 elementary schools in San Diego where more than 70 percent of students qualify for free or reduced price meals. Of these schools, 45 eventually implement BIC by the end of the sample period.

## 1.5 Empirical Strategy

### 1.5.1 Research Design

In estimating the effect of school-supplied breakfast on academic outcomes, the general equation of interest is

$$Y_{igst} = \alpha_i + \gamma_s + \lambda_t + \varphi_g + \delta BIC_{st} + X'_{igst}\beta + u_{igst} \tag{1.1}$$

where $i$ indexes individual students, $g$ is grade level, $s$ indexes the particular school and $t$ is a year index. The outcomes $Y_{igst}$ of interest are achievement gains, attendance and classroom behavior. $BIC_{st}$ is an indicator for whether school $s$ has implemented the Breakfast in the Classroom program. This allows for the possibility that students are exposed to BIC in one year, but move schools and are not exposed in later years. $\delta$ is the average impact of providing breakfast to students in the classroom on the outcome of interest. Identification is driven by variation in the years that each school implements the BIC program, and requires that the treatment $BIC_{st}$ is independent of unobserved outcome determinants. Because standardized tests differ by grade level, $\varphi_g$, indicators for each grade level $g$ are included to control for grade-level-wide deviations in performance among students. School fixed effects $\gamma_s$ are included to control for school-specific time-invariant factors that influence student outcomes. Year fixed effects $\lambda_t$ control for yearly district-wide fluctuations in student outcomes. The remaining covariates and student fixed effects are progressively added to the baseline specification. $X_{igst}$ is a set of student level characteristics which includes gender, ethnicity, student's English learner status, indicators for zip code of residence, the attended school's year-round status and the percentage of students at attending school $s$ who are English learners in year $t$. $\alpha_i$ is a student fixed effect and $u_{igst}$ is an unobserved error component. Standard errors are clustered at the school level.

Since BIC implementation is staggered at the school level, principal fixed effects and teacher covariates $T_{igst}$, and school-specific linear time trends $\theta_s t$ are progressively added to equation (1.1). Teacher covariates included are teacher's years of experience, ethnicity and Cross-cultural Language and Academic Development (CLAD) certification[9] status for student $i$ at school $s$ during grade $g$ and year $t$. Principal fixed effects are also included in $T$ to control for school administrators, which may be correlated with both BIC implementation and student outcomes. School-specific linear time trends $\theta_s t$ are added to ensure the estimates of $\delta$ are not merely picking up increasing or decreasing trends in outcomes over time within each school. The full model is represented by

$$Y_{igst} = \alpha_i + \gamma_s + \lambda_t + \theta_s t + \varphi_g + \delta BIC_{st} + X'_{igst}\beta + T'_{igst}\phi + u_{igst} \qquad (1.2)$$

It is important to note that the implementation of BIC introduces one of two changes in each school. A portion of schools within the sample previously provided universally free breakfasts (UFB) to all enrolled students as a before-school option with voluntary participation. All other schools offered before-school breakfasts which were free only for students qualifying for free or reduced price meals by household income according to federal guidelines. Thus, estimates of $\delta$ are informing us of the weighted average net effect of BIC over the traditional SBP in some schools and merely moving universally free breakfasts into part of the classroom day in other schools. That is,

$$\begin{aligned} \delta &= E(Y_{i1} - Y_{i0}|X_i) \\ &= (1-\rho)E(Y_{i1} - Y_{i0} \mid UFB = 0, X_i) + \rho E(Y_{i1} - Y_{i0} \mid UFB = 1, X_i), \end{aligned}$$

where $\rho$ is the proportion of students attending schools which already offered UFB before BIC, $Y_{ij}$ is the outcome of student $i$ under BIC if $j = 1$ and without BIC if $j = 0$. Adding and subtracting the counter-factual outcome of previous UFB participators under no UFB participation, $Y_{i,UFB=0}$, we have

$$\begin{aligned} \delta = (1-\rho)E(Y_{i1} - Y_{i0} \mid UFB = 0, X_i) + \rho E(Y_{i1} - Y_{i,UFB=0} \mid UFB = 1, X_i) \\ -\rho E(Y_{i0} - Y_{i,UFB=0} \mid UFB = 1, X_i). \end{aligned}$$

[9] CLAD is certification for teaching students who are English-learners in the state of California.

Thus $\delta = ATE_{BIC} - \rho ATE_{UFB}$, the average treatment effect (ATE) of an in-class universal breakfast intervention on all students relative to the standard SBP, minus the participation-weighted ATE of the before-school UFB for students in Provision 2 schools prior to BIC. To the extent that UFB has a positive effect on student outcomes, $\delta$ will underestimate the average causal effect of breakfast interventions on all students. To separately estimate the impacts of BIC for students with and without a prior UFB, both equations (1.1) and (1.2) are also estimated with an interaction between an indicator for BIC and an indicator for schools having a pre-BIC UFB program under Provision 2. The resulting equation is

$$Y_{igst} = \alpha_i + \gamma_s + \lambda_t + \varphi_g + \delta_1 BIC_{st} + \delta_2 BIC_{st} \times Prov2_s + X'_{igst}\beta + u_{igst} \qquad (1.3)$$

where $\delta_2$ is the average net impact of BIC on $Y_{igst}$ for schools without a previous UFB program. Since prior offerings of UFB under Provision 2 are likely nonrandom and perhaps correlated with unobserved determinants of student outcomes, estimates of $\delta_2$ should not be interpreted as a differential causal impact of BIC given prior UFB offerings. Given the staggered implementation of BIC, however, $\delta_1 + \delta_2$ is the impact of BIC for schools previously offering before-school UFB while $\delta_1$ is the impact for schools previously using the traditional SBP.

Finally, there is the question as to what extent any impacts of the BIC program can be attributed to the resulting increased breakfast consumption. It is possible that the implementation of the BIC program changes the structure of classroom time and nature of interactions between classroom peers. To investigate impacts due to increased breakfast participation, rather than the implementation and administration of the program itself, the previous models are estimated again interacting the indicator $BIC_{st}$ with the resulting proportional increase in enrolled students eating school breakfasts. For reference, Figure 1.3 plots the proportional increase in the ratio of breakfasts consumed to student enrollment after implementing BIC, by the percentage of FRP eligible students at each BIC school. The points are separated by schools with and without previous UFB offerings. BIC results in between a 25 and nearly 400 percent increase in the proportion of students eating school breakfasts.

### 1.5.2 Accounting for Non-random Treatment Assignment

As stated above, at least 70 percent of a school's student body must qualify for free or reduced meals for a school to be eligible for the BIC program. To the extent that income levels are correlated with unobserved determinants of our outcomes, BIC assignment based on income levels results in a possibly endogenous treatment. Thus, the sample is restricted to all elementary schools eligible to receive the BIC program–those with at least 70 percent of students qualifying for FRP meals. This approach is similar to estimating treatment effects by matching treatment and control observations along covariates or propensity scores (Rubin, 1974; Rosenbaum and Rubin, 1983; Hirano and Imbens, 2001). Additionally, students who attend both schools within the sample and schools below 70 percent of students FRP eligible during the years of observation are dropped from the sample to avoid confounding estimates with outcome determinants specific to schools outside the estimation sample[10]. Because of concern about non-random selection, the pilot school and the first wave of 5 schools implementing BIC are dropped from the sample[11].

The identifying assumption for $\delta$ in the above equations is

$$E\left[Y_{0igst} \mid X_{igst}, s, t, BIC_{st}\right] = E\left[Y_{0igst} \mid X_{igst}, s, t\right].$$

In other words, identification of $\delta$ requires that the outcomes of interest are independent of BIC implementation, conditional on observables. If BIC targets low performing schools, estimates could be biased upwards from an ATE over all sample schools–particularly if low performing schools realize larger benefits from breakfast interventions. On the other hand, if some causes of poor performance in targeted schools are independent of nutrition and meal interventions and a limiting factor on student outcomes (e.g. low quality instruction), the estimated effect of BIC will be describe a local average treatment effect (LATE) specific to such schools. This type of targeted treatment would understate the effects of BIC on the average school

[10]This equated to roughly 12 percent of the treated sample.
[11]Including the first wave showed similar results with slightly larger estimate magnitudes.

with above 70 percent FRP meal-eligible students.

Tables 1.3 and 1.4 compare pre-BIC means and report normalized differences of outcomes between schools about to receive BIC and non-BIC schools, and schools with and without Provision 2 status prior to BIC. There do not appear to be large differences in outcomes between these groups before the program began. There are however, sizable differences in the average percentages of English learner (EL) and Hispanic students between these groups. Additionally, there are relatively smaller, but significant differences in the percentage of white students and average income levels. This motivates the inclusion of student English learner status, ethnicity, and both the percentage of EL and FRP-eligible students in the attending school each year as covariates.

Figure 1.2 displays the staggered timing of BIC implementation in San Diego, highlighting the spatial and temporal variation in the rollout of the program. To test for targeted timing of BIC implementation on the basis of school characteristics, the number of months from the inception of the BIC program to the time a school implemented BIC is regressed on observables. These results are reported in Appendix Table A.1. The only significant predictor of timing is the school's percentage of EL students. Fortunately, this association is not large. 5 percent more of a school's student body having EL status corresponds to a one month earlier implementation of BIC.

As another check, the following equation is used to test for targeted BIC implementation on the basis of student achievements,

$$Y_{igst} = \alpha_i + \gamma_s + \lambda_t + \varphi_g + \sum_{\tau=0}^{m} \delta_{-\tau} BIC_{s,t-\tau} + \sum_{\tau=1}^{q} \delta_{+\tau} BIC_{s,t+\tau} + X'_{igst}\beta + T'_{igst}\phi + u_{igst} \tag{1.4}$$

which is similar to equations (1.1) and (1.2), but includes lagged and lead BIC treatment dummies. The set of coefficients $\{\delta_{+\tau}\}$ is tested to determine whether past outcomes $Y_{igst}$ Granger-cause[12] BIC implementation. Although the absence of sta-

[12]See Granger (1969) for more on Granger causality testing.

tistical significance for these anticipatory effects does not exclude the possibility of BIC targeting, it is suggestive that the identifying assumption is satisfied. Additionally, the lagged periods allow an investigation into whether any effects of the BIC program persist or change over time. Indicators for each of the three years prior to a school beginning BIC are included. Because three full years of the BIC program are observed only for the earliest wave of the sample, an indicator for the third year of BIC would merely capture a wave-specific year effect. Therefore, an indicator for the first year of BIC and another combining the second and third years of BIC are used for post-implementation effects. An additional specification of (1.4) using the interactions of (1.3) is also used to differentially test these coefficients for schools that did and did not offer UFB prior to BIC. These results are presented in Table 1.6.

The lack of significance and relatively smaller magnitudes for the leads in Table 1.6 are reassuring that BIC is not targeted on the basis of CST performance. Likewise, the coefficients for the first and later years of BIC suggest that non-Provision 2 schools continue to realize positive gains in both ELA and math scores beyond the introduction year of the program. Indeed, the estimates for math and ELA are slightly larger in successive years, though not statistically different from the effects in the first year. These continued gains can be explained by some schools implementing BIC after the beginning of the academic year. Thus, the successive years are the first in which a full treatment schedule is administered. This suggests the impact from BIC is not merely a Figlio and Winicki (2005) style effect from meals being served immediately before testing periods. That is, the length of time over which breakfasts are served matters for test score gains.

### 1.5.3 Self-selection into Treatment via School Choice

As described above, the staggered implementation of the BIC program and its nearly complete treatment rate in the classroom makes it possible to sidestep within-school non-random treatment selection. The availability of school choice options, however, means it may be possible for students to select into a treatment school from an untreated school or vice versa. If treatment selection via school choice occurs, it is possible the estimates of interest will be biased. Figure 1.4 plots the proportion of students within BIC schools, separated by implementation year, that are exercising school choice over time. Years are normalized to the number of years since BIC was implemented. School choice options have increased in popularity district-wide over time, but there does not appear to be any discernible change in the rising rate of school choice, save for the proportion in wave 2 leveling off just before BIC was implemented.

To check for obvious patterns of school selection along the BIC status of schools, an indicator for student *i* attending her residential boundary school in year *t* is regressed on the BIC status of the boundary school and the variables of equations (1.2) and (1.3). This tests for any association between the BIC program and the proportion of students choosing to attend their local boundary school rather than exercising school choice. The first two columns of Table 1.7 show that implementing BIC at a student's local school weakly predicts a 1.3 percent increase in the likelihood a student attends that school, given it is not Provision 2 status. There is no significant association for students residing in a Provision 2 boundary.

In the third and fourth columns of Table 1.7, an indicator for student *i* at school *s* in year *t* exercising school choice[13] is regressed on the attended school's BIC status using equations (1.2) and (1.3). This second specification tests the association between a school's BIC status and changes in the proportion of the school's student body that have chosen that school place of the assigned boundary school. The resulting estimates suggest that schools with an active BIC program contain roughly 2.9% fewer students attending from outside the school boundary than other school. An average of about one third of students in San Diego public schools exercise a school choice option. These estimates agree with that number, as an increase in the proportion of local resident students attending any given school decreases the

[13]School choice is defined as a student attending a school different than the location determined by her address of residence and school zone boundaries.

percentage of students within that school who are exercising school choice by twice that amount. Although the estimates of student movement are statistically significant, the magnitudes do not suggest a large selection bias will result from these movements[14].

## 1.6 Results

### 1.6.1 Impact of BIC on Standardized Test Scores

Table 1.8 reports the regression results for student gains in CST math and English language arts (ELA) scores. CST scores are standardized using statewide means and variances separately by grade level and year. Units of the estimated average impacts of BIC are standard deviations. To check the sensitivity of the estimates, student fixed effects, principal fixed effects and teacher covariates, followed by school specific linear time trends are progressively added to pairwise columns moving left to right.

Estimates of the overall average net impact of BIC on ELA gains range between 3 and 5 percent of a standard deviation, but are not statistically different from zero. Including an interaction for previous UFB (Provision 2), however, reveals gains of roughly 11 percent of a standard deviation from BIC in schools without prior UFB. On the other hand, there does not appear to be any significant impact for students in schools offering UFB prior to BIC implementation. Similarly for gains in math, BIC is associated with a roughly 15 percent standard deviation increase for non-Provision 2 schools, even after accounting for school administration, teachers and school-specific trends over time, although significance does fall to the 10 percent level with linear time trends. Again, there is no significant impact for Provision 2 schools.

[14]For instance, consider the case where the newly attending boundary students–making up 1.5 percent of the treatment group–positively select into the BIC program. If each of these students realize an unusually large 0.5 standard deviation increase in math score gains as a result, their inclusion would inflate the coefficient estimates of about 0.15 in Table 1.8 by $(.0015 * 0.5)/(0.146 - .0015 * 0.5)$, or about 5.4 percent.

These results suggest a significantly positive impact from school breakfasts on academic performance when meals are offered universally free. On the other hand, moving the meal into the classroom schedule may not meaningfully change this impact. This could be interpreted in two ways. First, it would appear that those who benefit significantly from school breakfasts are likely to use the program when it is offered universally free, regardless of whether it is before school or administered during classroom time. This interpretation suggests there may be a social stigma surrounding free meals when eligibility is determined by income levels, perhaps discouraging voluntary participation in students who would benefit from school breakfasts. A second interpretation is that the BIC program only increases student gains in non-Provision 2 schools because students in these schools differ in treatment efficacy for some unobserved reason. There is no evidence to support large differences on average between these two groups, however. As seen in Table 1.4, there only appear to be significant differences in the average proportions of EL students and FRP-eligible students, which are controlled for. Although the average percentage of FRP-eligible students is roughly 9 percent lower for non-Provision 2 schools, Figures 1.1 and 1.3 show that non-Provision 2 schools are not distributed along one particular side of FRP eligibility for BIC participation levels or increases in breakfast participation when compared to Provision 2 schools. Therefore, the first interpretation seems more plausible.

### 1.6.2 Impact of BIC on Attendance Rates

Table 1.9 presents estimates of the impact of BIC on student attendance in elementary school. The first panel presents estimates for the impact of BIC on the percentage of days a student is absent during a given academic year. The unit here for the dependent variable is 1 percent, or nearly 2 days[15]. The dependent variable of the second panel is an indicator for the student being chronically absent, defined as being absent at least 10 percent of instructional days during the school year.

[15] There are 180 instructional days in an academic year.

Interestingly, there is no effect of BIC on attendance in either type of school. The magnitudes of these estimates are quite small as well, given that the units are percentage of days absent. These estimates are relevant for policy, as proponents of universal school meal programs have touted their potential to reduce tardiness and absences through both improved nutrition and direct incentive for attendance by meal provision[16]. This result seems intuitive given that moving breakfast from before school to during class time eliminates the meal incentive for early arrival. This does, however, suggest that improved attendance is not a mechanism through which BIC may improve student achievement.

### 1.6.3 Impact of BIC on Classroom Behavior

Results for impacts on student behavior are presented in Table 1.10. The behavioral measures on which students are evaluated are described above in Section 1.4. Each behavior measure is standardized by the district-wide means for each grade level. The coefficient estimate units are therefore standard deviations.

There are significant gains in student preparedness and respectfulness scores associated with moving universally free breakfasts from before school into the classroom. However, this effect is not found when implementing BIC in schools where UFB was not offered. The implementation of BIC is associated with an increase of just over half a standard deviation in student respectfulness scores at Provision 2 schools. Similarly, scores of student preparedness see an increase of over 30 percent of a standard deviation in Provision 2 schools.

Table 1.11 similarly tests coefficients for behavioral outcomes during different years of BIC. Unfortunately because the behavioral measures changed in San Diego just before BIC implementation, it is not possible to test for anticipatory effects of BIC with respect to behavioral scores. As in the other specifications, there are positive gains in some behavioral scores for schools offering UFB prior to BIC. These schools see modest increases in gains for student respect scores and

[16]See, for instance, the BIC fact sheet at the Food Research and Action Center: http://frac.org/wp-content/uploads/2009/09/universal_classroom_breakfast_fact_sheet.pdf

assignment completion. Again, the significant gains are in student respect and preparedness, with about 0.25 and 0.4 standard deviation increases, respectively, in the first and subsequent years of BIC.

There are two potential tradeoffs for students when moving breakfast from before school into the classroom. The first is that students no longer need to arrive as early to school to receive breakfast. This potentially allows students more time in the mornings for sleep or to prepare for school. Although not directly observable, these changes may contribute to better preparedness and behavior in the classroom. The second, is that conditional on arrival time not changing, students now have more socializing and leisure time before the school day begins. Moreover, students are no longer segregated by breakfast participation before school. Rather, almost all students participate in school breakfasts inside the classroom. The integration of students across income strata within a classroom may be one way in which the level of respect students show for others is impacted. That the impacts of BIC on student preparedness and respect are significantly larger for Provision 2 schools is consistent with the higher previous breakfast participation rates in these schools.

### 1.6.4 Are Gains Driven by Increased Breakfast Consumption?

It is unclear whether the impacts of BIC are driven by the consumption of school breakfasts or by the changes the program makes to the structure and administration of the school day. Table 1.12 repeats the specifications of Table 1.8, but includes an interaction of $BIC_{st}$ with the proportional increase in students eating breakfasts in each school after BIC began. A 100 percent increase in the fraction of students participating in school breakfasts roughly corresponds to a 0.05 standard deviation increase in ELA gains within schools that did not previously offer universally free breakfasts. When looking at the impact on math gains, a 100 percent increase corresponds to just over 0.08 standard deviation gains and about 0.07 standard deviations when including school specific linear time trends. Given that the average increase is nearly 200 percent, these estimates agree with those of Table

1.8, which are roughly twice as large.

Table A.2 in the appendix presents similar estimates using the increase in breakfast participation, but for lead and lagged years of the BIC program. The results closely match those of Tables 1.6 and 1.8. This suggests student gains are increasing in breakfast participation and are primarily driven by breakfast consumption, rather than the administration of the BIC program.

### 1.6.5 Do Impacts of BIC Differ Along the Distribution of Past Student Achievement?

To check for heterogeneous effects of BIC across the distribution of student achievement, students are assigned to fixed quartile groups determined by past achievement levels in ELA and math. The same specifications as Table 1.8 are run separately for each quartile. Tables 1.13 and 1.14 report the results for the main specifications including student fixed effects, principal fixed effects, teacher covariates and controls for school-specific outcome trends over time. Focusing on the full specification in the fourth column under each quartile, it becomes apparent that the overall results are primarily driven by impacts to students in the lower half of the achievement distribution. The majority of the impact in ELA gains corresponds to students in the second quartile of past ELA achievement. Among these students in schools not previously offering universally free breakfasts, the impact to ELA gains from BIC is an increase of roughly a quarter of a standard deviation. For gains in mathematics, the impact for schools not previously offering universally free breakfasts appears to be driven predominately by the first two quartiles of previous achievement levels. BIC corresponds to about a 0.2 standard deviation increase within the first quartile, with an imprecise estimate of 0.18 standard deviations for the second quartile. , There is some evidence of a positive impact in the upper quartile as well, though estimates for this quartile decrease in magnitude and precision once school-specific linear time trends are included.

These results are in line with some of the plausible mechanisms described in

Section 1.1, through which BIC can impact student achievement. First, we would expect the achievement impacts from BIC to be largest among lower performing students if inadequate nutrition in the morning is indeed a significant factor in diminished academic performance. Second, to the extent that achievement levels are correlated with household income, the proportionally larger increase in disposable income for low income households resulting from school-provided breakfasts will be more pronounced among lower-performing students. Third, BIC eliminates the segregation of school breakfast participants from non-participants before class begins. If this segregation line is correlated with student achievement, increased peer interaction throughout the achievement distribution after BIC implementation may particularly benefit students in lower achievement quantiles as well. The empirical framework of this paper does not allow for disentanglement or separate estimation of the impacts through each of these potential mechanisms. However, the fact that the impacts of BIC on student gains are most pronounced for students in the lower achievement quantiles within schools exhibiting lower household income levels suggests universal breakfast programs such as BIC have the potential to mitigate achievement gaps across socioeconomic strata.

## 1.7 Conclusion

This paper exploits the staggered implementation of a new, "Breakfast in the Classroom" program (BIC) in estimating the causal effects of providing universally free school breakfasts on academic achievement, student attendance rates and classroom behavior. The results show that implementing an in-class, universally free breakfast program in elementary schools with relatively low income levels increases breakfast consumption in these schools by an average of 183 percent. The BIC program increases English-language arts and math gains by an average of 11 and 15 percent of a standard deviation, respectively, in schools that did not previously offer universally free breakfasts. The overall impacts to achievement gains in these schools are driven by the impacts to students at the lower end of the achievement distribution. The gains from BIC persist after the first, partial year of the program. Moreover, estimates of gains in later years are larger than in the initial year, though the magnitudes are not statistically different. This suggests benefits to achievement occur throughout the year and are not driven solely by meals at the time of standardized testing. There are no significant changes in gains, however, for schools which already offered universally free meals outside the classroom prior to BIC.

In-class breakfasts do not impact school attendance rates, regardless of whether or not the school already offered universally free breakfasts. Moving universally free breakfasts from before school into the classroom increases gains in some student behavioral scores, particularly in students' level of preparedness for the school day. This result agrees with notions that additional free time previously forgone when eating breakfast before school is beneficial to classroom behavior, perhaps through increased sleep or preparation time.

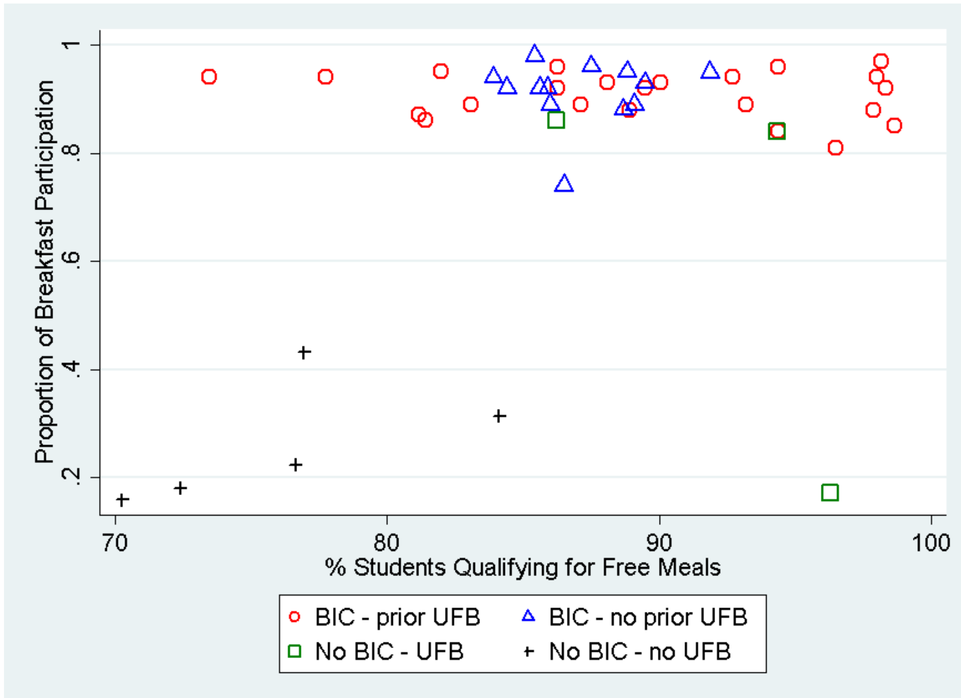
These findings suggest school breakfasts are under-utilized by students who would benefit from them when they are not offered universally free, perhaps because of social stigmas surrounding low-income requirements for free meal eligibility. Additionally, making breakfasts universally free in schools with low income levels is a relatively inexpensive method of achieving test score gains and may help in reducing achievement gaps across income levels. The estimated gains in math from BIC are at least as large as those from increasing teacher quality by one standard deviation (Rockoff, 2004; Aaronson, Barrow and Sander, 2007; Dobbie, 2011) and greater than those from decreasing classroom size by ten students (Rivken, Hanushek and Kain, 2005). As of June 2012, school breakfasts cost state and federal governments a maximum combined total of $2.06 in reimbursements per meal. Average costs to San Diego elementary schools for these meals before reimbursement are between $1.08 and $1.20 per meal, including labor and equipment costs. Given 180 instruction days per academic year, this corresponds to an average cost of between $195 and $216 per newly participating student, per year, with the portion paid by government reimbursements depending on the distribution of household incomes among schools. For comparison, reducing class size from 30 to 20 students is estimated to cost 915 dollars per student and about 435 dollars per student for a reduction from 24 to 20 students (Reichardt, 2000). Given the relatively higher costs of reducing classroom size or increasing teacher quality, providing universally free meals in low income schools appears to be a compelling policy for mitigating achievement gaps across socioeconomic backgrounds.

Considering the literature on achievement gains and earnings, the results here also indicate that providing universally free breakfasts is a policy with the potential for high returns. For instance, Chetty, Friedman, and Rockoff (2011) estimate that a one standard deviation increase in teacher value-added for one year leads to a 25,000 dollar increase in the average student's lifetime earnings–a present discounted value of about 4,600 dollars assuming a 5 percent discount rate. Given that the estimated gains from BIC are on par with those of a one standard deviation increase in teacher quality, the discounted per-dollar return to providing universally free meals in terms of future earnings is over 12 dollars in terms of federal funding. The estimated per-dollar return is much greater–roughly 21–in terms of average per-student spending at the school level.

Further research on school meal programs is needed. From a policy perspective, it is useful to quantify any impacts of universally free meal programs in schools with higher income levels. If the large impacts found here can also be had in wealthier schools, it may warrant policy revisions or expansions of Provision 2 of the National School Lunch Act to incentivize wider user of universally free school meals. Additionally, more research is needed on the existence of social stigmas associated with assistance programs to better understand the determinants of participation. There is also much to be learned about if and how the effects of these programs are related to the nutritional quality of the foods served. Future such additions to the literature will help inform the current policy debates on reforming the content, delivery and funding of school meals.

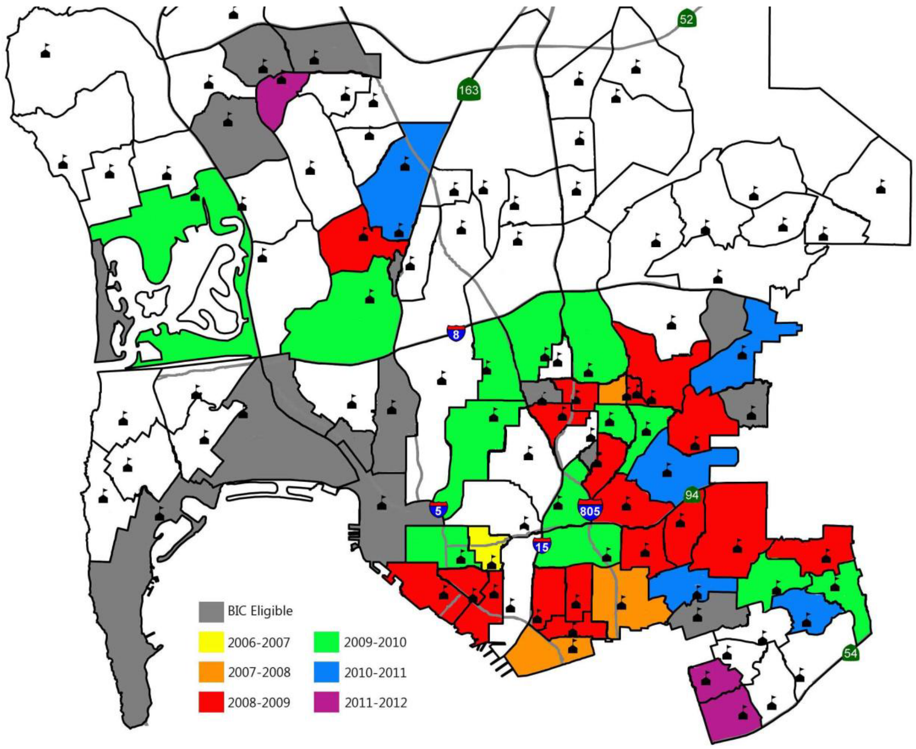
This chapter will be prepared for future submission for publication.

## 1.8 Chapter 1 Figures and Tables



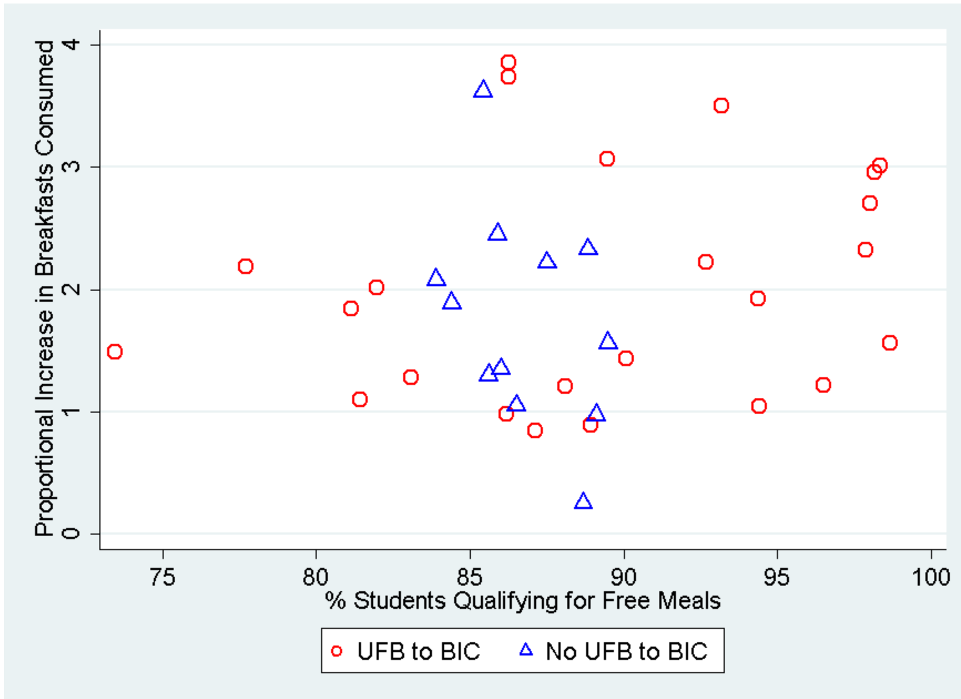

Notes: Figure presents average proportion of students within each sample school participating in school breakfasts. Symbols correspond to different BIC status and pre-BIC universally free meal status. Averages are calculated using 2010-2011 academic year data.

**Figure1.1**: Proportion of Students Participating in School Breakfast by School



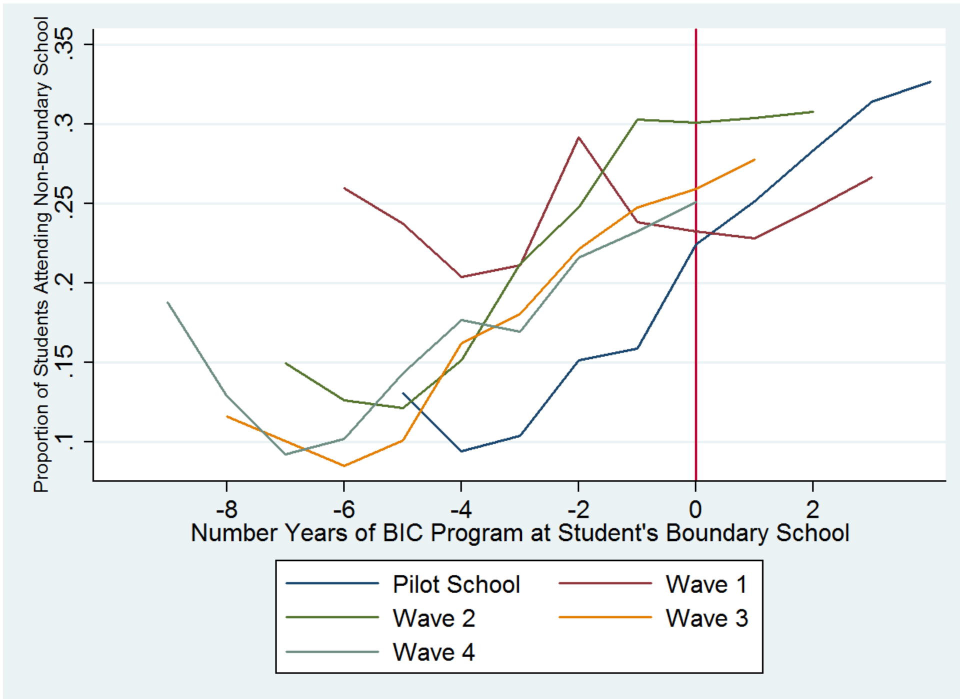

Notes: Figure presents implementation timing of the BIC program across San Diego elementary schools. BIC eligible schools are schools where at least 70 percent of the student body qualifies for free school meals. Areas in grey are eligible schools that have yet to have implemented the BIC program as of July 2012.

**Figure1.2**: Implementation of Breakfast in the Classroom Program in San Diego




Notes: Figure presents proportional increase of students participating in school breakfasts after BIC implementation, by school. Symbols differentiated schools that did or did not offer universally free meals prior to BIC. Units of increase on the vertical axis are 100 percent increases in the proportion of enrolled students eating school breakfasts. Increases are calculated using 2010-2011 academic year data.

**Figure1.3**: Change in the Proportion of School Breakfast Participators After BIC




Notes: Figure presents proportion of students exercising school choice by attending a school other than the location determined by residence, over time. Years are normalized by year of the BIC program. Lines are plotted separately by year the BIC program was implemented.

**Figure1.4**: Students Exercising School Choice, by BIC Implementation Wave

**Table1.1**: Free or Reduced Price Meal Eligibilities

**Federal Income Guidelines for Free or Reduced Price Meals (*as of July 2011*)**

| | Federal Poverty Guidelines | Maximum Household Income (*US Dollars*) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Free Meals (130% of Poverty Guideline) | | | Reduced Price Meals (185% of Poverty Guideline) | | |
| Household Size | Annual | Annual | Monthly | Weekly | Annual | Monthly | Weekly |
| 1 | 10,890 | 14,157 | 1,180 | 273 | 20,147 | 1,679 | 388 |
| 2 | 14,710 | 19,123 | 1,594 | 368 | 27,214 | 2,268 | 524 |
| 3 | 18,530 | 24,089 | 2,008 | 464 | 34,281 | 2,857 | 660 |
| 4 | 22,350 | 29,055 | 2,422 | 559 | 41,348 | 3,446 | 796 |
| 5 | 26,170 | 34,021 | 2,836 | 655 | 48,415 | 4,035 | 932 |
| 6 | 29,990 | 38,987 | 3,249 | 750 | 55,482 | 4,624 | 1,067 |
| 7 | 33,810 | 43,953 | 3,663 | 846 | 62,549 | 5,213 | 1,203 |
| 8 | 37,630 | 48,919 | 4,077 | 941 | 69,616 | 5,802 | 1,339 |

Source: Federal Register, Vol. 76, No. 58, 5/25/11, p. 16725.

Notes: Threshold income levels adjusted annually based on the Consumer Price Index. Reported levels apply to the 48 contiguous United States, the District of Columbia, Guam and the Territories.

**Table1.2**: Federal School Breakfast Reimbursement Rates

| | **July 2011 - June 2012** | | |
|---|---|---|---|
| | Non-Severe Need | Severe* Need | Price to Student (San Diego) |
| Free | 1.51 | 1.80 | 0.00 |
| Reduced Price | 1.21 | 1.50 | 0.00 |
| Paid | 0.27 | 0.27 | 1.00 |

Source: Federal Register, Vol. 76, No. 139, 7/20/11, p. 43259.

Notes: Reported reimbursement rates are in US dollars and apply to the 48 contiguous United States, the District of Columbia, Guam and the Territories.
* Schools where at least 40 percent of lunches served two years prior were free or reduced price qualify as "severe need" and receive higher rates of reimbursement.

**Table1.3**: Summary Statistics: Last Year Before Implementation

| | Never Selected | Selected for BIC | Difference |
|---|---|---|---|
| CST Math | -0.335 | -0.343 | 0.007 |
| | (0.084) | (0.034) | (0.075) |
| CST ELA | -0.378 | -0.442 | 0.064 |
| | (0.084) | (0.035) | (0.076) |
| % Days Absent | 4.948 | 4.415 | 0.533 |
| | (0.500) | (0.142) | (0.393) |
| Behavior: Respects | -0.442 | -0.28 | -0.162 |
| | (0.358) | (0.098) | (0.262) |
| Behavior: Interest | -0.36 | -0.366 | 0.006 |
| | (0.335) | (0.077) | (0.228) |
| Behavior: Preparedness | -0.608 | -0.278 | -0.33 |
| | (0.371) | (0.078) | (0.243) |
| Behavior: Completes | -0.509 | -0.284 | -0.226 |
| | (0.344) | (0.075) | (0.229) |
| % Free or Reduced Price | 79.503 | 87.335 | -7.832*** |
| | (2.001) | (1.064) | (2.077) |
| English Learner | 0.444 | 0.605 | -0.161*** |
| | (0.053) | (0.025) | (0.051) |
| Teacher Experience | 12.728 | 11.814 | 0.915 |
| | (1.413) | (0.521) | (1.218) |
| Black | 0.259 | 0.144 | 0.115** |
| | (0.071) | (0.019) | (0.055) |
| Hispanic | 0.546 | 0.726 | -0.181*** |
| | (0.063) | (0.028) | (0.059) |
| White | 0.091 | 0.034 | 0.057*** |
| | (0.027) | (0.006) | (0.020) |

*, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses. Estimates averaged during year before rollout of BIC program, separated by treated and non-treated schools.

**Table1.4**: Summary Statistics: Comparison by Provision 2 Status

| | Not Provision 2 | Pre-BIC Provision 2 | Difference |
|---|---|---|---|
| CST Math | -0.272 | -0.33 | 0.058 |
| | (0.067) | (0.055) | (0.087) |
| CST ELA | -0.306 | -0.423 | 0.118 |
| | (0.079) | (0.037) | (0.088) |
| % Days Absent | 4.345 | 4.319 | 0.025 |
| | (0.351) | (0.215) | (0.416) |
| % Free or Reduced Price | 80.087 | 89.512 | -9.425*** |
| | (1.426) | (1.565) | (2.113) |
| English Learner | 0.474 | 0.685 | -0.212*** |
| | (0.048) | (0.018) | (0.052) |
| Teacher Experience | 11.948 | 11.174 | 0.774 |
| | (1.333) | (0.764) | (1.556) |
| Black | 0.197 | 0.103 | 0.094* |
| | (0.042) | (0.020) | (0.047) |
| Hispanic | 0.590 | 0.790 | -0.200*** |
| | (0.055) | (0.034) | (0.065) |
| White | 0.128 | 0.026 | 0.101** |
| | (0.046) | (0.007) | (0.048) |

*, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses. Estimates averaged during year before rollout of BIC program for Provision 2 and non-Provision 2 schools.

**Table1.5**: Tabulation of Schools by Breakfast Program

| | BIC School | | |
|---|---|---|---|
| | No | Yes | Both |
| Not Provision 2 | 16 | 19 | 35 |
| Pre-BIC Provision 2 | 6 | 26 | 32 |
| Both | 22 | 45 | 65 |

Notes: Includes all schools where at least 70% of students qualify for free or reduced price meals.

## **Table1.6**: Lead and Lagged Estimates of Student Gains

| | Math Gains | | | | | | ELA Gains | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *BIC(t-3)* | 0.070 | 0.025 | | 0.086 | 0.029 | | 0.012 | -0.025 | | 0.021 | -0.041 | |
| | (0.049) | (0.055) | | (0.054) | (0.057) | | (0.037) | (0.045) | | (0.037) | (0.048) | |
| *BIC(t-3) + BIC(t-3)*UFB* | | 0.080 | | | 0.098 | | | 0.025 | | | 0.042 | |
| | | (0.052) | | | (0.062) | | | (0.041) | | | (0.047) | |
| *BIC(t-2)* | 0.047 | -0.003 | | 0.097 | 0.033 | | -0.013 | -0.028 | | 0.010 | -0.044 | |
| | (0.085) | (0.085) | | (0.094) | (0.098) | | (0.054) | (0.065) | | (0.054) | (0.070) | |
| *BIC(t-2) + BIC(t-2)*UFB* | | 0.025 | | | 0.066 | | | -0.036 | | | -0.008 | |
| | | (0.087) | | | (0.093) | | | (0.058) | | | (0.055) | |
| *BIC(t-1)* | 0.070 | 0.099 | | 0.179 | 0.199 | | 0.019 | 0.015 | | 0.057 | -0.011 | |
| | (0.107) | (0.113) | | (0.132) | (0.139) | | (0.074) | (0.077) | | (0.081) | (0.097) | |
| *BIC(t-1) + BIC(t-1)*UFB* | | -0.031 | | | 0.061 | | | -0.028 | | | 0.002 | |
| | | (0.112) | | | (0.133) | | | (0.082) | | | (0.084) | |
| *BIC Year 1* | 0.114 | 0.256* | 0.168*** | 0.333* | 0.436** | 0.233*** | 0.064 | 0.15 | 0.137*** | 0.144 | 0.153 | 0.165*** |
| | (0.139) | (0.143) | (0.058) | (0.172) | (0.169) | (0.059) | (0.099) | (0.104) | (0.046) | (0.105) | (0.121) | (0.050) |
| *BIC Year 1 + BIC Year 1*UFB* | | -0.024 | -0.012 | | 0.135 | 0.031 | | -0.018 | 0.003 | | 0.008 | 0.002 |
| | | (0.143) | (0.056) | | (0.196) | (0.086) | | (0.112) | (0.040) | | (0.126) | (0.055) |
| *BIC Years 2-3* | 0.162 | 0.306 | 0.204** | 0.427** | 0.520** | 0.298*** | 0.126 | 0.212* | 0.191*** | 0.259** | 0.236 | 0.244*** |
| | (0.169) | (0.188) | (0.101) | (0.202) | (0.213) | (0.106) | (0.111) | (0.115) | (0.062) | (0.123) | (0.148) | (0.077) |
| *BIC Year 2-3 + BIC Year 2-3*UFB* | | 0.056 | 0.052 | | 0.258 | 0.134 | | 0.062 | 0.073 | | 0.147 | 0.138* |
| | | (0.172) | (0.080) | | (0.219) | (0.101) | | (0.122) | (0.059) | | (0.139) | (0.071) |
| Observations | 67,157 | 67,157 | 67,157 | 67,157 | 67,157 | 67,157 | 66,461 | 66,461 | 66,461 | 66,461 | 66,461 | 66,461 |
| F-test p-values | | | | | | | | | | | | |
| BIC(t-3)=BIC(t-2)=BIC(t-1)=0 | 0.245 | | | 0.242 | | | 0.519 | | | 0.656 | | |
| BIC(t)=BIC(t+1,2)=0 | 0.609 | | | 0.11 | | | 0.231 | | | 0.032 | | |
| All BIC(t) = 0 | 0.335 | 0.065 | | 0.024 | 0.005 | | 0.305 | 0.02 | | 0.069 | 0.012 | |
| BIC(t)=BIC(t+1,2)=BIC*UFB=0 | | 0.125 | 0.071 | | 0.047 | 0.008 | | 0.09 | 0.01 | | 0.042 | 0.013 |
| All BIC(t) and BIC(t)*UFB = 0 | | 0.032 | | | 0.004 | | | 0.022 | | | 0.03 | |
| Principal FE / Teachers | No | No | No | Yes | Yes | Yes | No | No | No | Yes | Yes | Yes |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student gains in math and English language arts CST scores. Scores are normalized using statewide mean and variance for each grade and year. Each model includes school, grade, year and student fixed effects, indicators for school type and year-round status, and a set of student characteristic covariates. Reported coefficients are for lead and lagged years of treatment, separately by pre-BIC universally free meal offerings.

**Table1.7**: Changes in School Choice During BIC Implementation

| | Dependent Variable | | | |
|---|---|---|---|---|
| | Attending Local School | | Exercising School Choice | |
| BIC at Local School | 0.0153* | 0.0129* | | |
| | [0.00869] | [0.00695] | | |
| BIC at Local * Local is Provision2 | | 0.00383 | | |
| | | [0.0145] | | |
| Local School is Provision2 | | -0.000689 | | |
| | | [0.0139] | | |
| BIC at Attending School | | | -0.0356*** | -0.0287*** |
| | | | [0.0110] | [0.0104] |
| BIC at Attending* Attending is Provision2 | | | | -0.00745 |
| | | | | [0.0201] |
| Attending is Provision2 | | | | -0.0632*** |
| | | | | [0.0223] |
| N | 152,786 | 152,786 | 152,786 | 152,786 |
| p-values | | | | |
| BIC at Local + Interaction = 0 | | 0.2998 | | |
| BIC + Interaction = 0 | | | | 0.0469 |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in brackets are clustered at the school level. Dependent variables are 1) an indicator for the student attending the boundary school assigned by address of residence, and 2) an indicator for the student attending a school other than her boundary school. Each model includes all fixed effects, covariates and school-specific linear time trends specified in the previous regressions.

**Table1.8**: Estimates of Student Achievement Gains

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | CST ELA Gains | | | | | | | |
| *BIC* | 0.008 | 0.088*** | 0.031 | 0.108** | 0.051 | 0.114** | 0.046 | 0.112* |
| | (0.029) | (0.033) | (0.034) | (0.045) | (0.044) | (0.051) | (0.046) | (0.063) |
| *BIC + BIC*UFB* | | -0.018 | | -0.005 | | -0.001 | | -0.016 |
| | | (0.032) | | (0.038) | | (0.054) | | (0.069) |
| Observations | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 |
| | CST Math Gains | | | | | | | |
| *BIC* | 0.01 | 0.095** | 0.029 | 0.141** | 0.102* | 0.193*** | 0.094 | 0.147* |
| | (0.040) | (0.048) | (0.048) | (0.055) | (0.061) | (0.055) | (0.065) | (0.080) |
| *BIC + BIC*UFB* | | -0.017 | | -0.022 | | 0.025 | | 0.046 |
| | | (0.046) | | (0.056) | | (0.086) | | (0.090) |
| Observations | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 |
| Student FE | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Principal FE/Teachers | No | No | No | No | Yes | Yes | Yes | Yes |
| School Time Trends | No | No | No | No | No | No | Yes | Yes |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student-level gains in math and English language arts CST scores. Scores are normalized using statewide mean and variance for each grade and year. Each model includes school, grade, and year fixed effects, indicators for school type and year-round status, as well as a set of student level covariates. Reported coefficients are 1) the impact of BIC for schools not previously offering universally free meals, and 2) the sum of coefficients that are the impact for schools previously offering universally free meals.

**Table1.9**: Estimates of Student Absences

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Percentage of School Days Absent | | | | | | | |
| *BIC* | -0.050 | -0.009 | -0.100 | -0.024 | 0.017 | 0.028 | 0.132 | 0.105 |
| | (0.146) | (0.144) | (0.125) | (0.128) | (0.134) | (0.118) | (0.145) | (0.115) |
| *BIC + BIC*UFB* | | -0.076 | | -0.144 | | -0.010 | | 0.145 |
| | | (0.157) | | (0.138) | | (0.205) | | (0.222) |
| | Chronic Absentee | | | | | | | |
| *BIC* | -0.003 | -0.006 | -0.010 | -0.011 | -0.008 | -0.009 | -0.001 | 0.002 |
| | (0.008) | (0.009) | (0.006) | (0.008) | (0.007) | (0.008) | (0.008) | (0.010) |
| *BIC + BIC*UFB* | | -0.001 | | -0.009 | | -0.006 | | -0.003 |
| | | (0.008) | | (0.007) | | (0.010) | | (0.011) |
| Observations | 157,295 | 157,295 | 157,295 | 157,295 | 157,295 | 157,295 | 157,295 | 157,295 |
| Student FE | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Principal FE/Teachers | No | No | No | No | Yes | Yes | Yes | Yes |
| School Time Trends | No | No | No | No | No | No | Yes | Yes |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are percentage of days in academic school year student was absent and whether student was chronically absent during the school year (defined as absent at least 10 percent of school days). Each model includes school, grade, and year fixed effects, indicators for school type and year-round status, as well as a set of student level covariates. Reported coefficients are 1) the impact of BIC for schools not previously offering universally free meals, and 2) the sum of coefficients that are the impact for schools previously offering universally free meals.

**Table1.10**: Estimates of Student Behavior Gains

| | Respects People and Property | | Shows Interest in Learning | | Prepares and Organizes | | Completes Assignments When Due | |
|---|---|---|---|---|---|---|---|---|
| *BIC* | 0.093 | -0.155 | 0.182 | 0.111 | 0.194 | 0.089 | -0.096 | -0.185 |
| | (0.262) | (0.159) | (0.150) | (0.160) | (0.137) | (0.177) | (0.132) | (0.157) |
| *BIC + BIC*UFB* | | 0.603** | | 0.284 | | 0.325*** | | 0.071 |
| | | (0.266) | | (0.217) | | (0.111) | | (0.137) |
| Observations | 38,180 | 38,180 | 38,227 | 38,227 | 38,208 | 38,208 | 38,219 | 38,219 |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student gains in individual in-class behavioral scores. Scores are normalized using district-wide mean and variance for each gradelevel. Each model includes school, grade, year and student fixed effects; indicators for school type and year-round status; a set of student level covariates; teacher ethnicity, years of experience and certification level; principal fixed effects and school-specific linear time trends. Reported coefficients are 1) the impact of BIC for schools not previously offering universally free meals, and 2) the sum of coefficients that are the impact for schools previously offering universally free meals.

**Table1.11**: Lead-Lagged Estimates of Student Behavior Gains

| | Respects People and Property | Shows Interest in Learning | Prepares and Organizes | Completes Assignments When Due |
|---|---|---|---|---|
| *BIC Year 1* | 0.053 | 0.011 | 0.044 | -0.022 |
| | (0.069) | (0.097) | (0.068) | (0.057) |
| *BIC Year 1 + BIC Year 1*UFB* | 0.223*** | 0.066 | 0.386*** | 0.196** |
| | (0.069) | (0.070) | (0.102) | (0.075) |
| *BIC Years 2-3* | 0.137 | -0.053 | 0.076 | 0.047 |
| | (0.087) | (0.125) | (0.119) | (0.091) |
| *BIC Year 2-3 + BIC Year 2-3*UFB* | 0.273*** | 0.069 | 0.418*** | 0.225** |
| | (0.086) | (0.095) | (0.124) | (0.106) |
| Observations | 38,628 | 38,675 | 38,655 | 38664 |
| F-test p-values | | | | |
| All BIC coef. = 0 | 0.025 | 0.803 | 0.005 | 0.171 |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student gains in individual in-class behavioral scores. Scores are normalized using district-wide mean and variance for each gradelevel. Each model includes school, grade, year and student fixed effects, indicators for school type and year-round status, a set of student level covariates, teacher ethnicity, years of experience and certification level, and school principal fixed effects. Reported coefficients are 1) the impact of BIC for schools not previously offering universally free meals, and 2) the sum of coefficients that are the impact for schools previously offering universally free meals. Coefficients are estimated separately for the first year of treatment and the following years of treatment. Units are standard deviation changes.

**Table1.12**: Increases in Breakfast Consumption and Student Achievement Gains

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | CST ELA Gains | | | | |
| *BIC * % change-in-meals* | 0.014 | 0.052*** | 0.025* | 0.057** | 0.023 | 0.051** | 0.022 | 0.048* |
| | (0.010) | (0.016) | (0.014) | (0.022) | (0.025) | (0.024) | (0.026) | (0.026) |
| *BIC*change + BIC*change*UFB* | | 0.008 | | 0.016 | | 0.000 | | -0.001 |
| | | (0.010) | | (0.017) | | (0.036) | | (0.046) |
| Observations | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 | 65,951 |
| | | | | CST Math Gains | | | | |
| *BIC * % change-in-meals* | 0.014 | 0.056** | 0.025 | 0.079*** | 0.05 | 0.089*** | 0.052 | 0.065* |
| | (0.013) | (0.022) | (0.021) | (0.025) | (0.035) | (0.027) | (0.036) | (0.038) |
| *BIC*change + BIC*change*UFB* | | 0.008 | | 0.009 | | 0.019 | | 0.042 |
| | | (0.014) | | (0.024) | | (0.054) | | (0.056) |
| Observations | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 | 66,634 |
| Student FE | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Principal FE / Teachers | No | No | No | No | Yes | Yes | Yes | Yes |
| School Time Trends | No | No | No | No | No | No | Yes | Yes |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student gains in math and English language arts CST scores. Scores are normalized using statewide mean and variance for each grade and year. Each model includes school, grade, and year fixed effects, indicators for school type and year-round status, as well as a set of student level covariates. Reported coefficients are 1) the impact of BIC for schools not previously offering universally free meals, and 2) the sum of coefficients that are the impact for schools previously offering universally free meals.Units are standard deviation changes per 100 percent increase in the proportion of enrolled students eating breakfast.

**Table1.13**: Estimates of ELA Gains by Achievement Quartile

| | CST ELA Gains | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | First Quartile | | | | Second Quartile | | | |
| *BIC* | 0.006 | 0.092 | -0.028 | 0.016 | 0.100* | 0.233*** | 0.096 | 0.262*** |
| | (0.075) | (0.100) | (0.077) | (0.130) | (0.056) | (0.060) | (0.063) | (0.083) |
| *BIC + BIC*UFB* | | -0.047 | | -0.059 | | 0.005 | | -0.030 |
| | | (0.084) | | (0.099) | | (0.069) | | (0.079) |
| Observations | 20,194 | 20,194 | 20,194 | 20,194 | 17,200 | 17,200 | 17,200 | 17,200 |
| | Third Quartile | | | | Fourth Quartile | | | |
| *BIC* | -0.032 | -0.021 | -0.019 | 0.014 | 0.068 | 0.080 | 0.064 | 0.070 |
| | (0.068) | (0.070) | (0.073) | (0.080) | (0.060) | (0.071) | (0.049) | (0.058) |
| *BIC + BIC*UFB* | | -0.034 | | -0.036 | | 0.039 | | 0.038 |
| | | (0.100) | | (0.119) | | (0.104) | | (0.129) |
| Observations | 12,671 | 12,671 | 12,671 | 12,671 | 15,886 | 15,886 | 15,886 | 15,886 |
| Student FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Principal FE/Teachers | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School Time Trends | No | No | Yes | Yes | No | No | Yes | Yes |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student-level gains in math and English language arts CST scores. Scores are normalized using statewide mean and variance for each grade and year. Each model includes school, grade, and year fixed effects, indicators for school type and year-round status, as well as a set of student level covariates. Reported coefficients are 1) the impact of BIC for schools not previously offering universally free meals, and 2) the sum of coefficients that are the impact for schools previously offering universally free meals. Quartiles of past achievement levels in ELA and math are taken from the distribution of average standardized cst score for years proir to the implementation of BIC.

**Table1.14**: Estimates of Math Gains by Achievement Quartile

CST Math Gains

| | First Quartile | | | | Second Quartile | | | |
|---|---|---|---|---|---|---|---|---|
| *BIC* | 0.091 | 0.245*** | 0.088 | 0.217** | 0.112 | 0.252** | 0.085 | 0.182 |
| | (0.068) | (0.052) | (0.080) | (0.099) | (0.077) | (0.110) | (0.085) | (0.141) |
| *BIC + BIC*UFB* | | -0.012 | | -0.010 | | 0.024 | | 0.020 |
| | | (0.085) | | (0.114) | | (0.096) | | (0.102) |
| Observations | 19,446 | 19,446 | 19,446 | 19,446 | 16,730 | 16,730 | 16,730 | 16,730 |
| | **Third Quartile** | | | | **Fourth Quartile** | | | |
| *BIC* | 0.003 | 0.106 | 0.025 | 0.048 | 0.119* | 0.148*** | 0.101 | 0.095 |
| | (0.121) | (0.198) | (0.136) | (0.234) | (0.070) | (0.051) | (0.083) | (0.094) |
| *BIC + BIC*UFB* | | -0.067 | | 0.012 | | 0.067 | | 0.141 |
| | | (0.131) | | (0.154) | | (0.151) | | (0.150) |
| Observations | 12,951 | 12,951 | 12,951 | 12,951 | 17,507 | 17,507 | 17,507 | 17,507 |
| Student FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Principal FE/Teachers | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School Time Trends | No | No | Yes | Yes | No | No | Yes | Yes |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student-level gains in math and English language arts CST scores. Scores are normalized using statewide mean and variance for each grade and year. Each model includes school, grade, and year fixed effects, indicators for school type and year-round status, as well as a set of student level covariates. Reported coefficients are 1) the impact of BIC for schools not previously offering universally free meals, and 2) the sum of coefficients that are the impact for schools previously offering universally free meals. Quartiles of past achievement levels in ELA and math are taken from the distribution of average standardized cst score for years proir to the implementation of BIC.

# Chapter 2

# Relative Age, Curricula Sorting, Achievement and Long-Run Educational Outcomes: New Evidence from San Diego

## 2.1 Introduction

In recent years, the idea that a student's school starting age can affect outcomes later in life has received much attention in both academic literature and popular media. Additionally, there has been a trend among states over the last few decades to move the entry cutoff date for children to reach the age of 5 earlier in the calendar year. Figure 2.1 shows the large portion of states which have moved cutoff dates to September or earlier between 1975 and 2005. There are important tradeoffs when policymakers increase the school starting age. Moving earlier the cutoff date for entry into kindergarten raises the average age of entering students as well as the average age of students entering the labor market. Consequently, students affected by the earlier cutoff dates forgo an additional year of post-education earnings and human capital accumulation. Given this tradeoff of later school entry, it is important to understand what the size of any school entry age effects are when debating the overall effects of school entry laws.

Conventional wisdom on the matter has held that students who enter school at a relatively older age are likely to enjoy higher levels of understanding and learning from the same classroom experience than their younger counterparts. In general, the empirical evidence suggests relatively older students perform better in school, and this effect diminishes as absolute age increases. The literature has yielded mixed findings, however, on this effect with respect to educational outcomes as adults. Some studies have found evidence that students who enter school relatively young are less likely to take the SAT or enroll in postsecondary education (Bedard and Dhuey, 2006). Others have instead found no evidence for such an effect (Dobkin and Ferreira, 2010). Furthermore, little work has been done exploring mechanisms or channels through which early relative age can affect later outcomes.

There are several ways that being relatively older can affect a student outcomes. Students who are relatively older are developmentally more mature and may therefore learn and retain information more effectively than their younger counterparts. Since students within a classroom are generally tested at the same time, relatively older students enjoy a cognitive maturity advantage at the time of testing as well. This effect will likely diminish as students get older and relative age differences become less significant. If successful learning complements learning later in life, it is possible for early advantages to compound throughout the life-cycle and impact long-run outcomes of students. For example, students who perform well early in their educational careers may receive differential treatment–such as access to more advanced curricula–which leads to the persistence of early age advantages.

On the other hand, relatively older entrants may experience disadvantages as well. Entering school at an older age implies being older in the secondary grades. In the U.S., compulsory schooling laws require students to attend school until a certain state-specified age[1]. Therefore, students entering school at a later age face the decision to drop out or continue schooling earlier in their academic careers, decreasing the average educational attainment of older students (Angrist and Krueger, 1991). Additionally, students who enter school relatively older spend a higher proportion of their early years at home rather than in school. When earlier exposure to classroom learning is relatively more beneficial to skill-building and development than time spent at home, entering school at a later age would be less advantageous in this regard.

Using unusually rich longitudinal data of students in San Diego, I investigate the overall net relationship between school entry age and classroom environments, curricula sorting and long-run educational outcomes. The data are comprised of student records from the time they enter the San Diego Unified School District (SDUSD) until they graduate high school or leave the district. Additionally, I incorporate postsecondary enrollment and degree attainment records for nearly all high school students who were enrolled in SDUSD schools[2]. These longitudinal data allow for analyses of secondary and postsecondary outcomes of students with

[1]California law mandates children attend school between 6 and 18 years of age

[2]NSC provides enrollment data for roughly 92 percent of postsecondary institutions in the US. There are 21 institutions in the state of California with an annual enrollment of at least 1,000 students which are not included in the NSC data. The majority of these institutions are small, career-oriented colleges. A complete list of participating institutions is available at http://www.studentclearinghouse.org/colleges/coreserv/docs/CoreParticipants.xls .

differing school starting ages in greater detail than previous studies. Furthermore, we can observe a host of educational differences during the elementary years that may arise between children who enter school relatively older and those who enter younger. By studying mediating factors in elementary school, we can open up the black box to understand the mechanisms through which school starting age influences both short-term and long-term outcomes.

The results suggest large achievement premiums in mathematics for older entrants–0.2 standard deviations on average–persist into at least grade eleven. Additionally, this group of students scores roughly one third a standard deviation higher on the mathematics portion of the California high school exit exam. They are also about 10 percentage points more likely to enroll in a four-year institution, as well as 15 percentage points more likely to maintain postsecondary enrollment for at least 3 of the first 4 years after high school. Two novel results provide evidence of mechanisms for the persistence of these school starting age effects. First, older entrants persistently score higher on classroom behavior evaluations throughout the elementary years. Second, older entrants tend to place higher in middle and high school math courses than their younger counterparts. To the extent that better classroom behavior evaluations and access to more advanced curricula translate into higher levels of learning, the early impacts from entering school relatively older can generate additional long-run benefits to student outcomes.

## 2.2 Relevant Literature on School Entry Age

### 2.2.1 Student Achievement

Bedard and Dhuey (2006) investigate whether the initial relative age of students entering school has an effect on long-run educational outcomes. They note that a student's relative age is potentially endogenous. Positive selection into an older age (delayed entry) can occur if parents, in hope of improving academic performance, delay their child's enrollment until they are older. Negative selection also occurs, either because younger students tend to perform relatively worse early on and therefore are more likely to be retained a grade and observed as relatively older the next year, or because entry of students who have developed relatively more slowly is intentionally delayed. To deal with this endogeneity, Bedard and Dhuey use assigned relative age – the difference between birth month and the regional school entry cutoff date – to instrument for observed relative age. Using international student-level data, they are able to exploit variation in school entry age across regions, separating seasonal birth effects from the effect of relative age. They find that relatively younger students score 4 to 12 percent lower and 2 to 9 percent lower than the oldest students in grades four and eight, respectively. Their results suggest the benefit of being relatively older is declining with age. To examine the persistence of this effect into adulthood, Bedard & Dhuey also look at the effect of relative age on the probabilities of taking the SAT and enrolling at a post-secondary institution. They find that in the U.S., the relatively youngest students (those born in the month prior to the cutoff-date) are underrepresented in the age distribution of SAT-takers and four-year college enrollees by 7.7 and 11.6 percent, respectively.

Datar (2006) uses data from the ECLS-K 1998 cohort to investigate the relationship between kindergarten entry age and academic achievement in reading and math. Similar to Bedard and Dhuey (2006), she uses an instrumental variables approach that exploits variations in students' ages at the cutoff date and school cutoff dates across regions. The results suggest a one-year delay in entry leads to an increase in math and reading scores between 0.6 and 0.8 standard deviations at kindergarten. Further, students who delayed entry had steeper test score gains than their peers as they moved into first and second grades, particularly for poor and disabled children. These steeper gains are suggestive of an entrance age effect beyond the benefit of greater age at testing. The study does not look at scores beyond the second grade, however, leaving open the question of whether and how these effects persist in later years.

Using U.S. data, Elder and Lubotsky (2009) investigate the effect of school-entry age on academic achievement. While they do estimate a positive effect of age on achievement beginning in Kindergarten, they find these effects taper off as students age and progress through the later elementary years. They also find that students from high-income families show the largest initial gains and most persistent benefits to entering school later. Elder and Lubotsky suggest this is evidence that the relative age effect is the result of greater parental investments in pre-school development on the part of wealthier parents, rather than a learning advantage for older students. Interestingly, they find the youngest students in a classroom are more likely to be diagnosed with disabilities such as attention deficit hyperactivity disorder (ADHD) and attention deficit hyperactivity disorder (ADD).

### 2.2.2 Tracking

Puhani and Weber (2007) look at school starting age effects among German students, where students are tracked as academic or nonacademic at age 10 with subsequent revisions during the secondary years. The stakes are high with respect to overall educational attainment in the German tracking system, as only the highest secondary track known as grammar school provides access to university education. They employ a fuzzy regression discontinuity design, using the discontinuity in school starting ages between birth months around a school entrance cutoff date. Their results show that students who enter relatively young are roughly 13 percent less likely to attend grammar school. They also find that in most secondary schools, this trend is reversed after students are re-tracked around 11 or 12 years after school entry, suggesting that the differences attributed to school entrance age diminish as students become older.

Dhuey and Lipscomb (2010) further investigate the finding of Elder and Lubotsky that younger students are disproportionately diagnosed with learning disabilities. They find that relatively older students are indeed less likely to receive special education services. Their conclusion is that initial maturity disadvantages are either perceived as ability deficiencies or are treated with special education in an attempt to close the initial performance gaps between young and older students. Although they do not look at curricula sorting outside of special education, their findings hold potentially important implications for the effect of relative age on eligibility for advanced curricula. A significant and positive effect of age on sorting into accelerated programs could merely be a mechanical result of disproportionately more students who are young being sorted into special education and automatically deemed ineligible for such programs.

### 2.2.3 Outcomes as Adults

Black, Devereux and Salvanes (2011) use population data from Norway to investigate the effect of school starting age on student outcomes as well. One of the main contributions of their paper is that they are able to use IQ test scores, for which there is variation in the age at which individuals are tested, to separate the effect of school starting age on IQ measures from the effect of being older at the time of testing. Black, Devereux and Salvanes also investigate the effect of school-entry age on overall years of education as adults. They find a strong negative correlation between school-entry age and years of educational attainment. Once they instrument for school-entry age using expected school-entry age, however, they find no statistically significant effect on the educational attainment of men or women.

Dobkin and Ferreira (2010) look at the effect of relative age on both the likelihood a student completes the 9th-12th grades and the likelihood of postsecondary enrollment. They note their estimates are of the net long-run effect of relative age, since older students tend to perform better but younger students tend to attain more education due to compulsory schooling laws in the U.S. Using a combination of IPUMS and the 2000 Decennial Census Long Form data, they use a regression discontinuity (RD) approach to estimate the local average treatment effect (LATE) of being born just after the school-entry cutoff dates in California and Texas. The RD design is useful here. Both students' relative age and unobservable determinants of academic success can vary with birth month, making it difficult to estimate an effect of relative age that is free from seasonal birth effects. By exploiting the discontinuity of relative age around the cutoff date, Dobkin and Ferreira are essentially comparing students born around the same time, where some start school a year later. They find a significantly negative LATE of being born after the cutoff date on the likelihood of completing the 9th - 12th grades. Moreover, this effect is increasing with high school grade level, suggesting compulsory schooling laws are preventing younger students from leaving school earlier. When looking at the likelihood of postsecondary enrollment, however, they do not find any significant effect of being born after the cutoff date which is evidence against a net long-run effect of relative age on postsecondary educational attainment. One possible explanation for this result is that two opposing effects negate any overall net effect. Students whose dropout decisions are affected by school starting age, due to compulsory schooling laws, exhibit a negative association between entry age and years of educational attainment. On the other hand, students with strong preferences for postsecondary education might exhibit a positive association between entry age and overall educational attainment.

### 2.2.4 Contributions of this Paper

To this author's knowledge, Dobkin and Ferreira (2010) provide the only recent study of the effects of school entry laws using students' exact day of birth. However, their data has the limitation that actual school entry date and the timing of grade progressions are unobserved. Therefore, the fuzzy discontinuity in entry age for those born near the cutoff is treated as a sharp discontinuity in the RD design, biasing estimates toward zero. Moreover, later changes to relative age–such as disproportionate grade retentions among the youngest students–are unobserved and potentially contaminate estimates of the causal effects of school starting age. Other authors have observed actual school entry and grade progressions, but have only been able to observe students' birth month (e.g. Bedard and Dhuey, 2006), resulting in less precise estimates than if day of birth was used.

This paper aims to contribute to the existing body of school starting age literature in three ways. First, it introduces another set of estimates of direct school starting age effects in which both date of birth and actual entry dates of students into each grade level are observed. While the existing evidence on the persistence of school entry age effects, particularly with respect to long term outcomes–is mixed, I find relatively older entrants exhibit higher levels of achievement in math than their younger counterparts well through the secondary years. Moreover, these students are more likely to enroll in a four-year postsecondary institution later in life. Second, I investigate possible mechanisms within schools by which these documented school starting age effects can persist beyond the early years of schooling. Older entrants score much higher marks in classroom behavioral evaluations, display better performance on non-verbal intelligence tests after controlling for age at test, and are more likely to be above-track in secondary math course progression. Third, through the use of rich secondary and postsecondary data, I document that entering school older increases the length of student's education as adults, particularly through postsecondary enrollment and persistence.

## 2.3 Data and Sample

The primary data come from administrative records on students and teachers maintained at San Diego Unified School District (SDUSD). San Diego has a large and diverse student body. SDUSD is the second largest public school district in California and has ranked between the eighth largest and twentieth largest in the U.S. over the years encompassed by our sample. 75 percent of the student body is non-white, over 30 percent are English language learners (EL) and roughly 60 percent qualify for the federal lunch-assistance program[3].

The data consist of longitudinally linked data on each student in the district, providing details year by year of students' school locations and classrooms, atten-

[3]See Betts, Zau and Rice (2003) for evidence that SDUSD is broadly representative of the student population of California.

dance, grades, behavioral evaluations, test scores, grade promotion and whether the student was retained or skipped a grade level. In addition, the database also includes information on the specific teachers of each student, and detailed qualifications of these teachers. Combining student records from SDUSD with postsecondary data from the National Student Clearinghouse (NSC), the panel of data follows students from Kindergarten through the Baccalaureate level. Among outcomes at the post-secondary level are indicators for any postsecondary enrollment in the first year after graduating from high school, the type of postsecondary institution attended in this first year (community college or four-year college), and indicators for obtaining a two-year degree within four years after graduation or a four-year degree within four or five years after graduation.

The sample used in the paper includes students with observed Kindergarten entries from Fall 1992 through Fall 2010. Most of the outcomes studied are observed beginning in 1996[4] and are observed through Spring of 2011. Because few of these cohorts reach their fourth year of postsecondary education by 2010, the analyses looking at postsecondary outcomes are supplemented with earlier cohorts where exact school entry is not observed. The supplemental estimates are reduced form, using student date of birth. Appendix Table B.1 presents total cohort sizes for each entry year observed from Kindergarten through grade 12. Table 2.1 provides summary statistics by grade level for outcomes used at the elementary and secondary levels. Summary statistics for other key variables are presented within their respective tables of results.

## 2.4 Empirical Methods and Identification

### 2.4.1 Endogeneity of School Starting Age

In estimating the effects of school starting age, a naive approach might revolve around an estimation equation of the form:

[4]The California Standards Test as well as postsecondary outcomes are the exceptions, with observations beginning in the Spring and Fall of 2002, respectively.

$$Y_i = \alpha + \beta X_i + \gamma Age_i + \varepsilon_i, \tag{2.1}$$

where $Age_i$ is the age at school entry for student $i$. One identification problem with such an approach is students may selectively start earlier or later than the cutoff date prescribes at the direction of parents or schools for reasons that are unobserved. Much of the previous literature has tried to account for this endogeneity by instrumenting for school starting age with expected or compliant school starting age. However, this method requires sufficient variation in cutoff dates to disentangle season of birth from school starting age. To the extent that season or month of birth may be correlated with other unobserved determinants of outcomes, it is a potentially confounding factor (Bound and Jaeger, 2000; Buckles and Hungerman, 2008). For these reasons, the paper focuses on a regression discontinuity (RD) design which exploits the discontinuous jump in school entry age among student with birth dates on either side of the California school entry age cutoff: December 2, the year the student turns five years of age.

A student born just after the cutoff date enters school approximately one year older than a student born just before, provided both comply with school-entry laws. Within a small interval around the cutoff date, it is unlikely parents are able to manipulate the prescribed school-entry age by accurately timing a child's birth date to a specific side of the cutoff. Moreover, students born within this interval around the cutoff are born in the same part of the calendar year, eliminating seasonal confounding factors.

Admittedly, parents may manipulate the entry age determined by the cutoff either by postponing entry or finding ways to enroll their child slightly early, such as private kindergarten enrollment. However the cutoff does provide a large discontinuity in the probability of entering older, with this probability discretely increasing for students born just after the cutoff in middle and high school, relative to those born just before. Figure 2.2 presents the proportion of students entering kindergarten in the academic year of their sixth birthday, plotted against the number of days born before or after the December 2 cutoff. Students born after the cutoff that manage to enter kindergarten in the year they turned five are exceedingly rare, a few percentage points. As for those who were born before the cutoff, the compliance rate is about 80 percent among those born in the weeks before December 2, rising to almost 100 percent among those who turned five months before the cutoff date. These features of the school entry age requirement suggest the distance in days of birth date from the cutoff date is an ideal forcing variable and that strong compliance with the state guidelines make a fuzzy RD design highly applicable.

### 2.4.2 Regression Discontinuity Design

Following Imbens and Lemieux (2008) and Lee and Lemieux (2010), the impacts of entering school relatively older are estimated using local linear regressions (LLR) of the form

$$Y_{ji} = \alpha_j + f_j(distance_i) + \phi_j' X_i + u_{ji} \tag{2.2}$$

$$D_{ji} = \gamma_j + g_j(distance_i) + \theta_j' X_i + u_{ji} \tag{2.3}$$

$$j = \begin{cases} l, & distance_i \in [-h, 0) \\ r, & distance_i \in [0, h] \end{cases}.$$

Regressions are fit separately within a bandwidth $h$ on either side of the cutoff date. Functions $f_j(.)$ and $g_j(.)$ are estimated using the form $f_j(distance_i) = \beta_j distance_i$. Checks that the significant results of the paper do not differ widely when using higher degree polynomials for $f_j(.)$ and $g_j(.)$ are provided in Appendix Tables B.6 through B.8.

In the above equations (2.2) and (2.3), $Y_{ij}$ is the outcome of interest for student $i$ born within $h$ days on the $j$ side of the cutoff date. $D_{ji}$ is an indicator that student $i$ on side $j$ delayed entry until the academic year she turns six years of age. The variable $distance_i = (DOB_i - c)$ is the number of days student $i$ is born after the cutoff, centered at zero. $X_i$ is a set of student level covariates–each predetermined before school entry–progressively included in the estimating equations. Equations (2.2) and (2.3) describe the discontinuities of the outcome $Y_i$ and delayed entry treatment $D_i$, respectively, around the cutoff date. Given valid conditions for the fuzzy RD design, the resulting estimand is a form of the Wald estimator:

$$\begin{aligned}\delta = & \frac{lim_{x\downarrow 0}E\left[Y \mid distance = x\right] - lim_{x\uparrow 0}E\left[Y \mid distance = x\right]}{lim_{x\downarrow 0}E\left[D \mid distance = x\right] - lim_{x\uparrow 0}E\left[D \mid distance = x\right]} \\ = & \frac{\alpha_r - \alpha_l}{\gamma_r - \gamma_l}\end{aligned} \tag{2.4}$$

This is the difference in average outcomes for students born on just either side of the cutoff, normalized by the discrete change in the likelihood of delayed entry as birth date moves across the cutoff. This provides a local average treatment effect (LATE) for students born near December 2. These estimates are numerically equivalent to 2SLS estimates of the form

$$Y_i = \alpha_l + f_l(dist_i) \cdot 1\{dist_i < 0\} + f_r(dist_i) \cdot 1\{dist_i \geq 0\} + \delta D_i + u_i. \tag{2.5}$$

An indicator for "born after cutoff", $Z_i = 1\{dist_i \geq 0\}$, is used to instrument for observed delayed entry $D_i$.

To choose estimation bandwidths, the Imbens and Kalyanaraman (2009) bandwidth algorithm (IK), which minimizes MSE in the sharp RD case, is used to calculate $h_{IK}$ for each outcome. For the fuzzy RD design, there is not clear case for an optimal bandwidth. A larger bandwidth increases precision as the within-bin sample grows, but may increase bias when the curvature of the specified functional form is a poor fit close to the discontinuity point. To account for the increase in standard errors that result from dividing the reduced form RD estimate by the first stage treatment estimate, the results reported below use the bandwidth $h = 2h_{IK}$. Evidence that the results are not sensitive to changes in bandwidths near $h$ are provided in Appendix Tables B.6 through B.8. Computationally, these calculations are done using techniques discussed in Nichols (2009). Bootstrapped standard errors are reported for the results that follow and are clustered at the student level for any panel analyses that involve more than one observation per individual.

### 2.4.3 RD Design Validity Tests

There are two main threats to the valid use of an RD design. The first is the possibility of manipulation of the forcing variable. In the framework of this paper, this would mean individuals are able to manipulate student birth dates in order to be on one side of the cutoff. This could occur purposefully, perhaps through misreporting of birth dates or parents' strategic timing of birth. Such behavior is likely to be correlated with unobservables that may also affect outcomes, nullifying the quasi-experimental nature of the RD design. The second threat to validity is the existence of irregularities in the distributions of other variables around the cutoff. Again, this can create issues with the validity of RD estimates if such variables are also correlated with the outcomes of interest.

Two specification checks are done to investigate whether these issues are likely. Although it is not possible to test for discontinuities along unobservables, I argue that if no such problems are found across the set of observable characteristics, they are unlikely to exist along unobservables determined before school entry.

Figure 2.3 presents the density of birth dates around the cutoff, where there do not appear to be any local discontinuities or relatively large variations in the distribution. Figure 2.4 plots the distributions of observable student characteristics which are determined before school entry around the cutoff. There do not appear to be any discontinuous jumps around the entry age cutoff. Table 2.2 presents RD estimates testing for sharp differences in these covariates around the cutoff. Estimates using the IK optimal bandwidth, as well as half and twice the IK bandwidth are reported. The results suggest there are no discontinuities that would invalidate the RD design. Additionally, one must be sure there do not exist other discontinuities outside of the one for use in the RD framework. Taken together, Figures 2.2, 2.3 and 2.4 do not show any evidence of such "outside discontinuities," strongly suggesting the RD design is valid for the purposes of this paper.

Finally, differential attrition on either side of the cutoff date is tested for. Given that this paper follows students from Kindergarten entry to beyond high school, there exists some attrition of students from the sample during such a large time interval. If students born on one side of the cutoff date leave the school district at a disproportionate rate, it is suggestive of selection into (or out of) the sample along potentially unobserved determinants of the outcomes of interest. Appendix Table B.2 uses the main specifications for the RD analyses across multiple bandwidths to test for any significant differences in attrition around the discontinuity in entry age. This test uses an indicator that the student observed at Kindergarten entry is not observed in the secondary and postsecondary outcomes sample. No significant difference in attrition rate is found.

## 2.5 Results

### 2.5.1 Student Achievement

The Math and English-language arts (ELA) portions of the California Standards Test (CST) are administered annually in California public schools to students in grades 2 through 11. The data contain CST scores for all San Diego students in grades 2 through 11 since the time the test was first administered–Spring 2002–through Spring 2011. Scores are standardized using statewide means and standard deviations for each grade level and year. These outcomes provide a way of testing whether entering school older has a significant impact on student achievement levels, relative to younger students. The longitudinal nature of the data also provide a glimpse into the persistence of any such differences between old and young entrants.

Table 2.3 presents RD estimates of the effect of delayed school entry, as induced by the cutoff date, on CST ELA and Math Scores for each grade level. Figures 2.5 and 2.6 plot these estimates by grade level. To deal with the possibility of multiple attempts by some students, only the first observed attempt of each grade level and test are used for each student. Because observed delayed Kindergarten entry–defined as entering the academic year the student turns six years of age–is used as the treatment, the estimation sample sizes for each outcome naturally decrease each grade level as some students move out of the district. Unless stated otherwise, all regressions include indicators for gender, ethnicity, English learner status and parental education level. Because CST Math tests vary by the student's level of math course progression in Grades 8 through 11, interactions for grade level by test version are included in the corresponding regressions.

Early relative achievement differences for older versus young entrants are remarkably large. For math scores, the oldest entrants score nearly 0.53 standard deviations higher than the youngest. These differences slowly decline as students matriculate through the elementary grades, where the difference is still a large 23 percent of a standard deviation in grade 6. Grades 7 and 8 show increasingly larger differences of 0.26 and 0.34 standard deviations, respectively. The size of the grade 8 estimate is particularly interesting, as this is the first point in students' educational careers where tracking into math courses occurs with substantially different curricula. This also signifies the first point where students face different CST math tests, corresponding to their current math curriculum. Moving to high school grades 9 through 11, the results begin to vary more. RD estimates are a weakly significant 0.17 standard deviations in Grade 9, larger and statistically significant at 0.31 standard deviations in Grade 10, and an insignificant 0.16 standard deviations in Grade 11.

The overall pattern of estimates as students progress through public school grades is in line with that found in the literature. Biologically older children exhibit much higher achievement levels than their youngest peers early on, but as Figures 2.5 and 2.6 illustrate, the size of this difference decreases over time. In fact, the fourth grade Math estimate here of 0.31 standard deviations is remarkably similar to the IV estimate of Bedard and Dhuey (2006) that finds a roughly 0.30 standard deviation difference for students born 11 months apart. Additionally, the IV estimates of Datar (2006) find roughly a 0.6 to 0.8 standard deviation difference in Kindergarten achievement between the oldest and youngest, which lines up reasonably well with the 0.5 standard deviation estimates found here in grade 2.

What does differ here from the previous literature is the extent to which the effects on math achievement persist. Estimates range between 0.17 and 0.34 standard deviation from middle school through grade 10, and of which are extremely large differences by any means. This apparent persistence of the school entry age effect is at odds with the finding of Elder and Lubotsky (2009), who find differences in math scores converge quickly and are only a few percentage points by grade eight. The variability found in grades 9 through 11 are somewhat concerning, but may potentially be explained by two things. First, there is a relatively wider range of math curricula and test versions in high school. Moreover, older entrants are also more likely to enroll in advanced-level math courses[5]. Although the particular math tests taken have been controlled for in the reported specifications, it may not completely eliminate the competing influences of older entrants facing more difficult material, but performing better than their younger peers conditional on difficulty level. Second, while the analyses do not include observations of repeat attempts for any grade, they do not exclude the possibility of comparing two first-time tenth grade students who differ along retention in, say, second grade. To investigate the extent to which these features of the data may influence estimates, students retained prior to high school are later excluded, with these results reported in Appendix Table B.3. Reduced form estimates including students who entered the district after Kindergarten are also reported in Table B.3. The pattern of results described above are robust to these changes in the estimation sample.

### 2.5.2 Retention

Given the evidence that the oldest entrants enjoy higher levels of achievement, and the fact that poor academic performance often leads to grade retentions, we might expect older entrants to be retained less often than their younger counterparts. Table 2.4 and Appendix Figure B.2 present RD estimates on whether or not a student was retained at all during each grade level. Columns (1) and (2) show that older entrants are 10 and 6 percentage points less likely to be retained

[5]This is outcome is discussed in Section 2.5.5

during Kindergarten and first grade, respectively. After first grade, there is no statistical difference in the incidence of retention among these groups, with point estimates dropping below one percent. This suggests the starkest differences stemming from school entry laws that are both perceived and large enough to warrant grade retention are essentially mitigated by second grade. Elder and Lubotsky (2009) also looked at school entry age and grade retentions. They find older students are roughly 15 percentage points less likely to be retained at all from Kindergarten through grade 8, and 13 percentage points less likely to be retained in first and second grade. The results of this section support their finding, and document that such retentions happen mostly in Kindergarten and to a lesser extent, in first grade.

### 2.5.3 Behavior

Elementary students in San Diego are evaluated by teachers on four measures of behavior within the classroom. These marks are numbered scores of the degree to which the student completes assignments, is attentive in class, disrupts classroom activities, and adapts to social interactions with peers. To the extent that these behavioral qualities may be associated with age and can impact the efficacy of classroom instruction, I test for sharp differences in the levels of classroom behavior evaluations around the cutoff between the youngest and oldest students. Because SDUSD changed the names and definitions of these behavioral measures in 2008, the original four measures are used to study the effects for students present in the secondary and postsecondary samples. Appendix Table B.5 reports results for newer entrants using the latest measures, which exhibit very similar patterns despite the somewhat subjective differences of these new measures.

Columns 2 through 4 of Table 2.5 show that, on average, older entrants score between 32 and 40 percent of a standard deviation higher on elementary behavior evaluations than the youngest entrants. Figure 2.7 plots each of the four standardized scores by day of birth, illustrating the stark differences around the cutoff date. Table 2.6 presents results for each behavior measure and grade level. The differences in behavior scores are remarkably persistent, with estimates for fifth grade not being statistically different than those for first grade. By fifth grade, older entrants still maintain scores roughly a half standard deviation higher for the "begins promptly" category. The scores for following directions, social behavior in the classroom, and self discipline are nearly as large, between 0.36 and 0.4 standard deviations in fifth grade.

These results show that relative maturity differences within the classroom drive differences in behavior evaluations. Whether these effects result from learned behavior early on, when a one year difference is relatively larger, or are indicative of differences in the perceptions of administrators is something that cannot be tested under the framework of this paper. This does however provide suggestive evidence that student behaviors, particularly behaviors thought to be associated with learning efficacy, are one mechanism by which differences in entry age might affect other student outcomes throughout the educational process.

### 2.5.4 Classroom Characteristics

It is possible that the apparent differences between young and old entrants affects student sorting in terms of classroom environments. If this is the case, the characteristics of a classroom that impact student outcomes–such as classroom size, peer quality, or even peer age–may depend on school entry age. Such sorting would suggest additional mechanisms through which the documented entrance age effects occur.

Linking students to teachers, classrooms and classmates during the elementary years, I estimate the effects of delayed entry on various classroom characteristics. This section focuses on Kindergarten through grade 5, since administrative records of the period structure of classrooms after elementary school do not allow for a tractable grouping of classrooms. Table 2.7 shows there is no significant effect of delayed entry on the average age of students' peers, classroom size or the years of teaching experience held by the student's teacher.

Table 2.8 presents estimates for the effect of delayed entry on peer quality, as measured by achievement levels on the CST, which is first administered in second grade. This outcome is calculated as the average score levels across all students, excluding student *i*, for the student's corresponding grade level and year. There does not appear to be any ability grouping of students resulting from entrance age or the higher achievement levels associated with delayed entry. The fact that observable classroom characteristics do not differ with entry age suggests these are not channels through which the school entry age effects are driven or mitigated.

### 2.5.5 Differentiated Curricula

#### Intelligence Testing and Gifted Programs

The Gifted and Talented Education (GATE) program is the most prevalent early method of sorting young students into accelerated curricula by ability within U.S. public schools. Typically, an intelligence test is administered to students during the primary grades in order to identify students with above average intelligence as eligible for GATE classes. In San Diego elementary schools, GATE consists of two programs; the Cluster program and the Seminar program. Virtually all students take the Raven's Progressive Matrices (RPM) test while in the second grade unless parents opt out of testing. The RPM is an untimed, non-reading test administered by a child psychologist. By testing visual and non-reading, non-verbal skills, it attempts to avoid ethnic and language biases while measuring intelligence. Students who score above the 98th percentile of the RPM are eligible to participate in the Cluster program. Additionally, students who score in the 95-97th percentile and also satisfy certain environmental, economic, language, emotional or health criteria considered to be academic disadvantages are eligible for the Cluster program as well[6].

Of the outcomes in this paper, the GATE identification test makes for a

[6] Because students who enter the district after the second grade are eligible for testing during grades 3 through 7, this section focuses on students' first RPM attempt in second grade.

uniquely different outcome in evaluating the effect of relative age in the classroom. To my knowledge, no other studies of relative or school starting age use student outcomes that do not depend on reading or curriculum-specific skills. Since the RPM is designed to measure more basic cognitive skills, such as abstraction and visual pattern recognition, we feel this outcome may provide valuable insight into any intelligence or cognitive development advantages that school-entry timing provides.

Because school psychologists administer the RPM to one school at a time, there is variation in the calendar dates of testing and hence age at testing, given age at entry. As Columns (5) and (6) of Table 2.9 show, however, this variation may not be sufficient to distinguish entry age effects from age at test effects. Delayed entry induced by birth to the right of the cutoff results in an average 306 additional days of age at the time of testing. Controlling for age at testing can provide insight into gifted education policy, however. In order to avoid age at test bias in GATE admissions, RPM scores are rescaled according to the student's at the time of testing. Figure 2.8 displays final RPM scores by day of birth, with the corresponding estimates presented in columns (1) and (2) of Table 2.9. Clearly, the youngest students are scoring at least as high, if not higher, than the oldest of their cohorts. Once age at test in controlled for, effectively eliminating the rescaling of RPM scores, column (3) shows that older students do in fact perform better at these tests of cognitive skills. Delayed entry corresponds to scores roughly 5.5 points higher than the youngest students on the RPM.

The lower halves of Figure 2.8 and Table 2.9 show that older entrants are not more likely to enroll in GATE. It is only after controlling for age at test that significant effects of school entry age appear: 17 percentage point increase in the likelihood of GATE Cluster enrollment. The collinearity of school entry age and age at test, however, prevent any causal inference from these estimates. Evidently, the rescaling of RPM scores does somewhat mitigate the apparent maturity advantage older students enjoy with respect to test scores. This practice does prevent older students from participating in GATE at higher rates than younger students, as we might expect given the large entry age advantages we have seen thus far for students in second grade.

Whether this is a policy failing or success of gifted education tests amounts to a question of equity versus efficiency. On the one hand, if space is limited in GATE classrooms and performance differences across entry age do diminish as students become older, the rescaling of scores can prevent an early crowding-out of young students before they can benefit from a GATE curriculum. On the other hand, if school entry age effects do persist, or even compound, a proportion of seats otherwise given to older students ready to benefit from advanced curricula may be allocated to students who benefit less from GATE solely because of RPM scaling. What can be said from the results of this section is that while older entrants do perform better at these types of cognitive skill tests, they are not more likely to participate in GATE. Thus, it does not seem likely that sorting into classrooms with advanced curricula is a mechanism by which school entry age effects can persist at the elementary level.

**Mathematics Course Progression**

Students in San Diego encounter differentiated mathematics curricula in middle school, with greater differentiation occurring in high school. Using course enrollment data, I test whether school entry age affects the levels of math courses students are placed in during the middle and high school grades. Each student-year observation is labeled with four levels of math progression as defined by California standards for each grade level and the SDUSD catalog of courses: behind-track, on-track, ahead-of-track and advanced level. “On-track” is used as the base category in the analyses that follow. With the exclusion of “behind-track,” the resulting dummy variables indicate whether a student is currently at or above the corresponding math course level.

Column (2) of Table 2.10 shows that differences in math course placement first occur in grade 8. This is not surprising, as there is little differentiation in grade 7 aside from support courses. Students induced to delay entry by the cutoff date are about 12 percentage points more likely to be ahead of the prescribed math course level. For eighth graders, this typically means the equivalent of high school algebra. Column (3) suggests older entrants are 7 percentage points more likely to be ahead in grade 9, though the estimates here are not precise enough for statistical significance. In grades 10 and 11, however, a significant difference of 10 and 14 percentage points for being ahead of normal progression is found. That no effect is found for grade 12 can be explained by the fact that California requires only three math courses for graduation: algebra, geometry and intermediate algebra. Many students finish these requirements before grade 12, leaving grade 12 students to self-select into mathematics courses as electives.

As shown in Section 5.1, older entrants perform considerably better in mathematics into the high school years. Given that high school students enrolled in higher level math courses also take correspondingly higher level math test on the CST, it is surprising that the effects of CST math achievement persist into the years of differentiated curricula. Although the effect size for grade 11 CST math scores diminishes to an insignificant 0.16 standard deviations, together with the increased likelihood of facing more difficult tests, this suggests achievement differences in math may not completely fade at all.

## 2.6 Outcomes as Adults

This section investigates how relative age affects long-run educational outcomes. Given the mixture of findings on the effect of relative age on educational attainment in the literature, these results will provide an interesting additional piece of evidence. As Dobkin and Ferreira (2010) note, these RD estimates are of the net effect of entering school older since older students perform better academically, but exhibit higher drop out rates when compulsory schooling laws are no longer binding in higher grade levels. Indeed, it is likely that being older than classmates negatively influences education attainment for students who would otherwise not have earned beyond a high school diploma since they outgrow compulsory school laws earlier in their educational careers. On the other hand, it is possible that relative age advantages have an upward influence on educational attainment for students with postsecondary aspirations.

Using the estimation methods described in Section 5.1, I estimate the effects of relative age on high school exit exam performance, the likelihood of high school completion, levels of postsecondary enrollment (2-year or 4-year) and levels of postsecondary degree attainment.

### 2.6.1 Secondary Attainment

In California, all students must pass a high school exit exam in order to earn a diploma. This test is comprised of two components: ELA and math. Students have several opportunities to pass the exam, beginning in grade 10. This section examines the effects of school entry age on California High School Exit Exam (CAHSEE) performance. Because the difficulty and composition of CAHSEE underwent changes in 2004, the estimation sample is restricted to students in grade 10 taking the exam in 2004 or later. This enables comparisons among students taking the same exam during the first year they attempt to pass CAHSEE.

Column 1 of Table 2.11 shows that among first-time tenth grade students, attempting CAHSEE for the first time, students who have entered relatively older score roughly 11 points higher on the math portion. There is no significant difference on the ELA portion for these students. The results for math represent more than a quarter of a standard deviation difference in scores. Table 2.12 shows that older students may be slightly more likely to graduate high school as well, with weakly significant results of between 10 and 15 percent increased likelihood of graduation. Together, these results suggest that of the students induced to delay entry by school entry age laws, the additional benefit of older entry may help a portion of these students graduate high school at the margin.

Particularly interesting is the absence, or perhaps masking, of compulsory schooling law effects. As Angrist and Kreuger (1991) show, students who enter school older and consequently age out of compulsory schooling laws earlier, tend to have lower rates of secondary completion. In this post-2000 San Diego data, however, there does not appear to be such an effect. Or rather, if this negative effect on attainment is present, it may be dominated by the competing positive effect of delayed entry. One explanation is that compulsory schooling laws do not hold as much importance as they used to in the educational attainment decision. Students choosing to forgo additional years of schooling may effectively no longer be constrained by compulsory schooling laws in California, either because 1) they ignore the law and drop out before 18 years of age, or 2) they tend to end their educational attainment after the secondary years since the average marginal time to graduation for any student turning 18 years of age in high school is a matter of months.

### 2.6.2 Postsecondary Outcomes

The previous sections of this paper have utilized observed Kindergarten entry for first stage estimates, allowing the RD estimates to account for the fact that some students born to the left of the cutoff voluntarily delay school entry. Unfortunately, the migration of students in and out of the district over time result in dwindling sample sizes that contain observed entry in the secondary years. This results in roughly half of students observed in the postsecondary data having Kindergarten entry records. For this reason, the postsecondary results report the usual set of estimates, but also report reduced form estimates using all observed students. The increased power using this larger sample allows for a quick check as to whether small sample size may be to blame when no significant effects are found.

Columns (1) and (2) of Table 2.13 present results for whether students enroll in any postsecondary institution during the first year after scheduled, or on-time graduation. The next two columns model enrollment in a four-year institution during the same time frame. Older entrants are about 12 percentage points more likely to enroll in a four-year institution after high school than the youngest entrants. Although the estimates in column (3) are weakly significant, the reduced form estimate utilizing the entire sample in column (4) remains roughly the same, but with more precision. Given the rather stable first stage estimates throughout the paper, the RD estimate of column (3) is likely reliable. Figure 2.10 plots these data against day of birth. It appears older entrants born to the right of the cutoff have a slight bump in enrollment rates, but the majority of the estimated effect is driven by the youngest students being dramatically less likely to enroll than their classmates.

Older entrants appear to have higher levels of postsecondary persistence, as well. Table 2.14 presents estimates for the number of years in the first four after high school for which students were enrolled in a postsecondary institution. Delayed entrants appear to be 15 to 20 percentage points more likely to be enrolled during all four of the initial years after high school. Despite this increased level of persistence, there do not appear to be significant effects on attainment. Table 2.15 contains estimates using any postsecondary degree or four-year degree as outcomes. These results look at degree conferral four full academic years after high school. The point estimates are roughly 4 and 6 percentage points for the reduce form and RD estimates, though the standard errors are just as large as the point estimates. Admittedly, the average completion time of a four-year degree is between four and five years. However, these results of no effect do not change when allowing for five years or more after high school.

## 2.7 Conclusion

The results of this paper point to a number of large impacts as a result of delayed school entry induced by minimum entry age guidelines. First, relative achievement differences in math between the oldest and youngest entrants are large and persist well into the secondary years. Point estimates suggest an average difference of roughly 0.2 standard deviation in high school. Though the persistence of entry age effects on achievement have been documented before, the previous literature has not been able to track grade specific specific achievement differences through the secondary years. Moreover, previous studies show a faster tapering of the effects on achievement when instrumenting for entry age with expected entry age. This paper also provides evidence that relative maturity differences do in fact provide cognitive advantages. Section 5.5.1 shows older entrants perform much better on intelligence exams testing non-verbal, cognitive tasks more than two years after Kindergarten entry.

Second, the results here suggest relative entry age has a large impact on postsecondary enrollment and persistence, with older entrants being at least 10 percentage points more likely to immediately enroll in a four-year institution after high school and persist in postsecondary enrollments through the following four years. Interestingly, there is no evidence of a strong interaction effect of delayed school entry and compulsory schooling laws on educational attainment. The fact that the data used by this paper are significantly more contemporary than studies showing a negative impact of older entry on attainment suggest that contemporary schooling laws in California may no longer be relevant in the attainment decision for students nearing the end of their secondary years.

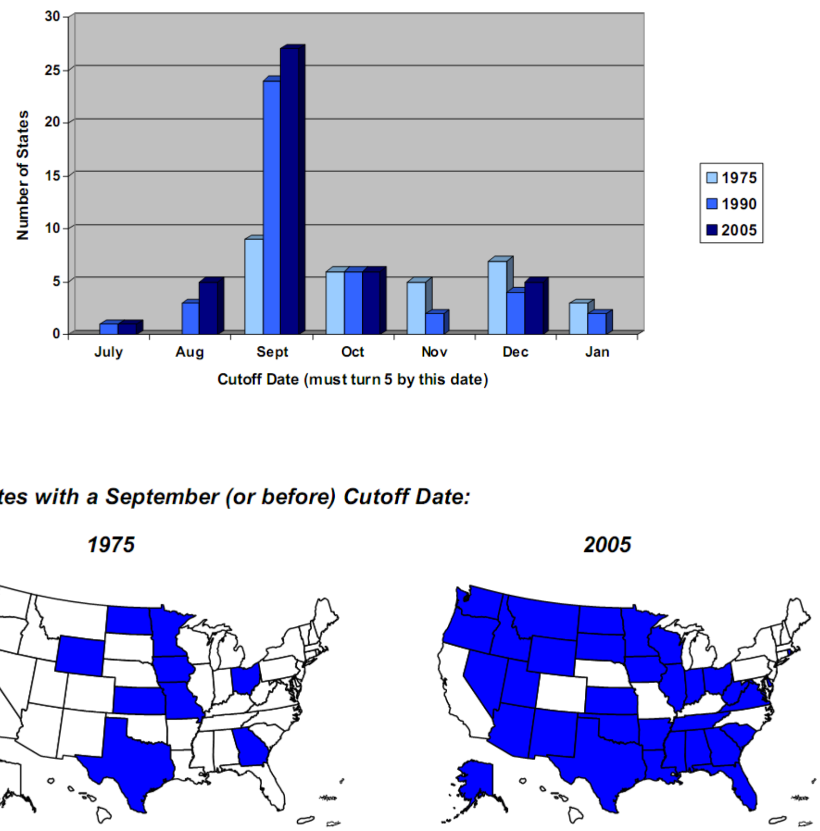
Third, this paper is the first to document other impacts that may function as mechanisms driving the persistence of entry age effects. There is no evidence of entry age impacts on classroom sorting–in terms of age, peer quality or classroom size–but older entrants exhibit better behavior in the elementary classroom and are likely to be placed in higher level math courses in middle and high school. Differences between the oldest and youngest entrants in classroom behavior are remarkably large and do not appear to diminish by the end of the elementary years. Although these differences appear to be monotonic with age, Figure 2.7 suggests there are non-linear decreases in behavior evaluations for the youngest students born within 30 days before the cutoff date. Students who enter older because of the cutoff are also 12 and 14 percentage points more likely to be above normal progression with respect to their math course placement levels at the end of middle and high school, respectively, than the youngest entrants. These results suggest the benefits to entering relatively older affect the education process through access to higher levels of curricula and perhaps more effective learning as measured by classroom behavior.

With the exception of the intelligence tests administered in second grade, this paper has not distinguished entry age effects from age-at-testing effects. Nor have entry age effects been distinguished from effects of relative age differences, due to lack of cutoff date variation in the data. From a policy perspective, however, these distinctions hold little importance. In practice, students are usually evaluated with their class unit simultaneously, which holds fixed time in school when comparing student achievement but can depend on ages at entry and testing. Moreover, since nearly every schooling system has some age threshold for first time enrollment, it is not feasible to altogether eliminate relative age differences within cohorts. Whether relative or absolute initial maturity advantages matter more in the human capital accumulation process, there is almost always a combined effect at the time students are evaluated for ability-sorting programs and curricula. With that said, the combined results of this paper highlight the usefulness of considering both relative and absolute student age in education policies surrounding student evaluations, placements and overall attainment levels.

More research is needed on the impacts of both relative and absolute student age on outcomes. For policy considerations, it would be particularly useful to know impacts and long-term effectiveness of age-based interventions for the relatively youngest students of cohorts. Additionally, research on policies changing only absolute entry age can shed light on whether shifts in cutoff dates are worth the opportunity cost of delayed labor market entry and how the impacts might vary by the direction of such shifts. For example, California recently made its first entry cutoff change in sixty years, moving the cutoff forward from December 2 to November 2 in 2012. The state will progressively be moving this date forward one month annually until 2014, reaching a final September cutoff. This will provide an opportunity to study the impacts of increasing absolute entry age while holding relative age differences around the cutoff date constant.

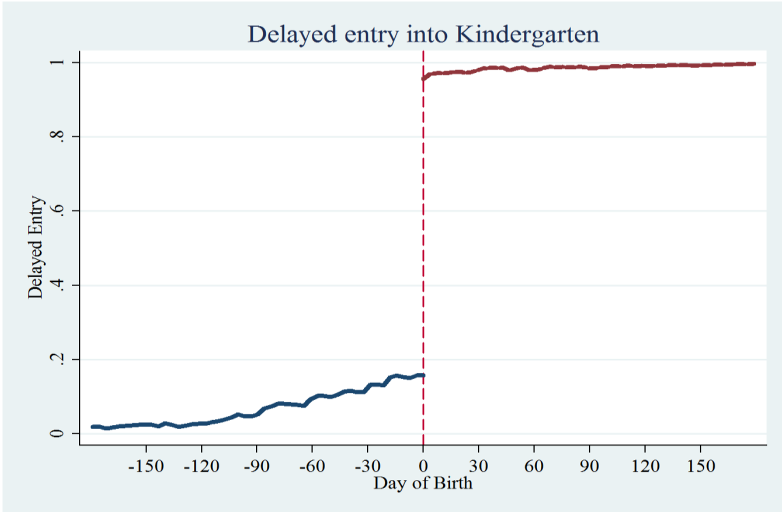
This chapter will be prepared for future submission for publication.

## 2.8 Chapter 2 Figures and Tables



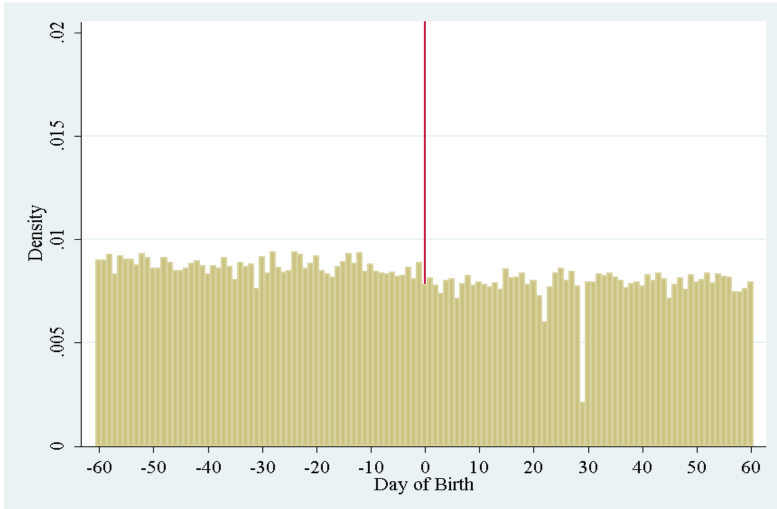

Source: Colasanti, Michael. Kindergarten Entrance Ages: A 30 Year Trend Analysis. March 2007. Education Commission of the States.

**Figure2.1**: Changes in School Entry Cutoff Dates by State



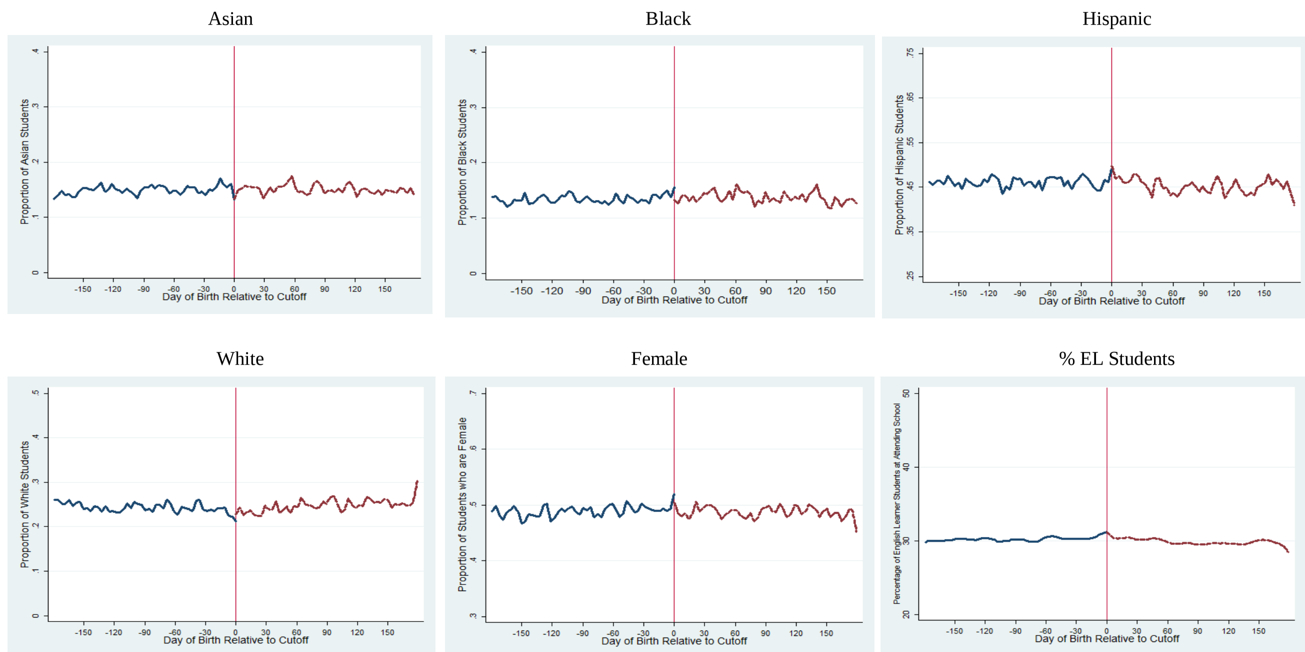

Notes: Figure presents the first stage discontinuity in treatment: the proportion of students with delayed entry into Kindergarten, by day of birth. Units are days after December 2, with December 2 centered at zero. Delayed entry is defined as Kindergarten entry during the academic year in which the student turns six years of age.

**Figure2.2**: Discontinuity in Delayed Kindergarten Entry



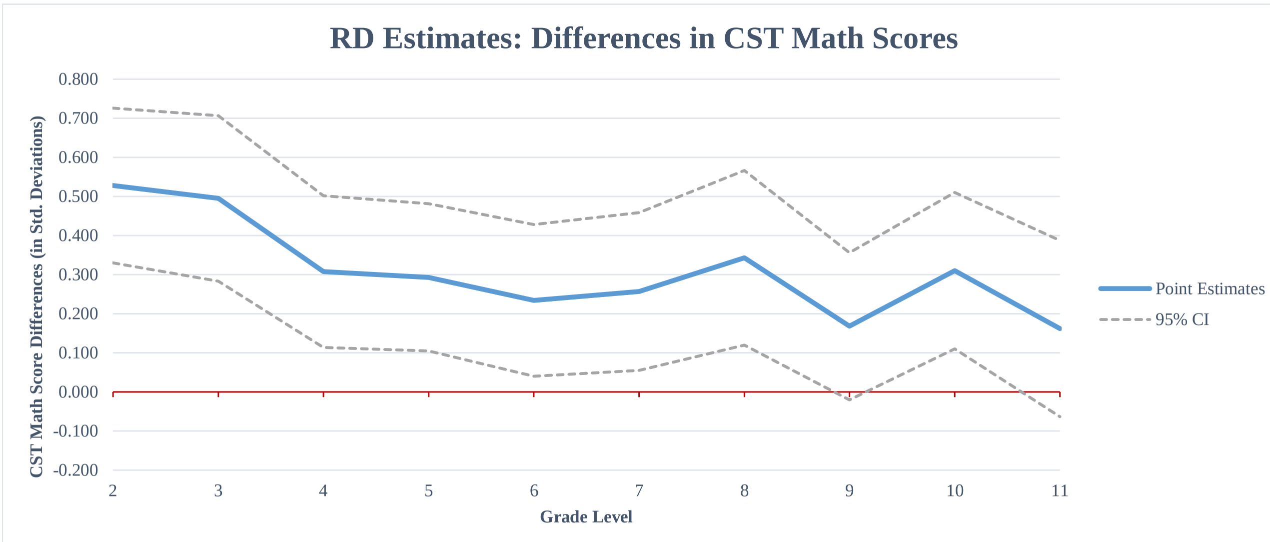

Notes: Figure is the density of birth dates. Units are days after December 2, with December 2 centered at zero.

**Figure2.3**: Density of Day of Birth



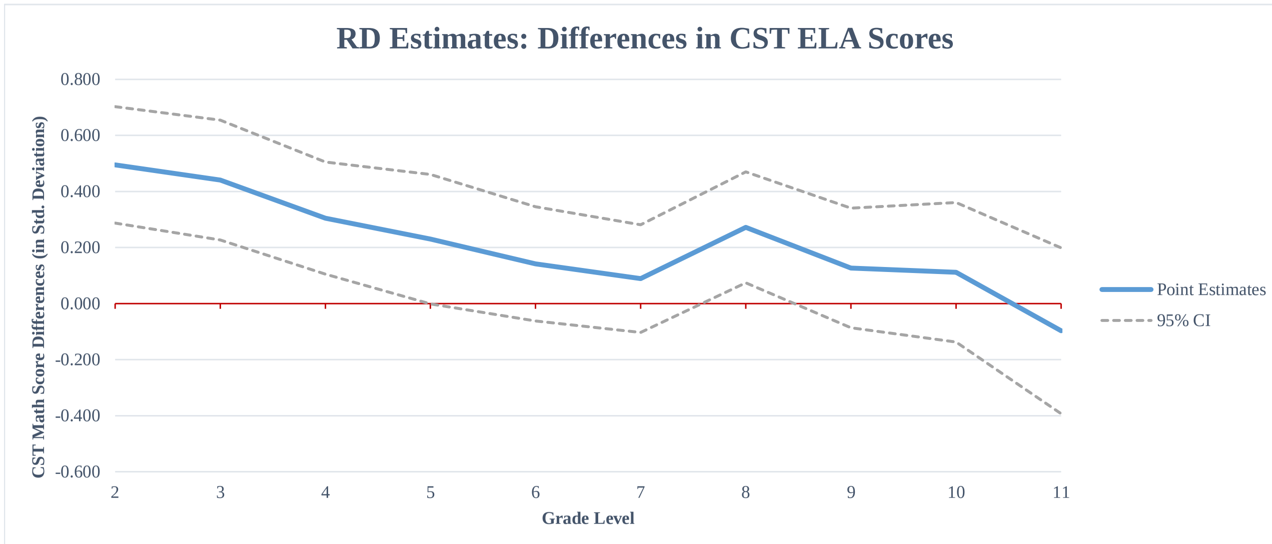

Notes: Plotted are the distributions of student characteristics determined prior to school entry, around the cutoff date. Importantly, there does not appear to be any discontinuities or clumping in the distributions of thes variables near the cutoff.

**Figure2.4**: Distributions of Covariates Around Cutoff



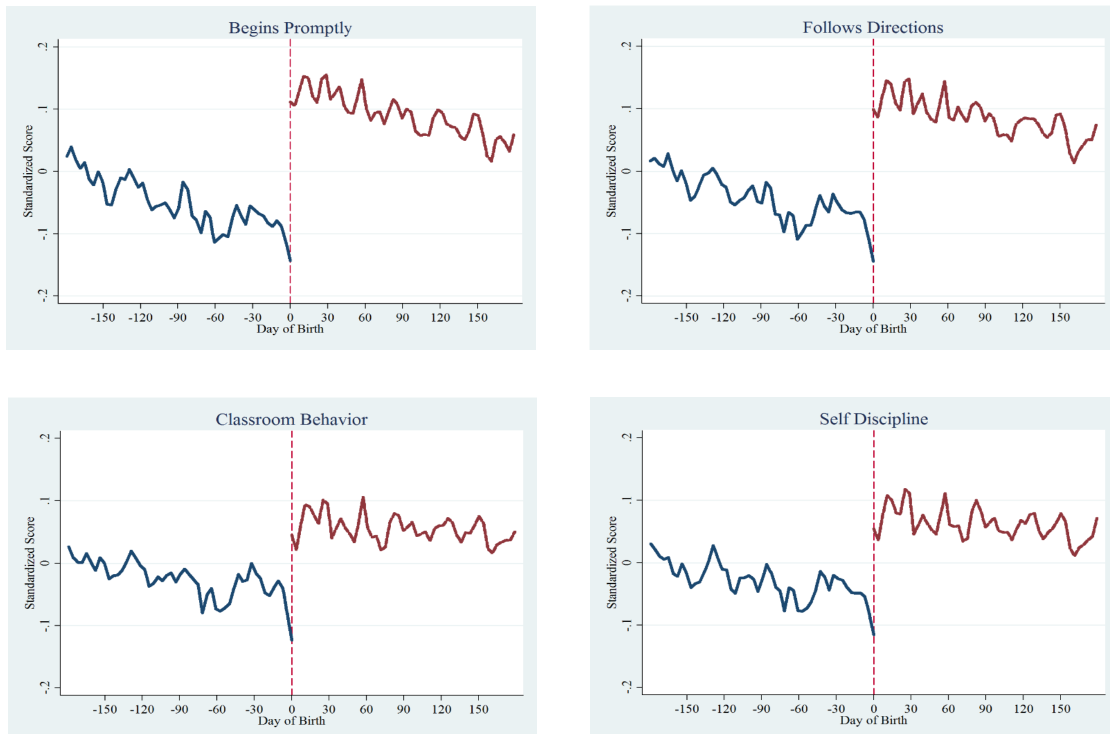

Notes: RD estimates differences in CST math scores for each grade level plotted with 95% confidence interval.

**Figure2.5**: RD Estimates of Math Achievement Differences by Grade Level



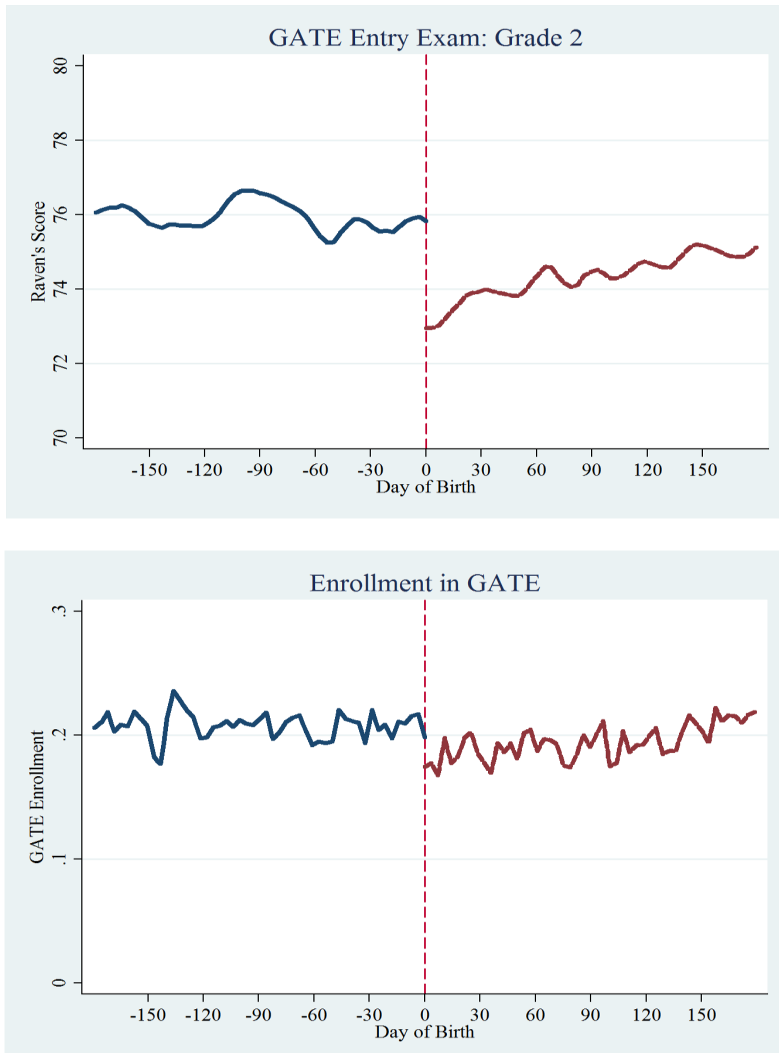

Notes: RD estimates differences in CST English languate arts scores for each grade level plotted with 95% confidence interval.

**Figure2.6**: RD Estimates of ELA Achievement Differences by Grade Level



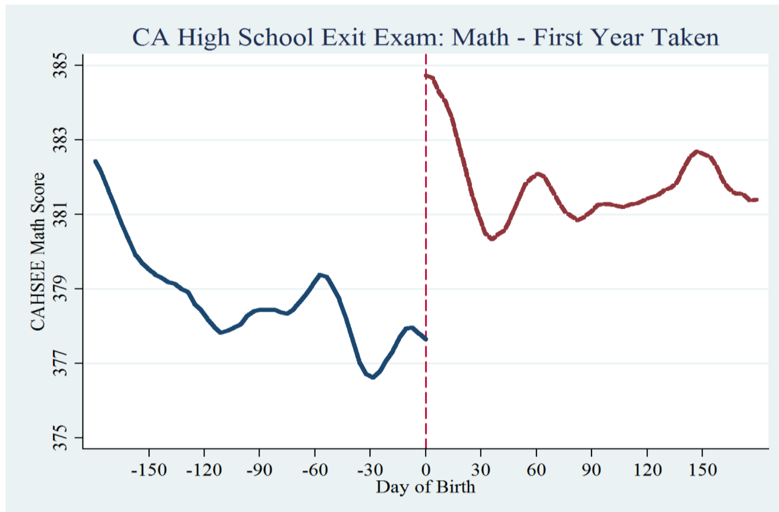

Notes: Figures present local polynomial plots of standardized behavioral evaluation scores by day of birth. Scores are standardized by grade level. Local linear regressions using these groupings of elementary students include indicators for grade level, as well as a standard set of covariates (see text). Standard errors from these specifications are clustered at the student level to account for repeat student observations across grade levels.

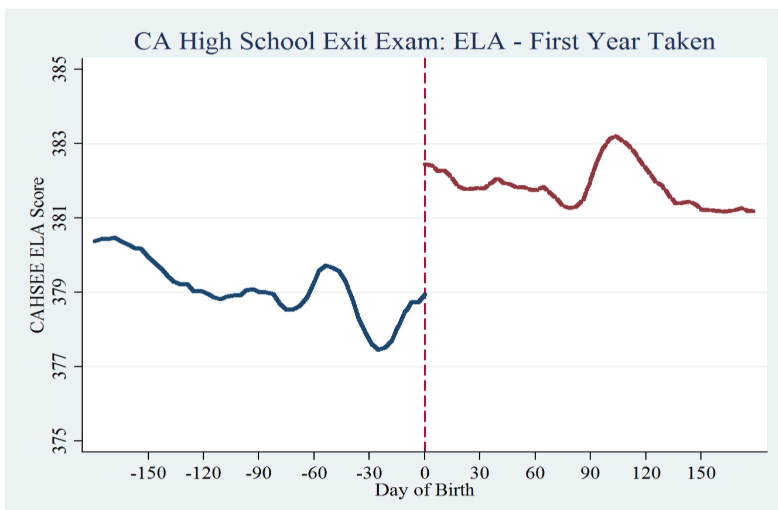
**Figure2.7**: Student Behavior Evaluations Around Birth Date Cutoff




Notes: The top pane presents Raven's Progressive Matrices (RPM) scores for students taking the test for the first time in second grade. Markedly higher scores illustrate the tendancy for RPM score scalings by age-at-test to benefit the youngest students. The bottom pane shows that despite this practice, the youngest students are not more likely to be enrolled in GATE classrooms.

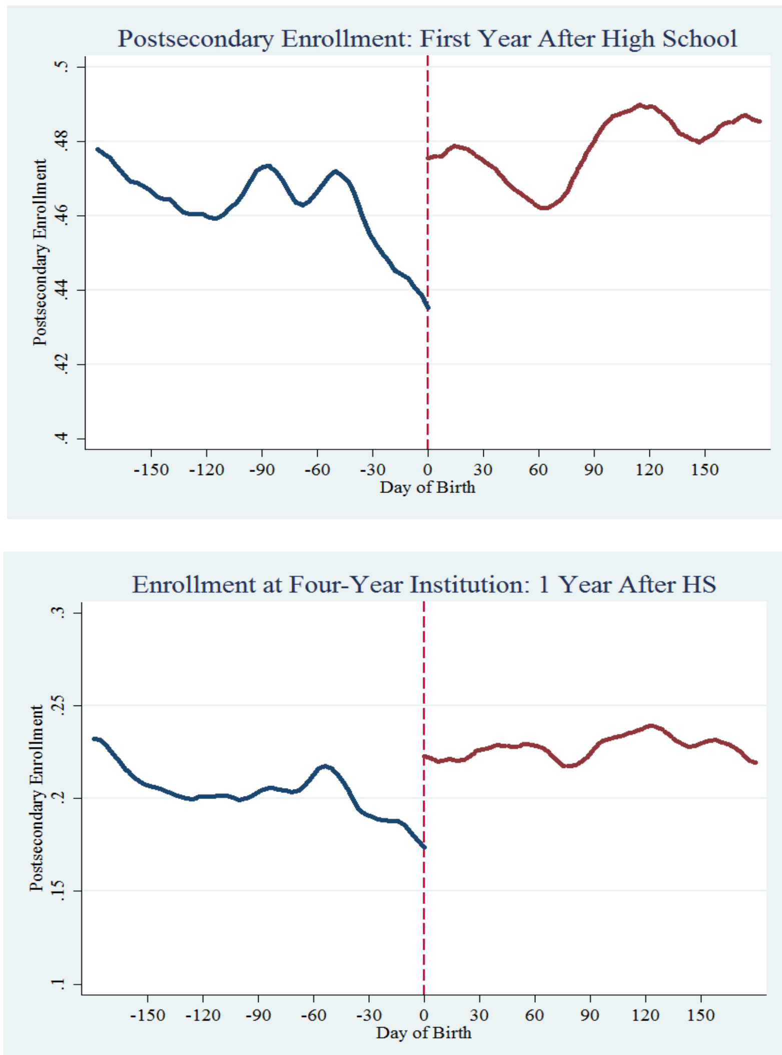
**Figure2.8**: GATE Testing and Enrollment by Day of Birth




Notes: Local polynomial plots of student California high school exit exam (CAHSEE) scores in grade 10, the first year students attempt the exam. Because of changes in the format and difficulty of the CAHSEE beginning with the 2003-2004 academic year, scores from Spring 2004 and later are used.

**Figure2.9**: California High School Exit Exam: First Attempt




Notes: Figure plots the proportion of observed high school students who enroll in any postsecondary (top) and a four-year postsecondary institution (bottom), by date of birth.

**Figure2.10**: Postsecondary Enrollments: First Year After High School

**Table2.1**: Summary Statistics by Grade Level

| | Grade K | Grade 1 | Grade 2 | Grade 3 | Grade 4 | Grade 5 | Grade 6 | Grade 7 | Grade 8 | Grade 9 | Grade 10 | Grade 11 | Grade 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # Grade Retentions | 0.029 | 0.030 | 0.017 | 0.012 | 0.007 | 0.006 | 0.009 | 0.010 | 0.011 | 0.056 | 0.070 | 0.058 | 0.200 |
| | (0.169) | (0.171) | (0.132) | (0.108) | (0.086) | (0.076) | (0.096) | (0.102) | (0.106) | (0.263) | (0.285) | (0.255) | (0.561) |
| Total # Retentions | 0.145 | 0.158 | 0.168 | 0.176 | 0.185 | 0.197 | 0.211 | 0.227 | 0.245 | 0.267 | 0.272 | 0.249 | 0.283 |
| | (0.405) | (0.421) | (0.435) | (0.447) | (0.462) | (0.476) | (0.494) | (0.512) | (0.533) | (0.556) | (0.564) | (0.551) | (0.607) |
| Raven's Progressive Matrices Score | | | 76.757 | 76.465 | 76.287 | 76.184 | 76.006 | 75.884 | 75.861 | 75.706 | 76.113 | 76.649 | 76.658 |
| | | | (24.295) | (24.435) | (24.600) | (24.703) | (24.907) | (25.082) | (25.240) | (25.467) | (25.429) | (25.334) | (25.656) |
| CST Math | | | 0.037 | 0.031 | -0.043 | -0.040 | -0.023 | 0.023 | 0.085 | -0.163 | -0.382 | -0.581 | |
| | | | (1.017) | (1.018) | (0.993) | (0.998) | (0.989) | (1.018) | (1.030) | (0.888) | (0.821) | (0.814) | |
| CST ELA | | | 0.087 | 0.071 | 0.030 | 0.027 | 0.024 | -0.007 | 0.020 | 0.058 | 0.055 | 0.0996 | |
| | | | (1.034) | (1.034) | (1.031) | (1.045) | (1.042) | (1.017) | (1.024) | (1.011) | (1.014) | (0.993) | |
| % Days Absent | 5.643 | 4.892 | 4.394 | 4.187 | 4.244 | 4.207 | 4.501 | 4.920 | 5.325 | 5.347 | 5.614 | 5.279 | 5.644 |
| | (6.405) | (5.988) | (5.692) | (5.512) | (5.563) | (5.495) | (5.883) | (6.579) | (7.430) | (8.949) | (9.252) | (8.382) | (8.529) |
| Behavior Evaluations | | | | | | | | | | | | | |
| Respects Others | 2.753 | 2.754 | 2.752 | 2.754 | 2.757 | 2.775 | | | | | | | |
| | (0.448) | (0.449) | (0.452) | (0.448) | (0.454) | (0.437) | | | | | | | |
| Shows Interest | 2.772 | 2.747 | 2.725 | 2.704 | 2.682 | 2.693 | | | | | | | |
| | (0.424) | (0.443) | (0.460) | (0.474) | (0.500) | (0.492) | | | | | | | |
| Student Preparedness | 2.686 | 2.647 | 2.621 | 2.625 | 2.606 | 2.637 | | | | | | | |
| | (0.501) | (0.533) | (0.553) | (0.553) | (0.572) | (0.553) | | | | | | | |
| Completes Assignments | 2.677 | 2.636 | 2.621 | 2.625 | 2.604 | 2.625 | | | | | | | |
| | (0.511) | (0.540) | (0.547) | (0.550) | (0.572) | (0.562) | | | | | | | |
| Begins Promptly | | 2.999 | 3.031 | 3.073 | 3.021 | 3.064 | | | | | | | |
| | | (0.886) | (0.894) | (0.905) | (0.924) | (0.930) | | | | | | | |
| Follows Directions | | 2.953 | 2.980 | 3.020 | 3.016 | 3.062 | | | | | | | |
| | | (0.910) | (0.913) | (0.918) | (0.921) | (0.928) | | | | | | | |
| Classroom Behavior | | 3.012 | 3.039 | 3.080 | 3.049 | 3.094 | | | | | | | |
| | | (0.895) | (0.906) | (0.900) | (0.915) | (0.922) | | | | | | | |
| Self Discipline | | 2.933 | 2.957 | 2.987 | 2.968 | 3.011 | | | | | | | |
| | | (0.951) | (0.969) | (0.966) | (0.981) | (0.992) | | | | | | | |
| Observations | 237,128 | 196,827 | 166,307 | 144,374 | 127,338 | 113,294 | 100,307 | 88,550 | 77,515 | 68,593 | 57,306 | 45,808 | 37,906 |

Notes: Means reported by grade level. Standard deviations in parentheses. CST scores are standardized by grade level and year using statewide means and std. deviations. Behavior scores are standardized in the analyses, but presented as raw scores here for reference. Scores range from 1 to 4 in 0.25 point increments.

**Table2.2**: Tests for Covariate Discontinuities Around Cutoff Date

| RD Estimates of Breaks at Discontinuity | Female (1) | Asian (2) | Black (3) | Hispanic (4) | White (5) | Other (6) | English Learner (8) |
|---|---|---|---|---|---|---|---|
| IK Optimal Bandwidth | -0.096 | -0.009 | 0.017 | 0.003 | -0.028 | 0.007 | 0.014 |
| | (0.075) | (0.054) | (0.030) | (0.076) | (0.063) | (0.007) | -0.054 |
| | [0.203] | [0.863] | [0.575] | [0.965] | [0.653] | [0.340] | [0.796] |
| IK Optimal Bandwidth * 0.5 | -0.025 | -0.024 | 0.008 | -0.022 | 0.026 | . | -0.033 |
| | (0.043) | (0.030) | (0.029) | (0.043) | (0.037) | . | (0.031) |
| | [0.570] | [0.428] | [0.792] | [0.603] | [0.479] | . | [0.296] |
| IK Optimal Bandwidth * 2 | -0.011 | -0.011 | -0.001 | -0.018 | 0.026 | 0.006 | -0.024 |
| | (0.047) | (0.036) | (0.038) | (0.052) | (0.042) | (0.010) | -0.036 |
| | [0.806] | [0.763] | [0.986] | [0.732] | [0.530] | [0.553] | [0.532] |
| Calculated IK Bandwidth | 2.236 | 1.873 | 1.987 | 2.149 | 1.858 | 1.351 | 1.882 |
| | | | Means and Std. Deviations | | | | |
| | 0.489 | 0.151 | 0.133 | 0.457 | 0.247 | 0.005 | 0.405 |
| | (0.500) | (0.358) | (0.339) | (0.498) | (0.431) | (0.070) | (0.491) |
| Observations | 100,422 | 100,422 | 100,422 | 100,422 | 100,422 | 100,422 | 100,422 |
| | | Parent Educational Attainment | | | | | |
| | % Free Lunch (9) | Not HS Grad (11) | HS Grad (12) | Some College (13) | College Grad (14) | | Post-graduate (15) |
| IK Optimal Bandwidth | -2.138 | 0.073 | 0.015 | -0.087 | -0.009 | | 0.021 |
| | (2.098) | (0.047) | (0.033) | (0.062) | (0.058) | | (0.029) |
| | [0.308] | [0.122] | [0.652] | [0.162] | [0.879] | | [0.469] |
| IK Optimal Bandwidth * 0.5 | -1.248 | 0.048* | . | -0.054 | 0.007 | | . |
| | (3.283) | (0.028) | . | (0.035) | (0.033) | | . |
| | [0.704] | [0.085] | . | [0.121] | [0.830] | | . |
| IK Optimal Bandwidth * 2 | -1.033 | 0.059* | 0.025 | -0.062 | 0.013 | | 0.015 |
| | (1.454) | (0.030) | (0.042) | (0.040) | (0.039) | | (0.038) |
| | [0.478] | [0.052] | [0.552] | [0.121] | [0.7446] | | [0.694] |
| Calculated IK Bandwidth | 7.077 | 1.978 | 1.999 | 2.209 | 2.104 | | 1.807 |
| | | | Means and Std. Deviations | | | | |
| | 63.341 | 0.114 | 0.197 | 0.196 | 0.182 | | 0.141 |
| | (31.541) | (0.318) | (0.398) | (0.397) | (0.386) | | (0.348) |
| Observations | 96,721 | 100,422 | 100,422 | 100,422 | 100,422 | | 100,422 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1), p-values in brackets. Each column reports estimates for an observable as a dependant variable in the main regression discontinuity specification for differences above and below the date of birth cutoff for entry to public school in California. Optimal Bandwidths are calculated using the Imbens and Kalyanaraman (IK) algorithm. Estimates using 100, 50 and 200 percent of the calculated optimal bandwidth are reported for each variable. Missing values correspond to cases where the number of observations were too few to properly estimate standard errors.

**Table2.3**: RD Estimates of Delayed Entry on Student Achievement

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Grade 2 | Grade 3 | Grade 4 | Grade 5 | Grade 6 | Grade 7 | Grade 8 | Grade 9 | Grade 10 | Grade 11 |
| Estimator | | | | | CST ELA Scores | | | | | |
| Reduced Form | 0.405*** | 0.353*** | 0.248*** | 0.190* | 0.119 | 0.075 | 0.229*** | 0.103 | 0.091 | -0.084 |
| | (0.087) | (0.087) | (0.084) | (0.097) | (0.087) | (0.084) | (0.085) | (0.088) | (0.103) | (0.129) |
| First Stage | 0.818*** | 0.800*** | 0.811*** | 0.827*** | 0.837*** | 0.850*** | 0.843*** | 0.812*** | 0.806*** | 0.858*** |
| | (0.027) | (0.014) | (0.031) | (0.032) | (0.026) | (0.028) | (0.026) | (0.031) | (0.035) | (0.047) |
| **RD Estimate** | **0.495*** ** | **0.441*** ** | **0.305*** ** | **0.230* ** | **0.142** | **0.089** | **0.272*** ** | **0.127** | **0.112** | **-0.097** |
| | **(0.106)** | **(0.109)** | **(0.102)** | **(0.118)** | **(0.104)** | **(0.098)** | **(0.101)** | **(0.109)** | **(0.127)** | **(0.151)** |
| Observations | 82,548 | 79,132 | 75,347 | 73,941 | 70,448 | 66,851 | 64,507 | 60,238 | 50,344 | 39,696 |
| Estimator | | | | | CST Math Scores | | | | | |
| Reduced Form | 0.432*** | 0.397*** | 0.250*** | 0.241*** | 0.195** | 0.217** | 0.297*** | 0.138* | 0.254*** | 0.130 |
| | (0.084) | (0.087) | (0.081) | (0.079) | (0.083) | (0.087) | (0.098) | (0.079) | (0.082) | (0.093) |
| First Stage | 0.818*** | 0.802*** | 0.811*** | 0.821*** | 0.833*** | 0.843*** | 0.866*** | 0.821*** | 0.819*** | 0.807*** |
| | (0.027) | (0.014) | (0.028) | (0.026) | (0.026) | (0.026) | (0.029) | (0.030) | (0.033) | (0.038) |
| **RD Estimate** | **0.528*** ** | **0.495*** ** | **0.308*** ** | **0.293*** ** | **0.234** ** | **0.257** ** | **0.343*** ** | **0.168* ** | **0.310*** ** | **0.162** |
| | **(0.101)** | **(0.108)** | **(0.099)** | **(0.096)** | **(0.099)** | **(0.103)** | **(0.114)** | **(0.096)** | **(0.102)** | **(0.115)** |
| Observations | 82,824 | 79,289 | 76,192 | 73,959 | 70,472 | 67,398 | 63,612 | 59,219 | 48,805 | 37,817 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). CST scores are standardized by grade level and year using statewide means and standard deviations. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**Table2.4**: RD Estimates of Delayed Entry on Grade Retention

| | (1) Kindergarten | (2) Grade1 | (3) Grade 2 | (4) Grade 3 | (5) Grade 4 |
|---|---|---|---|---|---|
| Estimator | Any Grade Level Specific Retention | | | | |
| Reduced Form | -0.077*** | -0.042** | -0.002 | -0.005 | -0.006 |
| | (0.019) | (0.017) | (0.014) | (0.013) | (0.011) |
| First Stage | 0.808*** | 0.835*** | 0.840*** | 0.854*** | 0.869*** |
| | (0.024) | (0.026) | (0.032) | (0.040) | (0.043) |
| **RD Estimate** | **-0.096*** ** | **-0.050** ** | **-0.002** | **-0.005** | **-0.007** |
| | **(0.024)** | **(0.020)** | **(0.017)** | **(0.015)** | **(0.013)** |
| Observations | 237,128 | 196,827 | 166,307 | 144,374 | 127,338 |
| Estimator | Grade Level Retention, Including Covariates | | | | |
| Reduced Form | -0.079*** | -0.050** | -0.007 | -0.009 | -0.003 |
| | (0.023) | (0.019) | (0.016) | (0.014) | (0.011) |
| First Stage | 0.799*** | 0.822*** | 0.842*** | 0.813*** | 0.820*** |
| | (0.027) | (0.029) | (0.034) | (0.015) | (0.016) |
| **RD Estimate** | **-0.099*** ** | **-0.060** ** | **-0.009** | **-0.011** | **-0.003** |
| | **(0.028)** | **(0.024)** | **(0.019)** | **(0.017)** | **(0.013)** |
| | Means and Std. Deviations | | | | |
| | 0.065 | 0.144 | 0.154 | 0.161 | 0.168 |
| | (0.246) | (0.351) | (0.361) | (0.367) | (0.374) |
| Observations | 237,128 | 196,827 | 166,307 | 144,374 | 127,338 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Grade level retention is defined as any repeat year for the specific grade level modeled. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**Table2.5**: RD Estimates of Delayed Entry on Student Behavior

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | Behavior Evaluations: Grades 1-5 | | | | |
| Estimator | | Begins Promptly | Follows Directions | Classroom Behavior | Self Discipline | |
| Reduced Form | | 0.324*** | 0.311*** | 0.283*** | 0.253*** | |
| | | (0.054) | (0.054) | (0.052) | (0.053) | |
| First Stage | | 0.804*** | 0.804*** | 0.804*** | 0.803*** | |
| | | (0.008) | (0.008) | (0.008) | (0.008) | |
| **RD Estimate** | | **0.403*** ** | **0.387*** ** | **0.352*** ** | **0.315*** ** | |
| | | **(0.067)** | **(0.067)** | **(0.064)** | **(0.066)** | |
| Observations | | 239,201 | 239,219 | 239,194 | 238,989 | |
| | Kindergarten | Grade1 | Grade 2 | Grade 3 | Grade 4 | Grade 5 |
| Estimator | | | Percentage of Days Absent | | | |
| Reduced Form | -0.684 | -0.368 | -0.339 | -0.279 | -0.675** | -0.533 |
| | (0.527) | (0.330) | (0.319) | (0.305) | (0.299) | (0.373) |
| First Stage | 0.774*** | 0.787*** | 0.798*** | 0.802*** | 0.800*** | 0.825*** |
| | (0.021) | (0.019) | (0.019) | (0.011) | (0.019) | (0.010) |
| **RD Estimate** | **-0.884** | **-0.467** | **-0.425** | **-0.348** | **-0.844** ** | **-0.646** |
| | **(0.682)** | **(0.419)** | **(0.400)** | **(0.380)** | **(0.374)** | **(0.453)** |
| Observations | 112,292 | 98,256 | 90,519 | 84,895 | 81,145 | 78,566 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Behavior evaluations are standardized by grade level. RD estimates reported are the differences in dependent variables between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels. Specifications which group grade levels also include indicators for each grade. The reported standard errors are clustered at the student level when grouped by grade level.

**Table2.6**: RD Estimates of Student Behavior by Grade

| | (1) Grade1 | (2) Grade 2 | (3) Grade 3 | (4) Grade 4 | (5) Grade 5 |
|---|---|---|---|---|---|
| Estimator | | | Begins Promptly | | |
| Reduced Form | 0.343** | 0.332*** | 0.436*** | 0.239** | 0.407*** |
| | (0.146) | (0.111) | (0.126) | (0.106) | (0.131) |
| First Stage | 0.796*** | 0.783*** | 0.796*** | 0.791*** | 0.827*** |
| | (0.020) | (0.037) | (0.019) | (0.035) | (0.020) |
| **RD Estimate** | **0.431**** | **0.424***** | **0.548***** | **0.302**** | **0.492***** |
| | **(0.182)** | **(0.141)** | **(0.158)** | **(0.134)** | **(0.159)** |
| Observations | 53,601 | 50,523 | 47,881 | 45,799 | 41,397 |
| | | | Follows Directions | | |
| Reduced Form | 0.433*** | 0.382*** | 0.318*** | 0.232** | 0.329*** |
| | (0.138) | (0.102) | (0.102) | (0.100) | (0.120) |
| First Stage | 0.796*** | 0.783*** | 0.804*** | 0.790*** | 0.828*** |
| | (0.020) | (0.034) | (0.016) | (0.034) | (0.019) |
| **RD Estimate** | **0.545***** | **0.487***** | **0.396***** | **0.294**** | **0.397***** |
| | **(0.173)** | **(0.129)** | **(0.126)** | **(0.126)** | **(0.145)** |
| Observations | 53,616 | 50,532 | 47,879 | 45,785 | 41,407 |
| | | | Classroom Behavior | | |
| Reduced Form | 0.331** | 0.347*** | 0.234** | 0.205** | 0.297*** |
| | (0.129) | (0.098) | (0.101) | (0.103) | (0.110) |
| First Stage | 0.795*** | 0.783*** | 0.804*** | 0.791*** | 0.831*** |
| | (0.019) | (0.034) | (0.016) | (0.035) | (0.017) |
| **RD Estimate** | **0.417***** | **0.443***** | **0.291**** | **0.258**** | **0.357***** |
| | **(0.161)** | **(0.125)** | **(0.125)** | **(0.130)** | **(0.133)** |
| Observations | 53,583 | 50,512 | 47,886 | 45,799 | 41,414 |
| | | | Self Discipline | | |
| Reduced Form | 0.210* | 0.306*** | 0.221* | 0.162 | 0.312*** |
| | (0.112) | (0.117) | (0.115) | (0.111) | (0.109) |
| First Stage | 0.793*** | 0.776*** | 0.798*** | 0.797*** | 0.832*** |
| | (0.017) | (0.041) | (0.018) | (0.038) | (0.017) |
| **RD Estimate** | **0.265*** | **0.394***** | **0.277*** | **0.203** | **0.376***** |
| | **(0.141)** | **(0.151)** | **(0.143)** | **(0.139)** | **(0.131)** |
| Observations | 53,574 | 50,404 | 47,879 | 45,754 | 41,378 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Behavior evaluations are standardized by grade level. RD estimates reported are the differences in dependent variables between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**Table2.7**: RD Estimates of Classroom Characteristics, by Grade

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Kindergarten | Grade1 | Grade 2 | Grade 3 | Grade 4 | Grade 5 |
| Estimator | | | Average Peer Age (in days) | | | |
| Reduced Form | -0.432 | 2.121 | -1.631 | -6.907 | 11.152 | 15.125 |
| | (5.187) | (7.348) | (9.279) | (10.083) | (9.625) | (10.555) |
| First Stage | 0.852*** | 0.803*** | 0.814*** | 0.813*** | 0.825*** | 0.829*** |
| | (0.016) | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) |
| **RD Estimate** | **-0.507** | **2.641** | **-2.003** | **-8.497** | **13.520** | **18.238** |
| | **(6.091)** | **(9.149)** | **(11.400)** | **(12.410)** | **(11.662)** | **(12.721)** |
| Observations | 34,108 | 86,781 | 80,674 | 75,775 | 72,291 | 66,123 |
| | | | Classroom Size | | | |
| Reduced Form | 0.138 | 0.810 | 0.161 | -1.527* | -0.147 | 0.432 |
| | (0.364) | (0.579) | (0.718) | (0.921) | (1.171) | (0.999) |
| First Stage | 0.847*** | 0.794*** | 0.799*** | 0.798*** | 0.813*** | 0.825*** |
| | (0.031) | (0.017) | (0.019) | (0.017) | (0.018) | (0.020) |
| **RD Estimate** | **0.163** | **1.020** | **0.201** | **-1.912*** | **-0.181** | **0.523** |
| | **(0.431)** | **(0.730)** | **(0.899)** | **(1.154)** | **(1.441)** | **(1.211)** |
| Observations | 34,108 | 86,781 | 80,674 | 75,775 | 72,291 | 66,123 |
| | | | Teacher Years of Experience | | | |
| Reduced Form | -0.136 | 0.411 | 0.414 | 0.116 | 0.273 | -0.197 |
| | (0.837) | (0.580) | (0.567) | (0.696) | (0.619) | (0.649) |
| First Stage | 0.845*** | 0.798*** | 0.798*** | 0.788*** | 0.819*** | 0.830*** |
| | (0.025) | (0.019) | (0.018) | (0.021) | (0.019) | (0.020) |
| **RD Estimate** | **-0.161** | **0.515** | **0.519** | **0.148** | **0.333** | **-0.237** |
| | **(0.990)** | **(0.726)** | **(0.710)** | **(0.883)** | **(0.755)** | **(0.782)** |
| Observations | 33,490 | 83,596 | 78,448 | 73,505 | 69,615 | 63,594 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). RD estimates reported are the differences in dependent variables between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**Table2.8**: RD Estimates of Delayed Entry on Peer Quality, by Grade Level

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Grade 2 | Grade 3 | Grade 4 | Grade 5 |
| Estimator | Average Peer CST Math Scores | | | |
| Reduced Form | 0.051 | -0.008 | 0.027 | -0.011 |
| | (0.049) | (0.063) | (0.053) | (0.071) |
| First Stage | 0.810*** | 0.819*** | 0.804*** | 0.857*** |
| | (0.032) | (0.037) | (0.032) | (0.040) |
| **RD Estimate** | **0.063** | **-0.010** | **0.034** | **-0.012** |
| | **(0.061)** | **(0.077)** | **(0.066)** | **(0.083)** |
| Observations | 75,589 | 71,954 | 68,998 | 63,328 |
| | Average Peer CST ELA Scores | | | |
| Reduced Form | 0.034 | 0.033 | 0.011 | 0.018 |
| | (0.051) | (0.053) | (0.047) | (0.067) |
| First Stage | 0.808*** | 0.807*** | 0.796*** | 0.858*** |
| | (0.034) | (0.032) | (0.031) | (0.040) |
| **RD Estimate** | **0.042** | **0.041** | **0.013** | **0.021** |
| | **(0.063)** | **(0.065)** | **(0.059)** | **(0.078)** |
| Observations | 75,314 | 71,796 | 68,239 | 63,304 |

Notes: Dependent variables are average CST scores across the student's classroom peers for each particular grade level. RD estimates reported are the differences in dependent variables between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels. Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

**Table2.9**: RD Estimates of Delayed Entry on GATE Testing and Enrollment

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Estimator | Raven's Progressive Matrices Scores | | | | Age at Testing | |
| Reduced Form | -1.627 | -2.095 | 3.445** | | 256.933*** | 256.376*** |
| | (1.338) | (1.332) | (1.494) | | (5.812) | (6.113) |
| First Stage | 0.828*** | 0.823*** | 0.620*** | | 0.836*** | 0.837*** |
| | (0.015) | (0.015) | (0.015) | | (0.011) | (0.011) |
| **RD Estimate** | **-1.965** | **-2.545** | **5.552**** | | **307.363***** | **306.367***** |
| | **(1.616)** | **(1.616)** | **(2.405)** | | **(5.861)** | **(6.219)** |
| Observations | 90,929 | 90,929 | 90,929 | | 90,929 | 90,929 |
| | GATE Enrollment | | | GATE Cluster Program | | |
| Reduced Form | 0.036 | 0.002 | 0.085* | 0.055 | 0.028 | 0.106*** |
| | (0.043) | (0.042) | (0.045) | (0.039) | (0.038) | (0.041) |
| First Stage | 0.851*** | 0.826*** | 0.612*** | 0.851*** | 0.826*** | 0.612*** |
| | (0.033) | (0.035) | (0.030) | (0.033) | (0.035) | (0.030) |
| **RD Estimate** | **0.042** | **0.003** | **0.138*** | **0.064** | **0.033** | **0.173***** |
| | **(0.050)** | **(0.051)** | **(0.073)** | **(0.046)** | **(0.047)** | **(0.067)** |
| Observations | 90,929 | 90,929 | 90,929 | 90,929 | 90,929 | 90,929 |
| Covariates | No | Yes | Yes | No | Yes | Yes |
| Age at Test | No | No | Yes | No | No | Yes |

| Means and Std. Deviations | Grade 2 RPM Scores | GATE Participation | Age at RPM Test |
|---|---|---|---|
| | 75.675 | 0.224 | 2842 |
| | (25.029) | (0.417) | (152.46) |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). GATE Cluster refers to the first tier of GATE enrollment. No significant results were found for GATE seminar, the highest level of GATE enrollment. RD estimates reported are the differences in dependent variables between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**Table2.10**: RD Estimates of Delayed Entry on Mathematics Progression

| | (1) Grade 7 | (2) Grade 8 | (3) Grade 9 | (4) Grade 10 | (5) Grade 11 | (6) Grade 12 |
|---|---|---|---|---|---|---|
| Estimator | Below Math Course Progression Prescribed by Grade Level | | | | | |
| Reduced Form | -0.015 | -0.080* | -0.056 | 0.048 | -0.008 | -0.064 |
| | (0.020) | (0.044) | (0.055) | (0.056) | (0.060) | (0.065) |
| First Stage | 0.862*** | 0.870*** | 0.844*** | 0.829*** | 0.844*** | 0.845*** |
| | (0.034) | (0.030) | (0.034) | (0.035) | (0.041) | (0.040) |
| **RD Estimate** | **-0.018** | **-0.092*** | **-0.067** | **0.058** | **-0.009** | **-0.076** |
| | **(0.024)** | **(0.051)** | **(0.064)** | **(0.068)** | **(0.071)** | **(0.077)** |
| Observations | 88,550 | 77,515 | 68,593 | 57,306 | 45,808 | 37,906 |
| | Above Proficient in Math Course Progression | | | | | |
| Reduced Form | 0.058 | 0.119*** | 0.061 | 0.087** | 0.114** | 0.048 |
| | (0.039) | (0.043) | (0.053) | (0.045) | (0.058) | (0.052) |
| First Stage | 0.857*** | 0.856*** | 0.841*** | 0.835*** | 0.816*** | 0.858*** |
| | (0.029) | (0.032) | (0.033) | (0.038) | (0.041) | (0.044) |
| **RD Estimate** | **0.068** | **0.139***** | **0.073** | **0.105**** | **0.140**** | **0.056** |
| | **(0.046)** | **(0.050)** | **(0.063)** | **(0.053)** | **(0.070)** | **(0.061)** |
| Observations | 88,550 | 77,515 | 68,593 | 57,306 | 45,808 | 37,906 |
| | | | Advanced-Track Math Course as Prescribed by Grade Level | | | |
| Reduced Form | | | 0.008 | 0.037 | 0.045*** | 0.066 |
| | | | (0.028) | (0.023) | (0.017) | (0.045) |
| First Stage | | | 0.848*** | 0.836*** | 0.841*** | 0.858*** |
| | | | (0.037) | (0.040) | (0.046) | (0.053) |
| **RD Estimate** | | | **0.009** | **0.044** | **0.054***** | **0.077** |
| | | | **(0.033)** | **(0.027)** | **(0.020)** | **(0.052)** |
| Observations | | | 68,593 | 57,306 | 45,808 | 37,906 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Course progressions are defined according to presribed courses for each grade level under the California state academic standards for mathematics. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**Table2.11**: RD Estimates of Delayed Entry on High School Exit Exam

| | Math CAHSEE Score | | ELA CAHSEE Score | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Estimator | | | | |
| Reduced Form | 9.119*** | 8.182** | 4.370 | 4.652 |
| | (3.116) | (3.340) | (2.998) | (3.563) |
| First Stage | 0.826*** | 0.825*** | 0.821*** | 0.818*** |
| | (0.024) | (0.027) | (0.024) | (0.032) |
| **RD Estimate** | **11.039*** ** | **9.917** ** | **5.320** | **5.687** |
| | **(3.741)** | **(4.001)** | **(3.631)** | **(4.316)** |
| Calculated Bandwidth | 20.38 | 20.38 | 19.75 | 19.75 |
| Observations | 26,099 | 19,379 | 26,037 | 19,344 |
| Include Retained Students? | Not Retained Current Grade | Never Retained | Not Retained Current Grade | Never Retained |
| | Means and Std. Deviations | | | |
| | 380.17 | | 380.40 | |
| | (38.449) | | (36.109) | |
| Observations | 26,099 | | 26,037 | |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). CAHSEE scores for years 2004 and later during the first year attempted are used. Local linear regressions with bandwidths calculated using the IK algorithm. Regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**Table2.12**: RD Estimates of Delayed Entry on Secondary Attainment Levels

| | Completed at least 11th Grade | | High School Graduation | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Estimator | | | | |
| Reduced Form | 0.130** | 0.086* | 0.125* | 0.102* |
| | (0.062) | (0.049) | (0.074) | (0.059) |
| First Stage | 0.850*** | | 0.841*** | |
| | (0.046) | | (0.041) | |
| **RD Estimate** | **0.153**** | | **0.149*** | |
| | **(0.073)** | | **(0.089)** | |
| Calculated Bandwidth | 4.14 | 4.43 | 5.10 | 4.73 |
| Observations | 44,463 | 122,234 | 44,463 | 122,234 |
| | Means and Std. Deviations | | | |
| | 0.771 | | 0.602 | |
| | (0.420) | | (0.490) | |

Notes: Outcomes are matriculation to grade 12 and high school graduation, conditional on observed high school entry. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels. Bootstrapped standard errors in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

**Table2.13**: RD Estimates of Delayed Entry on Postsecondary Enrollment

| | Postsecondary Enrollment: 1st year after HS | | Enrollment at 4-year Institution: 1st year after HS | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Estimator | | | | |
| Reduced Form | 0.098 | 0.151** | 0.106* | 0.104** |
| | (0.072) | (0.062) | (0.058) | (0.046) |
| First Stage | 0.826*** | | 0.839*** | |
| | (0.041) | | (0.044) | |
| **RD Estimate** | **0.119** | | **0.126*** | |
| | **(0.089)** | | **(0.070)** | |
| Calculated Bandwidth | 6.01 | 4.98 | 5.16 | 4.42 |
| Observations | 36,881 | 84,361 | 36,881 | 84,361 |
| | Means and Std. Deviations | | | |
| | 0.410 | | 0.184 | |
| | (0.492) | | (0.388) | |

Notes: Outcomes are any postsecondary enrollment and enrollment at a four-year institution, conditional on being observed in SDUSD during at least grade 11. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels. Bootstrapped standard errors in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

**Table2.14**: RD Estimates of Delayed Entry on Postsecondary Persistence

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Enrolled at least 2 years | | Enrolled at least 3 years | | Enrolled 4 years | | Number of First 4 Years Enrolled | |
| Estimator | | | | | | | | |
| Reduced Form | 0.136 | 0.118* | 0.209** | 0.160*** | 0.158* | 0.154*** | 0.435 | 0.373** |
| | (0.098) | (0.062) | (0.101) | (0.060) | (0.095) | (0.056) | (0.312) | (0.186) |
| First Stage | 0.766*** | | 0.770*** | | 0.771*** | | 0.755*** | |
| | (0.065) | | (0.070) | | (0.071) | | (0.059) | |
| **RD Estimate** | **0.178** | | **0.271**** | | **0.205*** | | **0.577** | |
| | **(0.131)** | | **(0.135)** | | **(0.124)** | | **(0.419)** | |
| Calculated Bandwidth | 5.37 | 4.90 | 4.86 | 4.63 | | | 6.48 | 6.30 |
| Observations | 18,111 | 51,824 | 18,111 | 51,824 | 18,111 | 51,824 | 18,111 | 51,824 |
| | Means and Std. Deviations | | | | | | | |
| | 0.392 | | 0.332 | | 0.261 | | 1.442 | |
| | (0.488) | | (0.471) | | (0.439) | | (1.743) | |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Outcomes are of the number of years during the first four after high school during which the student was enrolled at a postsecondary institution, conditional on being observed in SDUSD during at least grade 11. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Regressions include indicators for student gender, ethnicity, English learner status at school entry and parental

**Table2.15**: RD Estimates of Delayed Entry on Postsecondary Attainment

| | Any Postsecondary Degree Attainment | | Four-Year Degree Attainement | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Estimator | | | | |
| Reduced Form | -0.025 | 0.031 | 0.048 | 0.035 |
| | (0.054) | (0.040) | (0.046) | (0.034) |
| First Stage | 0.767*** | | 0.771*** | |
| | (0.066) | | (0.070) | |
| **RD Estimate** | **-0.033** | | **0.062** | |
| | **(0.071)** | | **(0.060)** | |
| Observations | 18,111 | 51,824 | 18,111 | 51,824 |
| | Means and Std. Deviations | | | |
| | 0.085 | | 0.070 | |
| | (0.278) | | (0.254) | |

Notes: Outcomes are any postsecondary degree attainment and four-year degree attainment, conditional on being observed in SDUSD during at least grade 11. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels. Bootstrapped standard errors in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

# Chapter 3

# Is “Gifted and Talented” Education a Gift that Keeps Giving?

## 3.1 Introduction

Within U.S. public schools, the most prevalent early method of sorting young students into curricula by ability is the Gifted and Talented Education (GATE) program. The state of California established GATE for public schools in 1980, providing funding to local education agencies to develop new opportunities for high and underachieving students identified as gifted or talented. The opportunities are typically a form of tracking, wherein students identified for GATE are grouped into classrooms with differentiated and accelerated curricula.

This paper investigates the impacts of tracking into Gifted and Talented Education (GATE) classrooms on long-run student outcomes. Virtually all second grade students in San Diego take an intelligence test to qualify for GATE enrollment. There are two well defined score thresholds, each of which grants eligibility for enrollment in one of two types of GATE programs. Because these criteria often result in very small GATE classroom sizes, the remaining seats are available to students scoring above a slightly lower threshold on the GATE entry test. This policy creates a large jump in the probability of a student participating in GATE, from zero to roughly 0.63 at the lower threshold score. Using rich longitudinal data following San Diego students into their postsecondary careers, I exploit this discontinuity and use an RD design to estimate the impacts of high achievement tracking on postsecondary enrollment, persistence and attainment. This design is particularly interesting as it estimates these impacts for students just at the margin of GATE eligibility relative to those just below; a group of students performing below the standard levels required for GATE entry who arguably represent the lowest level of ability within a GATE classroom.

There are at least a few ways in which GATE participation can affect the outcomes of students. First, GATE curricula are taught with more depth and complexity than the core curriculum. To the extent that more advanced curricula translate to higher levels of learning, GATE can increase achievement levels and boost students' educational trajectories. Bhatt (2009) looks at the impacts of GATE on achievement and finds some evidence of large gains immediately after participation–0.86 standard deviations on eighth grade math scores–with effects quickly fading out afterward. Her estimates for other grades are positive, but not statistically significant. Bhatt also finds GATE participation positively effects the likelihood students take advanced placement (AP) courses in high school. The results of Bui, Craig and Imberman (2011) differ with those of Bhatt in that they find little to no impact on achievement levels for students who marginally qualify for GATE enrollment. They do, however, find that students on the margin of GATE eligibility see decreases in course grades and relative class rankings after GATE enrollment. Second, GATE teachers are generally required to have more teaching experience and either GATE-specific certifications or a relevant post-graduate degree. Bui, Craig and Imberman (2011) find that students admitted to GATE magnet schools have more highly qualified teachers than other students. Third, enrollment in a GATE classroom places students among peers of higher ability levels. Although Bhatt (2009) finds no evidence of GATE participation affecting peer quality, others who have studied impacts of GATE have found such effects. Davis et al. (2010) find that among the highest achievers, admittance to GATE decreases the likelihood of leaving the student's public school district. Bui, Craig and Imberman (2011) find large impacts of GATE on peer achievement levels; roughly 0.3 standard deviations. A recent working paper by Card and Giuliano (2013) find that students on the margin of being tracked into GATE classrooms realize a positive effect on test scores and, for boys, writing scores.

While the benefits to GATE exposure seem obvious, the intuition is not clear as to whether the GATE classroom experience generates positive or negative net impacts for students just on the margin of GATE eligibility. To the extent they face greater learning difficulties under advanced curricula or are ranked lower in performance relative to other GATE students, these students may be discouraged or even negatively impacted by participation. This paper contributes to the literature of gifted education programs by investigating how the net effects of these programs impact students' educational attainment level as adults.

The results suggest there are benefits to educational attainment for students right on the margin of inclusion in GATE classrooms, but they are likely realized for students among this group who initially aspire for higher levels of postsecondary persistence. GATE participation does not increase secondary attainment levels or postsecondary enrollment rates, but does increase the rate of postsecondary degree attainment among students on the margin of participating by roughly 8 percentage points. The paper proceeds as follows. Section 3.2 provides a brief overview of the GATE program in San Diego and the testing method by which students are granted admission. Section 3.3 describes the data and empirical methods used. Section 3.4 discusses the results and Section 3.6 ends with concluding remarks. Robustness checks are discussed in Section 3.5.

## 3.2 Overview of Gifted and Talented Education in San Diego

### 3.2.1 GATE Eligibility Testing

Students are typically identified as eligible for enrollment in GATE by performance on an intelligence test administered during the primary grades. In the San Diego Unified School District (SDUSD), students take the Raven's Progressive Matrices (RPM) test while in the second grade; specifically, the exam taken is the Standard Progressive Matrices Plus Version. The RPM is an untimed, non-reading test administered by a child psychologist. It consists of 60 multiple choice tasks which become progressively more difficult and complex. Figure 3.1 presents an example of the types of tasks second graders perform on the RPM. The test is visual and tests non-reading cognitive skills in an attempt to avoid ethnic and language biases while measuring intelligence.

Raw RPM scores out of 60 are converted to scores ranging from 0 to 100, adjusting for the student's age at the time of testing. District policy states that students may not retest until grade 5 and only if the student's initial score was at least

85[1]. The latter requirement is waived if the student has scored in the "advanced" range on at least two sections of the California Standards Test (CST) in grade 5. Students who enter the district after the second grade are eligible for initial testing during grades 3 through 7.

### 3.2.2 GATE Curricula

At the elementary level, GATE consists of two programs; the Seminar program and the Cluster program. Cluster is the most widely used program, while Seminar is a less utilized and more advanced GATE program reserved for students scoring extremely high on the RPM exam. Curricula are compacted in both programs, allowing students to test out of material they have mastered quickly, and is taught in more depth than in standard classrooms. Broadly, GATE classrooms offer a curriculum with greater depth, complexity and acceleration than the core curriculum of standard classrooms. At the elementary level, Cluster classes meet for the full school day and are separate from non-GATE classrooms. In middle school, Cluster offers multiple-period core classes and offers some high school courses to students as well. At the high school level, certain periods of course offerings are offered as Cluster classes.

This paper focuses on GATE classrooms under the Cluster program. Students who score above 98 on the RPM are eligible to participate in the Cluster program. Additionally, students who score above 95 and also satisfy certain socioeconomic, language, emotional or health criteria considered to be academic disadvantages are eligible for the Cluster program as well. The traditional Cluster model requires that Cluster classrooms are comprised of at minimum 25-50% GATE-identified students. Importantly, any remaining openings beyond those available to GATE identified students are filled by students scoring at least a 90 on the RPM, but not high enough to be GATE identified. The procedure used to fill the re-

[1]In very rare cases, a student may be allowed to retest in grade 2 if an appeal from the student's parent can convincingly show that extenuating circumstances resulted in test performance not indicative of the student's abilities. Written appeals are reviewed quarterly by the district. This is dealt with in the empirical methods as described in Section 3.3.

maining seats in the GATE Cluster program provides an opportunity to estimate the impacts of exposure to classroom with advanced curricula. As Figure 3.2 illustrates, students scoring just below 90 on the RPM are not eligible for enrollment in GATE. Those scoring 90 or above, however, are likely to be tracked into GATE classrooms with more advanced curricula and peers exhibiting high levels of achievement–many with RPM scores above 95 or 98. Because students scoring 99 and above are tracked into Seminar classrooms[2]–a relatively rare and different type of classroom–this paper is focused on Cluster and students scoring 95 and below on the RPM. The remainder of this paper uses the terms GATE and Cluster interchangeably.

## 3.3 Empirical Methods and Data

### 3.3.1 Data and Estimation Sample

The primary data come from administrative records on students and teachers maintained at San Diego Unified School District (SDUSD). San Diego has a large and diverse student body. SDUSD is the second largest public school district in California and has ranked between the eighth largest and twentieth largest in the U.S. over the years encompassed by our sample. 75 percent of the student body is non-white, over 30 percent are English language learners (EL) and roughly 60 percent qualify for the federal lunch-assistance program[3].

The data consist of longitudinally linked records on each student in the district, providing details year by year of students' school locations and classrooms, attendance, grades, behavioral evaluations, test scores, grade promotion and whether

[2]Students who score above 99.9 on the Raven test are eligible for the Seminar portion of the GATE program. Additionally, students who score between 99.6 and 99.9 and also satisfy certain academic disadvantage criteria are also eligible for Seminar. The Seminar program places the top 0.1 percentile of students in accelerated classrooms no bigger than 25 students in size. These classes are full-day, self contained classrooms at the elementary level. They offer advanced, accelerated curriculum beyond that of the Cluster program. The curriculum presented in each Seminar classroom dynamically targets the students' ability levels.

[3]See Betts, Zau and Rice (2003) for evidence that SDUSD is broadly representative of the student population of California.

the student was retained or skipped a grade level. In addition, the database also includes information on the specific teachers of each student, and detailed qualifications of these teachers. Combining student records from SDUSD with postsecondary data from the National Student Clearinghouse (NSC), the panel of data follows students from Kindergarten through the Baccalaureate level. Among outcomes at the postsecondary level are indicators for any postsecondary enrollment in the first year after graduating from high school, the type of postsecondary institution attended in this first year (community college or four-year college), and indicators for obtaining a two-year degree within four years after graduation or a four-year degree within four or five years after graduation.

The sample used in the paper includes second grade students with observed initial RPM scores from 1995 through 1998. The postsecondary outcomes of interest are observed beginning in Fall 2002.To allow at least four years after high school for students to complete a postsecondary degree, the sample for postsecondary outcomes is restricted to the three cohorts taking their first RPM test in second grade between the 1994-1995 and 1996-1997 academic years.

### 3.3.2 Estimating the Effects of Gifted Education on Educational Attainment

This paper is interested in the causal effects of being sorted into Gifted Education classrooms on long-run educational outcomes. Students who participate in GATE programs tend to have high levels of academic ability–partly because of the testing requirements for eligibility–and are therefore likely to attain higher levels of education than other students, regardless of any benefits GATE may provide. The endogeneity of GATE participation is overcome by exploiting the discontinuity in GATE eligibility around a clearly defined threshold score on the Raven's Progressive Matrices (RPM) exam, which is administered to all San Diego students in second grade. As Figure 3.2 illustrates, this feature of GATE provides an excellent application of a fuzzy regression discontinuity (RD) design, where the probability of GATE enrollment has a large and discrete jump at scores of 90 on the RPM.

Following Imbens and Lemieux (2008) and Lee and Lemieux (2010), the impacts of GATE participation are estimated using local linear regressions (LLR) of the form

$$Y_{ji} = \alpha_j + f_j(RPM_i) + \phi_j' X_i + u_{ji} \tag{3.1}$$

$$D_{ji} = \gamma_j + g_j(RPM_i) + \theta_j' X_i + u_{ji} \tag{3.2}$$

$$j = \begin{cases} l, & RPM_i \in [-h, 0) \\ r, & RPM_i \in [0, h] \end{cases}.$$

In the above equations, $Y_{ij}$ is the outcome of interest for student $i$ scoring within $h$ points of the GATE enrollment threshold, on the $j$ side. $D_{ji}$ is an indicator that student $i$ on side $j$ was enrolled in a GATE classroom for at least one grade level during the elementary years. The variable $RPM_i = (Score_i - 90)$ is the observed Raven's score for student $i$ minus 90, the distance of student $i$'s score from the eligibility threshold for GATE enrollment centered at zero. $X_i$ is a set of student level covariates–each predetermined before testing–progressively included in the estimating equations. Equations (3.1) and (3.2) describe the discontinuities of the outcome $Y_i$ and delayed entry treatment $D_i$, respectively, around the RPM threshold. Regressions are fit separately within a bandwidth $h$ on either side of the cutoff date. Functions $f_j(.)$ and $g_j(.)$ are estimated using the form $f_j(RPM_i) = \beta_j RPM_i$. Given valid conditions for the fuzzy RD design, the resulting estimand is a form of the Wald estimator:

$$\delta = \frac{lim_{x\downarrow 0} E\left[Y \mid RPM = x\right] - lim_{x\uparrow 0} E[Y \mid RPM = x]}{lim_{x\downarrow 0} E\left[D \mid RPM = x\right] - lim_{x\uparrow 0} E[D \mid RPM = x]} = \frac{\alpha_r - \alpha_l}{\gamma_r - \gamma_l} \tag{3.3}$$

This is the difference in average outcomes for students scoring on just either side of 90 points on the RPM, normalized by the discrete change in the likelihood of GATE enrollment as scores move across the threshold. This provides a local average treatment effect (LATE) for students scoring 90 on the RPM. These estimates are numerically equivalent to 2SLS estimates of the form

$$Y_i = \alpha_l + f_l(RPM_i) \cdot 1\{RPM_i < 0\} + f_r(RPM_i) \cdot 1\{RPM_i \geq 0\} + \delta D_i + u_i. \quad (3.4)$$

An indicator for "GATE eligible", $Z_i = 1\{RPM_i \geq 0\}$, is used to instrument for observed GATE enrollment $D_i$.

To choose estimation bandwidths, the Imbens and Kalyanaraman (2009) bandwidth algorithm (IK), which minimizes MSE in the sharp RD case, is used as a starting point. For the fuzzy RD design, there is not a clear case for an optimal bandwidth. A larger bandwidth increases precision as the within-bin sample grows, but may increase bias when the curvature of the specified functional form is a poor fit close to the discontinuity point. To account for the increase in standard errors that result from dividing the reduced form RD estimate by the first stage treatment estimate, the main results reported use a bandwidth of 4 RPM score points on either side of the 90 point threshold for enrollment. This amounts to slightly less than twice the average IK optimal bandwidth calculated for the reduced form, which is roughly 2.2 on average. Evidence that the results are not sensitive to changes in bandwidth or functional form around this point is provided in the appendix and discussed below. Computationally, these calculations are done using techniques discussed in Nichols (2009). Bootstrapped standard errors are reported for the results that follow.

### 3.3.3 RD Design Validity Tests

There are two main threats to the valid use of an RD design. The first is the possibility of manipulation of the forcing variable. In the framework of this paper, this would mean individuals are able to manipulate RPM scores in order to be on one side of the cutoff. Though it is rare, some students retake the RPM in second grade through parental appeals to the district showing clear evidence that initial scores are not indicative of the student's abilities. These retest appeals can be granted if, for instance, it is clear the student was ill or facing other extenuating circumstances during the testing window. Such appeals from parents are likely to be correlated with unobservables that may also affect outcomes, nullifying the quasi-experimental nature of the RD design. To deal with this type of manipulation, only scores from each student's first RPM attempt during the standard second grade testing window are used. The second threat to validity is the existence of discontinuities in the distributions of other variables around the GATE eligibility threshold. This can also create issues with the validity of RD estimates if any such variables are also correlated with the outcomes of interest.

Two checks are done to investigate whether these issues are likely. Although it is not possible to test for discontinuities along unobservables, I argue that if no such problems are found across the set of observable characteristics, they are unlikely to exist along unobservables determined before school entry. Figure 3.3 presents the density of RPM scores around the GATE eligibility threshold. Given that RPM scores are adjusted for the student's age at the time of testing, second graders tend to score relatively high. This explains the truncated appearance of the distribution of scores. The portion observed is the left side of a normally distributed random variable. Importantly, there do not appear to be any local discontinuities or relatively large variations in the distribution. Figure 3.4 displays the distributions of observable student characteristics which are determined before GATE testing, plotted via local mean smoothing weighted using a triangle kernel within the bin-width described above. There do not appear to be any large jumps in these variables around the eligibility threshold. Table 3.2 presents RD estimates testing for sharp differences in these covariates around the threshold. The results suggest there are no large jumps in these variables at this point. Taken together, Figures 3.2, 3.3 and 3.4 do not show any evidence of such "outside discontinuities," strongly suggesting the RD design is valid for the purposes of this paper.

Finally, differential attrition on either side of the threshold is tested for. Given that this paper follows students from testing in second grade to beyond high school, there exists some attrition of students from the sample during such a large time interval. If students on one side of the eligibility threshold studied here leave the school district at a disproportionate rate, it is suggestive of selection into (or out of) the sample along potentially unobserved determinants of educational attainment. Appendix Table C.1 uses the main specifications for the RD analyses across multiple bandwidths to test for any significant differences in attrition around the threshold. This test uses indicators that the student is observed at high school entry and again for being observed during senior year of high school as outcomes. No significant differences in attrition rates are found.

## 3.4 Results

Reduced form, first stage and fuzzy RD estimates for the impacts of GATE enrollment on secondary attainments are reported in Table 3.3. The first stage estimates are of the difference in GATE enrollment rates between students scoring just at or above the 90 point admission threshold on the RPM and those scoring just below. The change in probability of participating in GATE at this point is 0.85 and is statistically significant at the 1 percent level. There does not appear to be any effect of GATE on whether or not students persist into their senior year in high school or if they graduate from high school. The reduced form and two stage RD estimates point to a statistically insignificant difference of 2 percentage points in the likelihood of completing at least grade eleven. Similarly, these estimates for high school graduation are between 4 to 6 percentage points, but not precise enough to be statistically different from zero. This is not surprising, as students scoring near 90 on the RPM represent higher levels of academic ability than the average student who does not complete high school.

Table 3.4 presents results for the impacts of GATE on postsecondary enrollments. For the postsecondary outcomes, the cohort testing into GATE during 1998 is dropped to allow students enough time to complete a postsecondary degree, at least four years after expected high school graduation. Columns (1) and (2) model enrollment at any postsecondary institution, while (3) and (4) model enrollment at a four-year institution. Although all RD estimates in Table 3.4 statistically insignificant, they are consistently positive. The estimate magnitudes for enrollment differences between marginally ineligible and eligible students participating in GATE are about 5 (3.5) percentage points for any college enrollment (four-year college enrollment). Table 3.5 presents estimates using outcomes of postsecondary persistence. Moving pairwise left to right, the first six columns model whether the student was enrolled at a postsecondary institution during at least the first two, three and four years after high school. Columns (7) and (8) model the number of years during the first four after high school in which the student was enrolled. Similar to Table 3.4, the RD estimates are all positive and statistically insignificant. As opposed to Table 3.4, the sizes of the standard errors in Table 3.5 are generally smaller than the point estimates. The point estimates are roughly 10 percentage points for postsecondary enrollment at least two or four years after high school. That the estimate of 6 percentage points for three years is relatively smaller than the estimate for two years can be explained by the fact that some students only attend a two-year vocational or community college. Overall, the results in Table 3.5 are suggestive of a positive net effect of GATE on postsecondary persistence for students at the margin of eligibility.

Estimates of the impacts of GATE participation on postsecondary attainments are presented in Table 3.6. Corresponding Figure 3.5 plots degree completion rates by Raven's score. The results here show that students just on the margin of eligibility who join a GATE classroom are more likely to complete a postsecondary degree than their counterparts scoring just below the eligibility threshold. These GATE participators are roughly 8 percentage points more likely to complete a postsecondary degree. Comparing columns (1) and (2) with (3) and (4), it is evident that the majority of this effect is driven by an increased likelihood of four-year degree attainment. While no statistically significant effects of GATE were found for postsecondary enrollments or persistence, the estimates showing positive effects on four-year degree attainment are strongly significant. Intuitively, this agrees with the notion that since all student in the sample are of ability levels to score reasonably well on the RPM, large differences in secondary attainment levels or postsecondary enrollment rate are not likely to exist. When it comes to postsecondary persistence and completing a degree, however, it appears being tracked into gifted classrooms does hold meaningful impacts for students at the margin of GATE eligibility.

## 3.5 Robustness Checks

The main tables presented in the paper report estimates using local linear regressions within bandwidths of 4 points on either side of the 90 point threshold for inclusion in a GATE classroom. These main tables include specifications with and without student level covariates. A review of the estimates across columns show the results are not sensitive to the inclusion of these variables.

In any application of a regression discontinuity design, it is important to check whether estimates are sensitive to the choice of bandwidth used. As discussed in Section 3.2, the primary bandwidths used for the analyses are rounded to 4 points to the left and 4 point to the right of the 90 point threshold. It should be noted that scores of 95 and above indicate students who are officially GATE identified and nearly always tracked into the more advanced classrooms. For the sake of comparing impacts for students not GATE identified, but on the margin of being tracked into GATE classrooms, local linear regressions are fit within bandwidths of 4 that do not overlap with the higher GATE threshold. Appendix Table C.2 presents the significant results of the paper–the impacts of GATE participation on postsecondary attainment–using various bandwidths for estimation. Moving vertically along the columns, estimates of the impacts on postsecondary degree attainments fluctuate by at most 3 percentage points, even when capturing students scoring above the higher threshold and students scoring 6 points below the lower threshold. The reduced form, first stage and RD estimates appear fairly robust to the bandwidth choice around the lower GATE eligibility threshold score.

Finally, the impacts of GATE on postsecondary attainment are estimated using equations that are linear, quadratic and cubic in RPM scores to determine whether the results are sensitive to the functional form used for $f_j(.)$ and $g_j(.)$ in equations (3.1) and (3.2) above. Appendix Table C.3 displays these results across polynomial degrees, again using multiple bandwidths. Within each outcome and bandwidth, moving horizontally across specifications results in very little movement of the point estimate magnitudes or significance levels. Estimates using a bandwidth of 6–which includes officially GATE-identified students–exhibit more sensitivity to functional form. The cubic specification yields larger estimates than the linear specification used for the main results, by about 4 percentage points. If the true data generating process is more closely approximated by a polynomial of higher degree than one, it does not appear that using a linear specification results in large biases and, if anything, underestimates the true magnitudes of the impacts.

## 3.6 Conclusion

The likelihood of enrolling in a GATE classroom dramatically increases from zero to about 0.85 for students scoring just at the testing threshold for GATE eligibility. GATE participation for students on the eligibility margin increases four-year degree attainment rates by roughly 8 percentage points. There are no statistically significant impacts of GATE participation on high school graduation or the levels or length of postsecondary enrollment for theses students, but estimate magnitudes are suggestive of a positive impact on postsecondary persistence after high school.

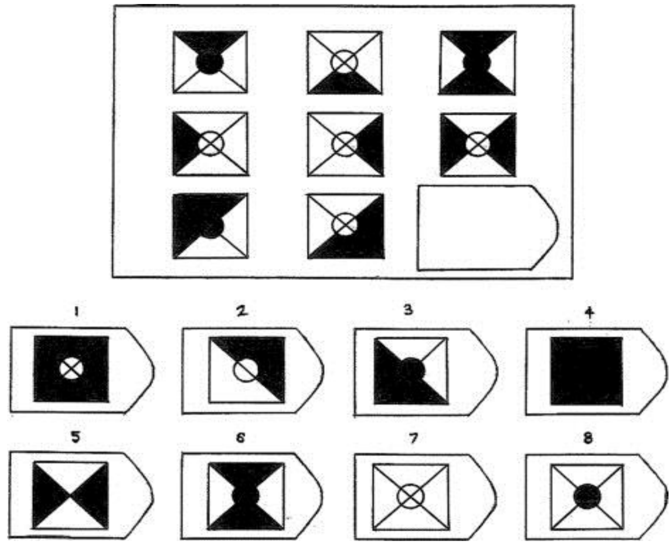
Together, these results suggest that for students performing around the RPM threshold for inclusion in a GATE classroom, being tracked into advanced classes makes little difference in terms of secondary attainment and postsecondary enrollment rates. This is in line with the previous literature on gifted education that finds no effects–or quickly fading effects–on student achievement levels. In terms of long-run educational outcomes, however, it is not until during a postsecondary education that effects appear in the form of final degree attainment. This holds important policy implications with respect to long-run educational attainment for the somewhat controversial practice of tracking students. Relative to secondary schooling, students exhibit much lower levels of persistence at the postsecondary levels. For instance, the National Center for Education Statistics (NCES) reports that among students enrolling at a postsecondary institution for the first time in Fall 2003, only 58 percent had attained a four-year degree by Spring 2009 (Radford et al., 2010).

Given that the long-term outcomes studied here have largely been absent from the literature on tracking, the results of this paper provide important considerations in tracking policies. If sorting students who score well on GATE entry tests–but below typical eligibility levels–into accelerated classrooms can increase postsecondary completion rates on the order of 8 percentage points, it suggests two things. First, the practice of high-ability tracking may have long-term benefits that justify such placements as early as the elementary years. Second, that these effects are manifested for students on the margin of inclusion in GATE classrooms suggests further expansion of ability-based tracking to include more students may be worthwhile. This is especially the case if students on this margin are less likely to complete a postsecondary education, ex ante, than the typical GATE-identified student.

Further research is needed within the framework of this paper. A planned future data collection will lengthen the number of years during which San Diego students are followed at postsecondary institutions. This will allow for the inclusion of later cohorts, for whom more detailed data on classroom experiences are available. A detailed analysis of educational differences in the secondary years resulting from GATE tracking will shed light on the mechanisms behind the postsecondary attainment effects found here.

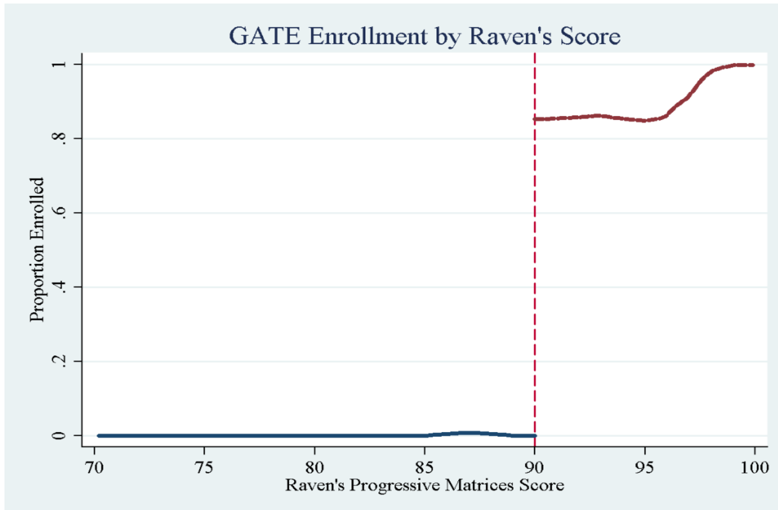
This chapter will be prepared for future submission for publication.

## 3.7 Chapter 3 Figures and Tables



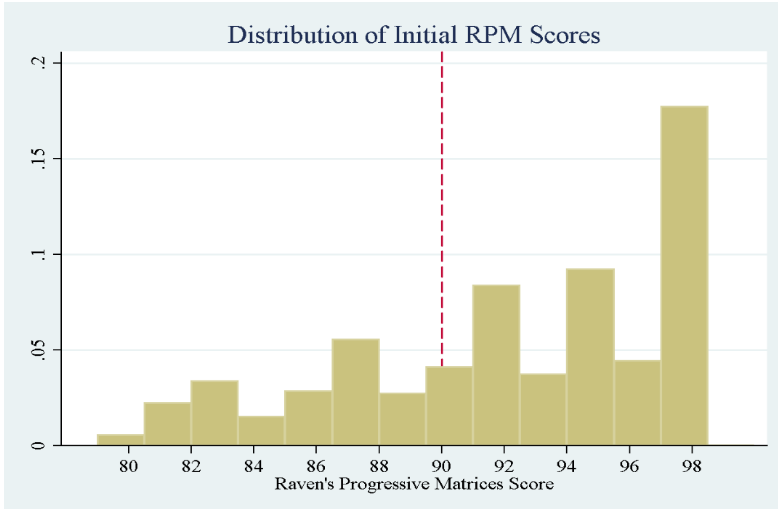

Notes: Example of a task item from the Raven's Progressive Matrices (RPM) exam. The RPM is an untimed, non-reading test administered by a child psychologist. It consists of 60 multiple choice tasks which become progressively more difficult and complex. The correct solution here is element number 4. Source: Kunda, McGreggor and Goel (2013).

**Figure3.1**: Example of Raven's Progressive Matrices Task



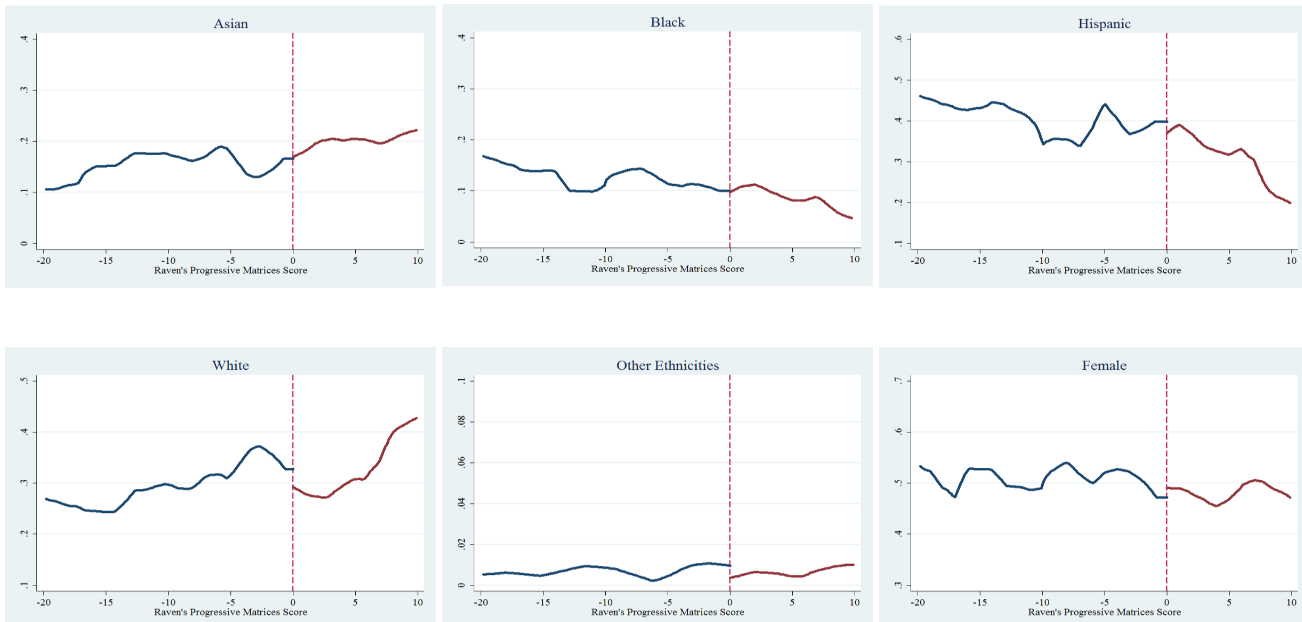

Notes: Figure presents the first stage discontinuity in treatment: the proportion of students participating in GATE, by Raven's score. Units are test scores, minus the eligibility threshold of 90.

**Figure3.2**: Proportion of GATE Enrollment by Raven's Score



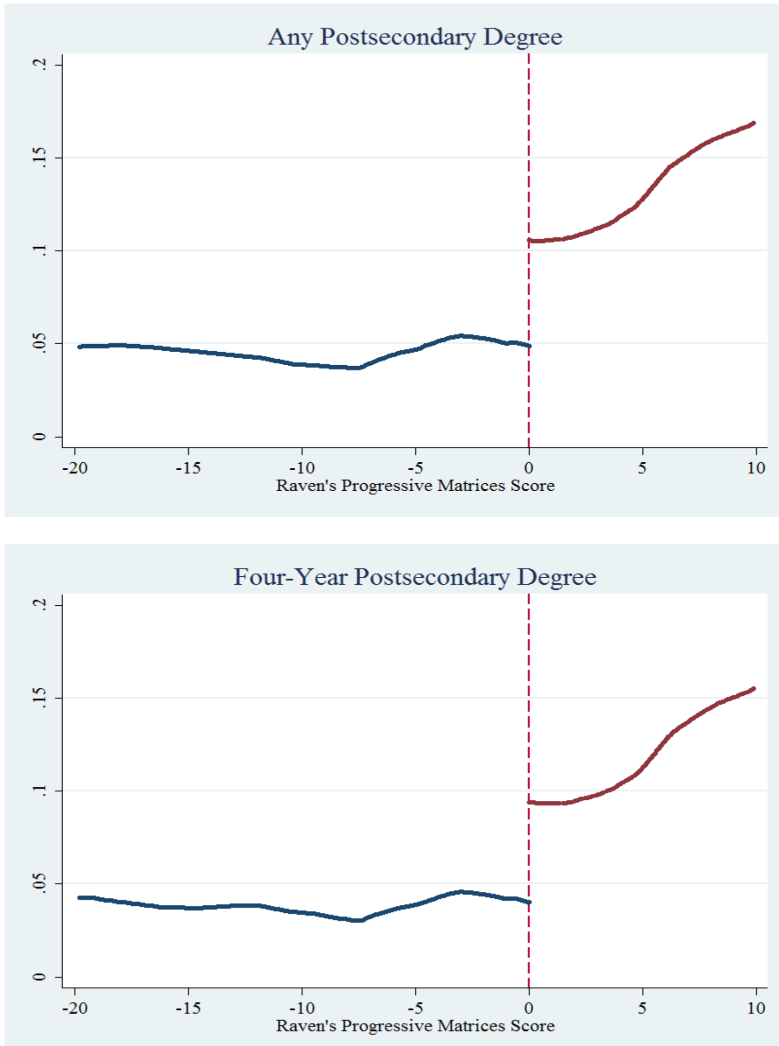

Notes: Figure is a histogram of Raven's Progressive Matrices scores among second graders taking the test for the first time in San Diego.

**Figure3.3**: Distribution of Raven’s Scores




Notes: Plotted are the distributions of student characteristics determined prior to testing, around the RPM eligibility threshold. Importantly, there does not appear to be any discontinuities or clumping in the distributions of thes variables near the GATE eligibility cutoff. Plots are of locally smoothed means weighted using a triangle kernel within the binwidths used in the analyses.

**Figure3.4**: Discontinuity Checks: Predetermined Observables




Notes: Figure plots the proportion of students who attain any postsecondary degree (top) and a four-year degree (bottom), by Raven's Progressive Matrices scores.

**Figure3.5**: Postsecondary Attainment

**Table3.1**: Summary Statistics

| | Estimation Sample | Allowing 4 Years After High School |
|---|---|---|
| | (1) | (2) |
| Female | 0.490 | 0.488 |
| | (0.500) | (0.500) |
| English Learner | 0.161 | 0.151 |
| | (0.368) | (0.358) |
| Asian | 0.192 | 0.179 |
| | (0.394) | (0.384) |
| Black | 0.086 | 0.082 |
| | (0.281) | (0.274) |
| Hispanic | 0.220 | 0.198 |
| | (0.415) | (0.399) |
| White | 0.348 | 0.336 |
| | (0.477) | (0.472) |
| Other | 0.153 | 0.205 |
| | (0.360) | (0.404) |
| Raven's Progressive Matrices Score | 92.256 | 92.49 |
| | (8.557) | (8.411) |
| GATE Cluster Paricipation | 0.612 | 0.640 |
| | (0.487) | (0.480) |
| GATE Seminar Participation | 0.072 | 0.065 |
| | (0.258) | (0.246) |
| Observations | 9,069 | 6,501 |

Notes: Sample means reported. Standard deviations in parentheses. Sample includes all students scoring at least 70 on their first Raven's Progressive Matrices test in second grade during the years of 1995 through 1998.

**Table3.2**: Tests for Discontinuities in Observables Around Eligibility Threshold

| RD Estimates of Breaks at Discontinuity | Female (1) | Asian (2) | Black (3) | Hispanic (4) | White (5) | Other (6) | English Learner (7) |
|---|---|---|---|---|---|---|---|
| Bandwidth = +/- 3 points | 0.017 | -0.008 | -0.011 | -0.069 | -0.022 | -0.012 | -0.058 |
| | (0.045) | (0.034) | (0.027) | (0.048) | (0.042) | (0.008) | (0.041) |
| | [0.703] | [0.820] | [0.666] | [0.150] | [0.597] | [0.139] | [0.155] |
| Bandwidth = +/- 4 points | 0.040 | -0.028 | 0.001 | -0.070 | 0.002 | -0.016 | -0.045 |
| | (0.062) | (0.046) | (0.037) | (0.068) | (0.058) | (0.012) | (0.058) |
| | [0.524] | [0.540] | [0.979] | [0.303] | [0.970] | [0.166] | [0.441] |
| Bandwidth = +/- 5 points | 0.046 | -0.022 | 0.005 | -0.070 | -0.005 | -0.016 | -0.036 |
| | (0.062) | (0.046) | (0.037) | (0.068) | (0.058) | (0.012) | (0.058) |
| | [0.452] | [0.631] | [0.886] | [0.300] | [0.930] | [0.172] | [0.538] |
| Bandwidth = +/- 6 points | 0.037 | 0.002 | 0.001 | -0.059 | -0.048 | -0.009 | -0.015 |
| | (0.052) | (0.039) | (0.031) | (0.057) | (0.049) | (0.010) | (0.049) |
| | [0.477] | [0.968] | [0.966] | [0.298] | [0.328] | [0.348] | [0.766] |
| Observations | 9,069 | 9,069 | 9,069 | 9,069 | 9,069 | 9,069 | 9,069 |

| | Parent Educational Attainment | | | | |
|---|---|---|---|---|---|
| | Not HS Grad (8) | HS Grad (9) | Some College (10) | College Grad (11) | Post-graduate (12) |
| Bandwidth = +/- 3 points | -0.024 | 0.001 | 0.009 | 0.018 | -0.032 |
| | (0.023) | (0.024) | (0.032) | (0.034) | (0.027) |
| | [0.295] | [0.958] | [0.785] | [0.607] | [0.225] |
| Bandwidth = +/- 4 points | -0.027 | 0.006 | 0.017 | 0.016 | -0.027 |
| | (0.033) | (0.033) | (0.043) | (0.047) | (0.038) |
| | [0.415] | [0.850] | [0.687] | [0.737] | [0.479] |
| Bandwidth = +/- 5 points | -0.024 | 0.007 | 0.015 | 0.018 | -0.026 |
| | (0.032) | (0.033) | (0.043) | (0.046) | (0.037) |
| | [0.465] | [0.823] | [0.720] | [0.695] | [0.478] |
| Bandwidth = +/- 6 points | -0.020 | 0.019 | 0.007 | 0.014 | -0.036 |
| | (0.027) | (0.028) | (0.036) | (0.039) | (0.031) |
| | [0.453] | [0.511] | [0.847] | [0.710] | [0.245] |
| Observations | 9,069 | 9,069 | 9,069 | 9,069 | 9,069 |

Notes: Bootstrapped standard errors in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$), p-values in brackets. Each column tests for discontinuities in observable characteristics around the GATE testing threshold score for enrollment eligibility. Estimates are reported for multiple bandwidths ranging from 3 to 6 points on either side of the threshold. Figure 3 presents the corresponding data plots.

**Table3.3**: GATE Participation and Secondary Attainment

| | Completed at least 11th Grade | | High School Graduation | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Reduced Form | 0.021 | 0.021 | 0.039 | 0.057 |
| | (0.060) | (0.061) | (0.062) | (0.064) |
| First Stage | 0.852*** | 0.837*** | 0.852*** | 0.837*** |
| | (0.011) | (0.013) | (0.011) | (0.013) |
| **RD Estimate** | **0.024** | **0.025** | **0.046** | **0.068** |
| | **(0.070)** | **(0.073)** | **(0.073)** | **(0.076)** |
| Observations | 9,069 | 9,069 | 9,069 | 9,069 |
| Covariates | No | Yes | No | Yes |
| | Means & Std. Deviations | | | |
| | 0.633 | | 0.552 | |
| | (0.482) | | (0.497) | |

Notes: RD estimates reported are the differences in dependent variables between students just eligible for GATE Cluster enrollment and those just under the threshold for eligibility. Local linear regressions are fit within a bandwidth of 4 points on either side of the threshold score of 90 (see text for more on bandwidth choice and robustness checks). Covariates include student age at testing date and a set of indicators for student gender, ethnicity and English learner status. Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

**Table3.4**: GATE Participation and Postsecondary Enrollment

| | Any Postsecondary Enrollment | | Enrollment at 4-year Institution | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Reduced Form | 0.065 | 0.048 | 0.063 | 0.031 |
| | (0.073) | (0.073) | (0.059) | (0.061) |
| First Stage | 0.880*** | 0.882*** | 0.880*** | 0.882*** |
| | (0.012) | (0.025) | (0.012) | (0.025) |
| **RD Estimate** | **0.074** | **0.054** | **0.071** | **0.035** |
| | **(0.082)** | **(0.083)** | **(0.067)** | **(0.069)** |
| Observations | 6,501 | 6,501 | 6,501 | 6,501 |
| Covariates | No | Yes | No | Yes |
| | Means & Std. Deviations | | | |
| | 0.455 | | 0.295 | |
| | (0.498) | | (0.456) | |

Notes: RD estimates reported are the differences in dependent variables between students just eligible for GATE Cluster enrollment and those just under the threshold for eligibility. Local linear regressions are fit within a bandwidth of 4 points on either side of the threshold score of 90 (see text for more on bandwidth choice and robustness checks). Covariates include student age at testing date and a set of indicators for student gender, ethnicity and English learner status. Second grade cohort tested during 1998 is dropped for postsecondary outcomes to allow time for degree completion. Bootstrapped standard errors in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

**Table3.5**: GATE Participation and Postsecondary Persistence

| | Enrolled at least 2 years | | Enrolled at least 3 years | | Enrolled 4 years | | Number of First 4 Years Enrolled | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Reduced Form | 0.091 | 0.090 | 0.070 | 0.055 | 0.089 | 0.081 | 0.229 | 0.225 |
| | (0.068) | (0.071) | (0.066) | (0.069) | (0.060) | (0.063) | (0.339) | (0.349) |
| First Stage | 0.880*** | 0.882*** | 0.880*** | 0.882*** | 0.880*** | 0.882*** | 0.902*** | 0.897*** |
| | (0.012) | (0.025) | (0.012) | (0.025) | (0.012) | (0.025) | (0.013) | (0.016) |
| **RD Estimate** | **0.103** | **0.102** | **0.080** | **0.062** | **0.101** | **0.092** | **0.254** | **0.251** |
| | **(0.078)** | **(0.081)** | **(0.075)** | **(0.078)** | **(0.068)** | **(0.071)** | **(0.376)** | **(0.389)** |
| Observations | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 |
| Covariates | No | Yes | No | Yes | No | Yes | No | Yes |
| | Means & Std. Deviations | | | | | | | |
| | 0.400 | | 0.357 | | 0.301 | | 2.181 | |
| | (0.490) | | (0.479) | | (0.459) | | (1.808) | |

Notes: RD estimates reported are the differences in dependent variables between students just eligible for GATE Cluster enrollment and those just under the threshold for eligibility. Local linear regressions are fit within a bandwidth of 4

**Table3.6**: GATE Participation and Postsecondary Attainment

| | Any Postsecondary Degree Attainment | | Four-Year Degree Attainement | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Reduced Form | 0.093*** | 0.073** | 0.091*** | 0.071** |
| | (0.031) | (0.032) | (0.028) | (0.030) |
| First Stage | 0.880*** | 0.882*** | 0.880*** | 0.882*** |
| | (0.012) | (0.025) | (0.012) | (0.025) |
| **RD Estimate** | **0.105*** ** | **0.083** ** | **0.104*** ** | **0.080** ** |
| | **(0.035)** | **(0.037)** | **(0.032)** | **(0.034)** |
| Observations | 6,501 | 6,501 | 6,501 | 6,501 |
| Covariates | No | Yes | No | Yes |
| | Means & Std. Deviations | | | |
| | 0.264 | | 0.245 | |
| | (0.441) | | (0.430) | |

Notes: RD estimates reported are the differences in dependent variables between students just eligible for GATE Cluster enrollment and those just under the threshold for eligibility. Local linear regressions are fit within a bandwidth of 4 points on either side of the threshold score of 90 (see text for more on bandwidth choice and robustness checks). Covariates include student age at testing date and a set of indicators for student gender, ethnicity and English learner status. Bootstrapped standard errors in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

**Table3.7**: Summary of RD Estimates for All Postsecondary Outcomes

| | Any Postsecondary Enrollment | Enrollment: Four-Year Institution | Enrolled 2/4 Years After High School | Enrolled 3/4 Years After High School | Enrolled 4/4 Years After High School | Any Postsecondary Degree | Four-Year Degree |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Reduced Form | 0.048 | 0.031 | 0.090 | 0.055 | 0.081 | 0.073** | 0.071** |
| | (0.073) | (0.061) | (0.071) | (0.069) | (0.063) | (0.032) | (0.030) |
| First Stage | 0.882*** | 0.882*** | 0.882*** | 0.882*** | 0.882*** | 0.882*** | 0.882*** |
| | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) |
| **RD Estimate** | **0.054** | **0.035** | **0.102** | **0.062** | **0.092** | **0.083**** | **0.080**** |
| | **(0.083)** | **(0.069)** | **(0.081)** | **(0.078)** | **(0.071)** | **(0.037)** | **(0.034)** |
| Observations | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 |

Notes: RD estimates reported are the differences in dependent variables between students just eligible for GATE Cluster enrollment and those just under the threshold for eligibility. Local linear regressions are fit within a bandwidth of 4 points on either side of the threshold score of 90 (see text for more on bandwidth choice and robustness checks). Covariates include student age at testing date and a set of indicators for student gender, ethnicity and English learner status. Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

# Appendix A

# Chapter 1 Supplemental Materials

**TableA.1**: Months to BIC Implementation

| | Months to BIC Implementation |
|---|---|
| % Students Free/Reduced Price Eligible | -0.166 |
| | [0.119] |
| CST Math Gains | -0.358 |
| | [0.578] |
| CST ELA Gains | 0.0173 |
| | [0.287] |
| % Day Absent | -0.222 |
| | [0.226] |
| % English learner students | -0.206** |
| | [0.0807] |
| Years of Teacher Experience | 0.0204 |
| | [0.0433] |
| Female | -0.06 |
| | [0.161] |
| White | 0.545 |
| | [0.766] |
| Black | 0.357 |
| | [0.431] |
| Other | 1.508 |
| | [0.904] |
| Observations | 5,666 |
| R-squared | 0.366 |
| F-test | 0.0107 |
| F-test: Outcomes and %free/reduced | 0.2327 |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in brackets are clustered at the school level. Dependent variable is the number of months from when the first BIC program implementations began to the BIC implementation date in the student's school. Independent variables are observations one year before the BIC rollout began.

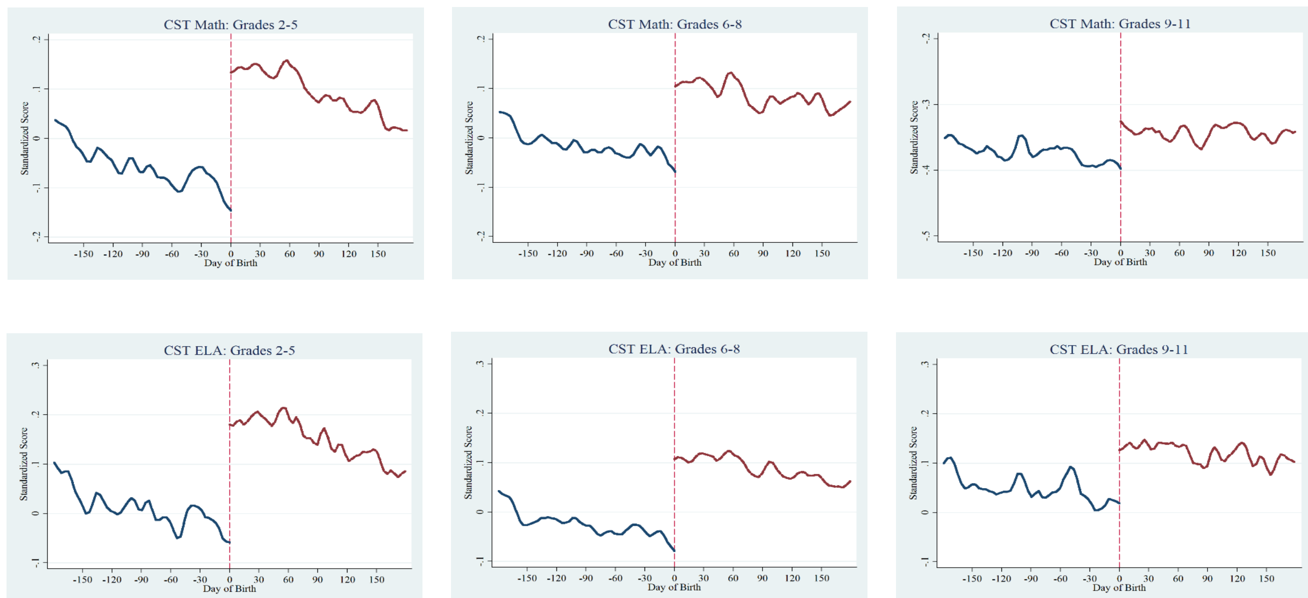
**TableA.2**: Lead-Lagged Estimates of Student Gains: Breakfast Consumption

| | Math Gains | | | ELA Gains | | |
|---|---|---|---|---|---|---|
| *BIC(t-3) * % change-in-meals* | 0.021 | 0.006 | | 0.014 | 0.012 | |
| | (0.024) | (0.024) | | (0.013) | (0.015) | |
| *BIC(t-3)*change + BIC(t-3)*change*UFB* | | 0.027 | | | 0.014 | |
| | | (0.031) | | | (0.017) | |
| *BIC(t-2) * % change-in-meals* | 0.009 | -0.024 | | 0.005 | 0.000 | |
| | (0.032) | (0.048) | | (0.018) | (0.023) | |
| *BIC(t-2)*change + BIC(t-2)*change*UFB* | | 0.014 | | | 0.004 | |
| | | (0.032) | | | (0.019) | |
| *BIC(t-1) * % change-in-meals* | 0.000 | 0.046 | | 0.019 | 0.031 | |
| | (0.047) | (0.065) | | (0.026) | (0.030) | |
| *BIC(t-1)*change + BIC(t-1)*change*UFB* | | -0.028 | | | 0.009 | |
| | | (0.048) | | | (0.026) | |
| *BIC Year 1 * % change-in-meals* | 0.050 | 0.134* | 0.096*** | 0.048 | 0.094** | 0.064** |
| | (0.075) | (0.077) | (0.033) | (0.041) | (0.039) | (0.024) |
| *BIC Year 1*change + BIC Year 1* change*UFB* | | -0.006 | 0.021 | | 0.013 | 0 |
| | | (0.087) | (0.055) | | (0.050) | (0.038) |
| *BIC Year 2-3 * % change-in-meals* | 0.090 | 0.192** | 0.152*** | 0.098* | 0.142*** | 0.110*** |
| | (0.097) | (0.092) | (0.038) | (0.056) | (0.048) | (0.029) |
| *BIC Year 2-3*change + BIC Year 2-3* change*UFB* | | 0.039 | 0.064 | | 0.067 | 0.053 |
| | | (0.108) | (0.072) | | (0.064) | (0.050) |
| Observations | 67,157 | 67,157 | 67,157 | 66,461 | 66,461 | 66,461 |
| F-test p-values | | | | | | |
| BIC(t-3)=BIC(t-2)=BIC(t-1)=0 | 0.676 | | | 0.497 | | |
| BIC(t)=BIC(t+1,2)=0 | 0.445 | | | 0.080 | | |
| All BIC(t) = 0 | 0.000 | 0.003 | | 0.089 | 0.006 | |
| BIC(t)=BIC(t+1,2)=BIC*UFB=0 | | 0.115 | 0.005 | | 0.006 | 0.004 |
| All BIC(t) and BIC(t)*UFB = 0 | | 0.000 | | | 0.006 | |

Notes: *, **, *** denote significance at the 10%, 5% and 1% levels respectively. Standard errors in parentheses are clustered at the school level. Dependent variables are student gains in math and English language arts CST scores. Scores are normalized using statewide mean and variance for each grade and year. Each model includes school, grade, year and student fixed effects, indicators for school type and year-round status,a set of student level covariates, teacher ethnicity, years of experience and certification level, and school principal fixed effects. Reported coefficients are for lead and lagged years of treatment, separately by pre-BIC universally free meal offerings. Units are standard deviation changes per 100 percent increase in meals served.

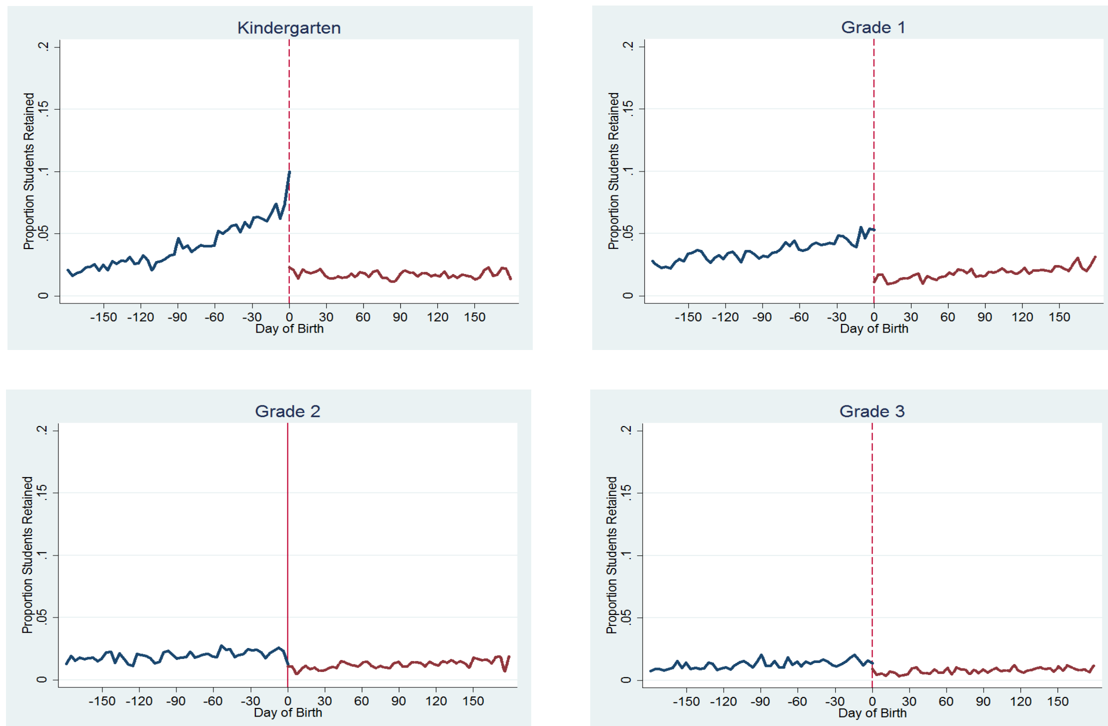
# Appendix B

# Chapter 2 Supplemental Materials



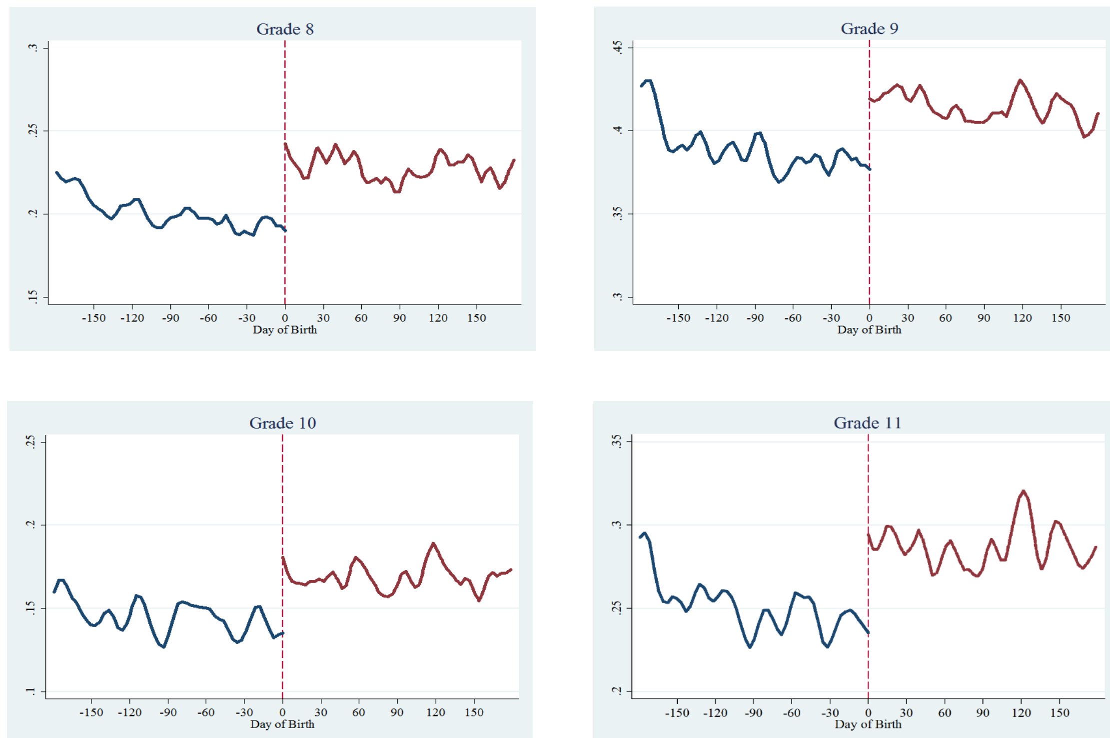

Notes: Figures present local polynomial plots of standardized CST Math and ELA scores by day of birth, separated by elementary, middle and high school grade levels. Scores are standardized using statewide mean and standard deviation for each academic year and grade level. Local linear regressions using these groupings include interaction indicators for grade level and test version as well as the standard set of covariates (see text). Standard errors from these specifications are clustered at the student level to account for repeat student observations across grade levels.

**FigureB.1**: Day of Birth and Student Achievement by Grade



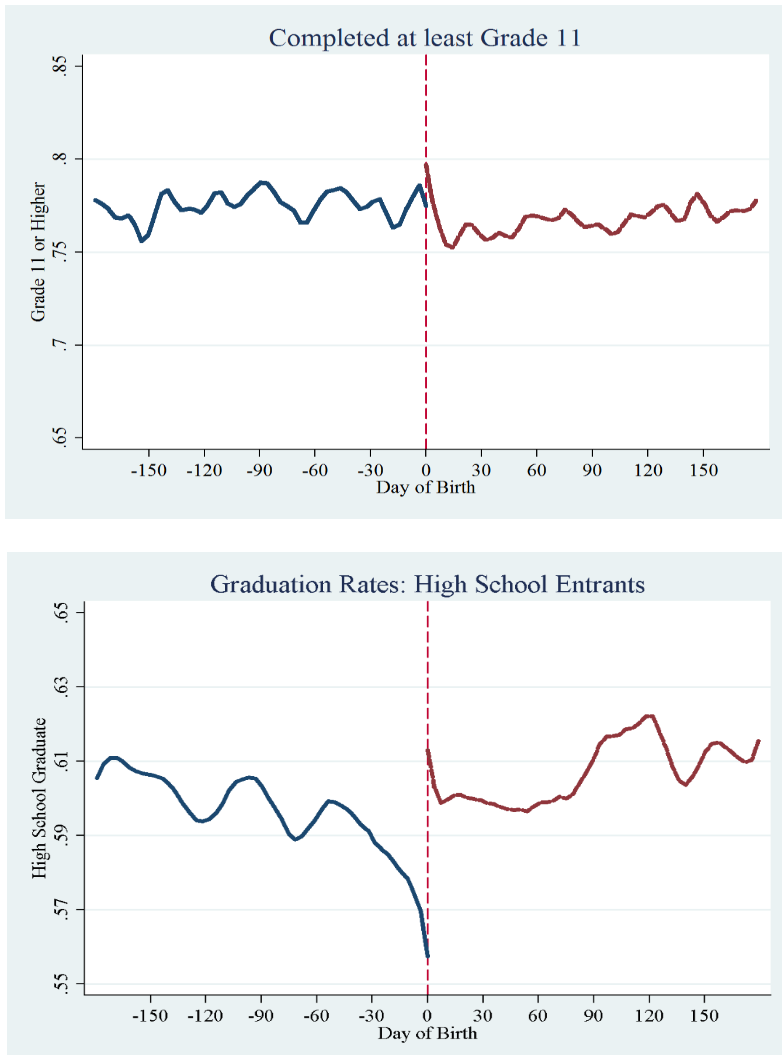

Notes: Figure plots the incedence of grade retentions during Kindergarten through grade 3, by day of birth. The youngest of entrants experience disproptionately higher rates of retention than their cohort peers in Kindergarten and first grade.

**FigureB.2**: Day of Birth and Grade Retentions



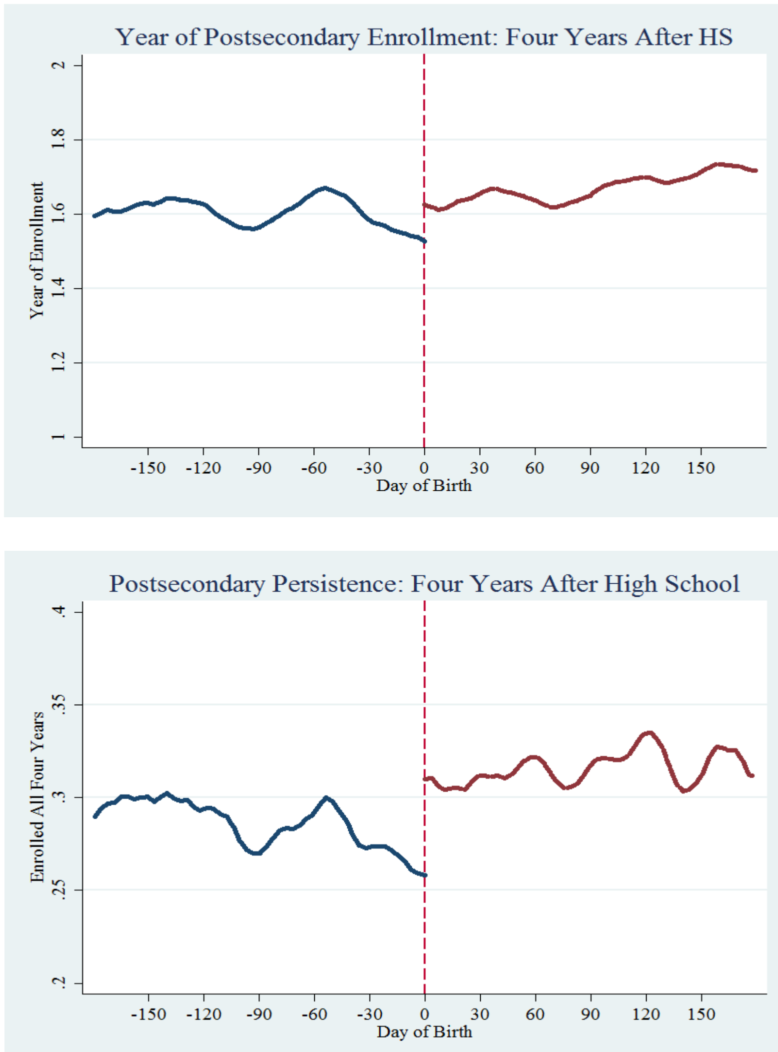

Notes: Local polynomial plots of the proportion of students above the normal math course progression level for each of grades 8 through 11. Course progressions are defined according to presribed courses for each grade level under the California state academic standards for mathematics.

**FigureB.3**: Math Enrollment Above Grade-Prescribed Course Level



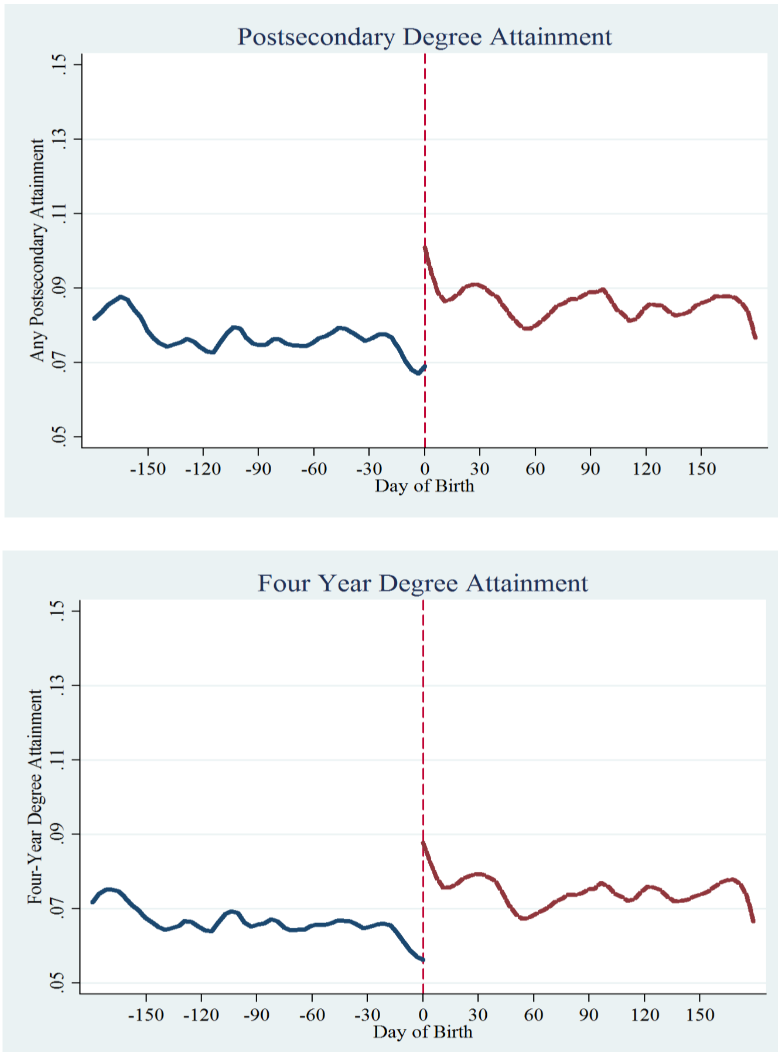

Notes: Figure plots the proportion of observed high school entrants matriculating to grade 12 (top) and graduating high school (bottom) respectively, by day of birth.

**FigureB.4**: Secondary Attainment




Notes: Figure plots the number of years during the first four after expected high school graduation during which students are enrolled in any postsecondary institution (top) and the proportion of high school students who are enrolled in a postsecondary institution during all four years after high school (bottom).

**FigureB.5**: Postsecondary Persistence




Notes: Figure plots the proportion of students who attain any postsecondary degree (top) and a four-year degree (bottom), by day of birth.

**FigureB.6**: Postsecondary Attainment

**TableB.1**: Cohort Sizes by Kindergarten Entry Year

| | Grade Level | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Kindergarten Entry Year | K (entry) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1992 | 13,216 | 11,747 | 10,316 | 9,374 | 8,711 | 8,219 | 7,825 | 7,448 | 7,073 | 6,898 | 6,392 | 5,823 | 5,371 |
| 1993 | 13,072 | 11,455 | 10,079 | 9,178 | 8,621 | 8,148 | 7,749 | 7,371 | 6,890 | 6,776 | 6,230 | 5,496 | 5,082 |
| 1994 | 13,664 | 11,987 | 10,563 | 9,690 | 9,074 | 8,597 | 8,254 | 7,728 | 7,311 | 7,154 | 6,465 | 5,553 | 5,429 |
| 1995 | 13,862 | 12,011 | 10,785 | 9,907 | 9,332 | 8,867 | 8,416 | 7,912 | 7,454 | 7,260 | 6,450 | 5,744 | 5,621 |
| 1996 | 13,814 | 12,114 | 10,851 | 10,081 | 9,472 | 8,891 | 8,367 | 7,922 | 7,406 | 6,860 | 6,443 | 5,828 | 5,606 |
| 1997 | 13,743 | 11,932 | 10,707 | 9,983 | 9,255 | 8,738 | 8,272 | 7,683 | 7,021 | 6,767 | 6,519 | 5,904 | 5,598 |
| 1998 | 13,488 | 11,943 | 10,806 | 9,909 | 9,248 | 8,729 | 8,194 | 7,343 | 7,041 | 6,910 | 6,549 | 5,929 | 4,971 |
| 1999 | 13,912 | 12,387 | 10,973 | 10,135 | 9,425 | 8,717 | 7,780 | 7,463 | 7,271 | 7,140 | 6,636 | 5,328 | |
| 2000 | 13,201 | 11,674 | 10,441 | 9,532 | 8,733 | 7,937 | 7,510 | 7,252 | 7,068 | 6,907 | 5,482 | | |
| 2001 | 12,724 | 11,229 | 9,947 | 8,929 | 8,008 | 7,578 | 7,257 | 7,105 | 6,860 | 5,848 | | | |
| 2002 | 12,098 | 10,704 | 9,329 | 8,136 | 7,652 | 7,379 | 7,142 | 6,899 | 6,074 | | | | |
| 2003 | 11,966 | 10,443 | 8,763 | 8,095 | 7,747 | 7,509 | 7,286 | 6,389 | | | | | |
| 2004 | 11,309 | 9,329 | 8,284 | 7,801 | 7,463 | 7,175 | 6,205 | | | | | | |
| 2005 | 11,333 | 9,664 | 8,847 | 8,321 | 7,795 | 6,773 | | | | | | | |
| 2006 | 10,793 | 9,421 | 8,651 | 7,983 | 6,773 | | | | | | | | |
| 2007 | 11,120 | 9,776 | 8,710 | 7,304 | | | | | | | | | |
| 2008 | 11,556 | 9,873 | 8,239 | | | | | | | | | | |
| 2009 | 11,185 | 9,134 | | | | | | | | | | | |
| 2010 | 11,072 | | | | | | | | | | | | |
| Number of Students by Grade Level | 237,128 | 196,827 | 166,307 | 144,374 | 127,338 | 113,294 | 100,307 | 88,550 | 77,515 | 68,593 | 57,306 | 45,808 | 37,906 |

Notes: Tabulation of San Diego students observed during first matriculation into each grade level, by Kindergarten entry year. Actual Kindergarten enrollment records were not available prior to 1992. Black boxed areas correspond to grade levels and entry years for which CST test scores are recorded in the data. The red, dotted box highlights the entry-year cohorts for which postsecondary outcomes are observable.

**TableB.2**: Tests for Differential Attrition Around Cutoff Date

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Reduced Form | 0.005 | 0.001 | -0.035 | -0.016 | -0.017 |
| | (0.047) | (0.079) | (0.050) | (0.035) | (0.024) |
| First Stage | 0.823*** | 0.872*** | 0.838*** | 0.806*** | 0.799*** |
| | (0.026) | (0.051) | (0.031) | (0.021) | (0.014) |
| **RD Estimate** | **0.006** | **0.002** | **-0.042** | **-0.019** | **-0.021** |
| | **(0.058)** | **(0.090)** | **(0.060)** | **(0.043)** | **(0.030)** |
| Observations | 82,116 | 82,116 | 82,116 | 82,116 | 82,116 |
| Bandwidth Used | $\frac{1}{2}h_{IK}$ | $h_{IK}$ | $2h_{IK}$ | $4h_{IK}$ | $8h_{IK}$ |
| Calculated Bandwidth | 1.153 | 2.306 | 4.612 | 9.224 | 18.448 |

Notes: The outcome variable is an indicator for the student, observed at Kindergarten entry, not being observed in the secondary and postsecondary outcomes sample. Bootstrapped standard errors in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

**TableB.3**: Student Achievement: Elementary, Middle and High School

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | CST ELA Scores | | | | | |
| Estimator | Grades 2-5 | | Grades 6-8 | | Grades 9-11 | |
| Reduced Form | 0.342*** | 0.320*** | 0.142*** | 0.165*** | 0.026 | 0.067 |
| | (0.053) | (0.068) | (0.052) | (0.045) | (0.058) | (0.046) |
| First Stage | 0.816*** | | 0.849*** | | 0.843*** | |
| | (0.018) | | (0.016) | | (0.024) | |
| **RD Estimate** | **0.419*** ** | | **0.168*** ** | | **0.031** | |
| | **(0.065)** | | **(0.061)** | | **(0.069)** | |
| Observations | 310,968 | 399,002 | 201,806 | 290,906 | 150,278 | 259,698 |
| | CST Math Scores | | | | | |
| | Grades 2-5 | | Grades 6-8 | | Grades 9-11 | |
| Reduced Form | 0.362*** | 0.344*** | 0.272*** | 0.245*** | 0.181*** | 0.130*** |
| | (0.048) | (0.044) | (0.064) | (0.058) | (0.052) | (0.047) |
| First Stage | 0.818*** | | 0.866*** | | 0.830*** | |
| | (0.016) | | (0.019) | | (0.020) | |
| **RD Estimate** | **0.443*** ** | | **0.314*** ** | | **0.218*** ** | |
| | **(0.058)** | | **(0.074)** | | **(0.063)** | |
| Observations | 312,264 | 400,967 | 201,482 | 290,512 | 145,841 | 250,999 |

Notes: CST scores are standardized by grade level and year using statewide means and standard deviations. RD estimates reported are the differences between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Regressions include indicators for grade level, student gender, ethnicity, English learner status at school entry and parental education levels. Bootstrapped standard errors, clustered at the student level, are in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

**TableB.4**: RD Estimates of Achievement Gains by Grade

| | Grade 3 | Grade 4 | Grade 5 | Grade 6 | Grade 7 | Grade 8 | Grade 9 | Grade 10 | Grade 11 |
|---|---|---|---|---|---|---|---|---|---|
| Estimator | | | | | | | | | |
| Reduced Form | -0.091 | -0.061 | -0.090 | 0.010 | -0.115 | 0.117** | -0.088 | -0.036 | -0.157* |
| | (0.068) | (0.059) | (0.071) | (0.082) | (0.071) | (0.060) | (0.069) | (0.070) | (0.089) |
| First Stage | 0.795*** | 0.810*** | 0.811*** | 0.864*** | 0.848*** | 0.869*** | 0.846*** | 0.806*** | 0.849*** |
| | (0.033) | (0.031) | (0.036) | (0.042) | (0.035) | (0.031) | (0.038) | (0.040) | (0.054) |
| **RD Estimate** | **-0.114** | **-0.076** | **-0.111** | **0.011** | **-0.135** | **0.135**** | **-0.104** | **-0.045** | **-0.185*** |
| | **(0.086)** | **(0.073)** | **(0.088)** | **(0.095)** | **(0.083)** | **(0.069)** | **(0.082)** | **(0.087)** | **(0.103)** |
| Observations | 67,759 | 65,016 | 62,978 | 60,571 | 57,753 | 55,340 | 51,085 | 47,409 | 37,865 |
| Estimator | | | | | | | | | |
| Reduced Form | -0.077 | -0.129 | -0.010 | 0.017 | -0.010 | 0.142* | -0.005 | 0.053 | -0.121 |
| | (0.086) | (0.079) | (0.068) | (0.064) | (0.062) | (0.083) | (0.090) | (0.083) | (0.086) |
| First Stage | 0.805*** | 0.811*** | 0.812*** | 0.836*** | 0.840*** | 0.885*** | 0.838*** | 0.830*** | 0.819*** |
| | (0.040) | (0.016) | (0.033) | (0.030) | (0.031) | (0.034) | (0.034) | (0.036) | (0.038) |
| **RD Estimate** | **-0.096** | **-0.159** | **-0.013** | **0.020** | **-0.012** | **0.160*** | **-0.006** | **0.064** | **-0.148** |
| | **(0.107)** | **(0.097)** | **(0.084)** | **(0.076)** | **(0.074)** | **(0.093)** | **(0.107)** | **(0.101)** | **(0.106)** |
| Observations | 67,903 | 65,611 | 63,682 | 60,520 | 57,892 | 55,171 | 49,806 | 45,683 | 35,676 |

| | CST ELA Gains | | | CST Math Gains | | |
|---|---|---|---|---|---|---|
| Estimator | Grades 3-5 | Grades 6-8 | Grades 9-11 | Grades 3-5 | Grades 6-8 | Grades 9-11 |
| Reduced Form | 0.073 | 0.029 | -0.083* | 0.097** | 0.055 | -0.014 |
| | (0.046) | (0.041) | (0.043) | (0.045) | (0.053) | (0.054) |
| First Stage | 0.811*** | 0.862*** | 0.841*** | 0.815*** | 0.869*** | 0.838*** |
| | (0.018) | (0.020) | (0.024) | (0.017) | (0.023) | (0.022) |
| **RD Estimate** | **0.090** | **0.034** | **-0.098*** | **0.119**** | **0.063** | **-0.017** |
| | **(0.057)** | **(0.047)** | **(0.051)** | **(0.055)** | **(0.061)** | **(0.064)** |
| Observations | 278,301 | 173,664 | 136,359 | 280,020 | 173,583 | 131,165 |

Notes: Bootstrapped standard errors, clustered at the student level, are in parentheses (*** p<0.01, ** p<0.05, * p<0.1). CST scores are standardized by grade level and year using statewide moments. One year differences are then taken for gains. Local linear regressions use bandwidths calculated using the IK algorithm. Indicators for grade level, student gender, ethnicity, English learner status at school entry and parental education levels are included.

**TableB.5**: RD Estimates of Student Behavior by Grade

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Kindergarten | Grade1 | Grade 2 | Grade 3 | Grade 4 | Grade 5 |
| Estimator | Shows Interest | | | | | |
| Reduced Form | 0.347* | 0.321** | 0.225 | 0.293** | 0.503*** | 0.287* |
| | (0.184) | (0.138) | (0.138) | (0.137) | (0.165) | (0.153) |
| First Stage | 0.807*** | 0.802*** | 0.807*** | 0.799*** | 0.819*** | 0.844*** |
| | (0.055) | (0.048) | (0.046) | (0.046) | (0.052) | (0.042) |
| **RD Estimate** | **0.429*** | **0.400**** | **0.279** | **0.367**** | **0.615***** | **0.340*** |
| | **(0.223)** | **(0.173)** | **(0.172)** | **(0.174)** | **(0.205)** | **(0.179)** |
| Observations | 34,614 | 30,441 | 27,027 | 23,674 | 22,243 | 21,860 |
| | Completes Assignments | | | | | |
| Reduced Form | 0.408** | 0.353** | 0.377** | 0.256* | 0.320** | 0.252 |
| | (0.161) | (0.168) | (0.149) | (0.133) | (0.163) | (0.178) |
| First Stage | 0.820*** | 0.787*** | 0.809*** | 0.799*** | 0.811*** | 0.844*** |
| | (0.050) | (0.056) | (0.048) | (0.046) | (0.056) | (0.050) |
| **RD Estimate** | **0.498**** | **0.448**** | **0.466**** | **0.320*** | **0.395**** | **0.299** |
| | **(0.194)** | **(0.212)** | **(0.186)** | **(0.166)** | **(0.201)** | **(0.209)** |
| Observations | 34,606 | 30,421 | 27,021 | 23,671 | 22,238 | 21,860 |
| | Respectfulness | | | | | |
| Reduced Form | -0.065 | 0.258** | 0.273* | 0.151 | 0.375** | 0.058 |
| | (0.163) | (0.130) | (0.141) | (0.137) | (0.183) | (0.141) |
| First Stage | 0.793*** | 0.798*** | 0.807*** | 0.799*** | 0.812*** | 0.845*** |
| | (0.056) | (0.041) | (0.046) | (0.046) | (0.055) | (0.043) |
| **RD Estimate** | **-0.082** | **0.323**** | **0.338*** | **0.189** | **0.462**** | **0.069** |
| | **(0.205)** | **(0.164)** | **(0.178)** | **(0.173)** | **(0.225)** | **(0.166)** |
| Observations | 34,613 | 30,414 | 27,014 | 23,672 | 22,245 | 21,865 |
| | Preparedness | | | | | |
| Reduced Form | 0.397*** | 0.318*** | 0.401*** | 0.220 | 0.252 | 0.119 |
| | (0.141) | (0.121) | (0.131) | (0.138) | (0.165) | (0.146) |
| First Stage | 0.843*** | 0.796*** | 0.805*** | 0.798*** | 0.813*** | 0.843*** |
| | (0.038) | (0.040) | (0.044) | (0.047) | (0.055) | (0.042) |
| **RD Estimate** | **0.471***** | **0.400***** | **0.498***** | **0.275** | **0.310** | **0.142** |
| | **(0.168)** | **(0.153)** | **(0.163)** | **(0.172)** | **(0.203)** | **(0.172)** |
| Observations | 34,612 | 30,429 | 27,026 | 23,672 | 22,241 | 21,862 |

Notes: Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Behavior evaluations are standardized by grade level. RD estimates reported are the differences in dependent variables between students born just before the school entry cutoff date and those induced to delay school entry by having birthdates just after December 2. The employed fuzzy regression discontinuity design uses local linear regressions with bandwidths calculated using the IK algorithm. Unless stated otherwise, all regressions include indicators for student gender, ethnicity, English learner status at school entry and parental education levels.

**TableB.6**: Student Behavior: Bandwidth and Functional Form Checks

| | $Bandwidth\ \ h = 2h_{ik}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Estimator | Begins Promptly | | Follows Directions | | Classroom Behavior | | Self Discipline | |
| Reduced Form | 0.324*** | 0.319*** | 0.311*** | 0.249** | 0.283*** | 0.260** | 0.253*** | 0.192* |
| | (0.078) | (0.118) | (0.077) | (0.116) | (0.076) | (0.116) | (0.077) | (0.116) |
| First Stage | 0.804*** | 0.801*** | 0.804*** | 0.801*** | 0.804*** | 0.801*** | 0.803*** | 0.801*** |
| | (0.015) | (0.020) | (0.015) | (0.020) | (0.015) | (0.020) | (0.015) | (0.020) |
| **RD Estimate** | **0.403*** ** | **0.399*** ** | **0.387*** ** | **0.311** ** | **0.352*** ** | **0.325** ** | **0.315*** ** | **0.240* ** |
| | **(0.096)** | **(0.146)** | **(0.096)** | **(0.144)** | **(0.095)** | **(0.144)** | **(0.096)** | **(0.145)** |
| | $Bandwidth\ \ h = h_{ik}$ | | | | | | | |
| | Begins Promptly | | Follows Directions | | Classroom Behavior | | Self Discipline | |
| Reduced Form | 0.372*** | 0.367*** | 0.305** | 0.341*** | 0.299** | 0.313*** | 0.260** | 0.307*** |
| | (0.132) | (0.093) | (0.127) | (0.091) | (0.130) | (0.090) | (0.128) | (0.091) |
| First Stage | 0.802*** | 0.823*** | 0.803*** | 0.823*** | 0.802*** | 0.823*** | 0.802*** | 0.823*** |
| | (0.022) | (0.032) | (0.022) | (0.032) | (0.022) | (0.032) | (0.022) | (0.032) |
| **RD Estimate** | **0.463*** ** | **0.446*** ** | **0.379** ** | **0.414*** ** | **0.373** ** | **0.381*** ** | **0.324** ** | **0.373*** ** |
| | **(0.164)** | **(0.112)** | **(0.158)** | **(0.110)** | **(0.161)** | **(0.109)** | **(0.159)** | **(0.109)** |
| Observations | 239,201 | 239,201 | 239,219 | 239,219 | 239,194 | 239,194 | 238,989 | 238,989 |
| IK Bandwidth $h_{ik}$ | 2.67 | 2.67 | 2.59 | 2.59 | 2.68 | 2.68 | 2.60 | 2.60 |
| Functional Form | Linear | Cubic | Linear | Cubic | Linear | Cubic | Linear | Cubic |

Notes: Bootstrapped standard errors, clustered at the student level, are in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Behavior evaluations are standardized by grade level. RD estimates are reported for both local linear regressions that are polynomials of degree one and three, respectively, in distance of birth date from the cutoff. Likewise, these sets of estimates are also reported using bandwidths equal to one and two time the calculated IK optimal bandwidth.

**TableB.7**: Postsecondary Attainment: Bandwidth and Functional Form Checks

| | Enrollment at 4-year Institution: 1st year after HS | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Estimator | $Bandwidth\ \ h = 2h_{ik}$ | | | |
| Reduced Form | 0.106* | 0.104** | 0.156* | 0.197*** |
| | (0.058) | (0.046) | (0.085) | (0.073) |
| First Stage | 0.839*** | | 0.870*** | |
| | (0.044) | | (0.069) | |
| **RD Estimate** | **0.126*** | | **0.179*** | |
| | **(0.070)** | | **(0.100)** | |
| Estimator | $Bandwidth\ \ h = h_{ik}$ | | | |
| Reduced Form | 0.141 | 0.191*** | 0.138* | 0.141*** |
| | (0.093) | (0.070) | (0.073) | (0.054) |
| First Stage | 0.884*** | | 0.835*** | |
| | (0.079) | | (0.053) | |
| **RD Estimate** | **0.159** | | **0.165*** | |
| | **(0.107)** | | **(0.088)** | |
| Observations | 36,881 | 84,361 | 36,881 | 84,361 |
| IK Bandwidth $h_{ik}$ | 2.582 | 2.209 | 2.582 | 2.209 |
| Functional Form | Linear | | Cubic | |

Notes: RD estimates are reported for both local linear regressions that are polynomials of degree one and three, respectively, in distance of birth date from the cutoff. Likewise, these sets of estimates are also reported using bandwidths equal to one and two time the calculated IK optimal bandwidth. Bootstrapped standard errors are in parentheses. (*** p<0.01, ** p<0.05, * p<0.1).

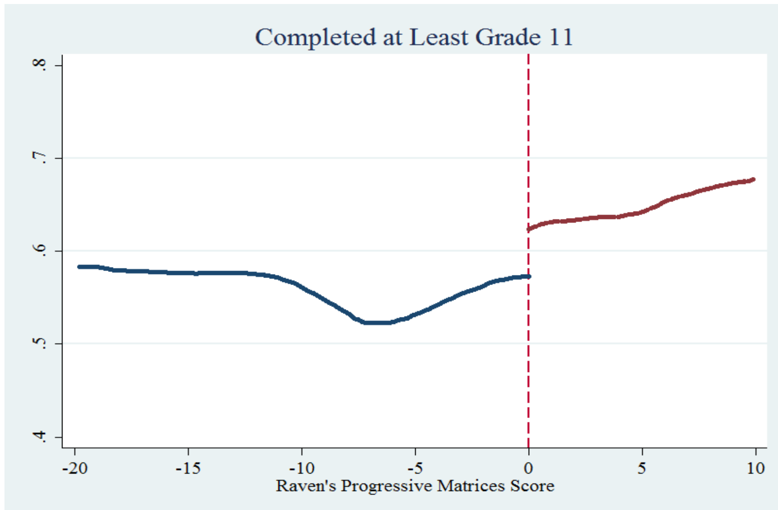
**TableB.8**: Postsecondary Persistence: Bandwidth and Functional Form Checks

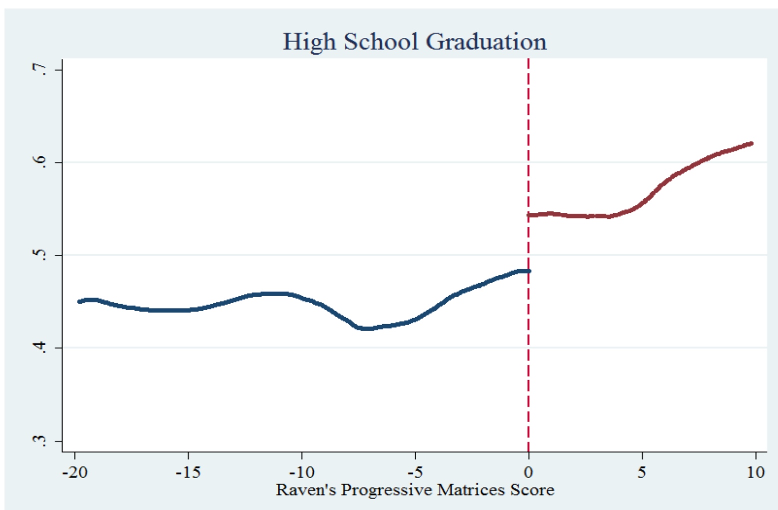
| | Enrolled All 4 years After High School | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Estimator | *Bandwidth* $h = 2h_{ik}$ | | | |
| Reduced Form | 0.158* | 0.154*** | 0.225 | 0.158* |
| | (0.095) | (0.056) | (0.146) | (0.091) |
| First Stage | 0.771*** | | 0.775*** | |
| | (0.071) | | (0.125) | |
| **RD Estimate** | **0.205*** | | **0.290** | |
| | **(0.124)** | | **(0.195)** | |
| | *Bandwidth* $h = h_{ik}$ | | | |
| Reduced Form | 0.276* | 0.180** | 0.291** | 0.179*** |
| | (0.144) | (0.084) | (0.115) | (0.062) |
| First Stage | 0.801*** | | 0.806*** | |
| | (0.118) | | (0.071) | |
| **RD Estimate** | **0.344*** | | **0.361**** | |
| | **(0.190)** | | **(0.145)** | |
| Observations | 18,111 | 51,824 | 18,111 | 51,824 |
| IK Bandwidth $h_{ik}$ | 2.307 | 2.169 | 2.307 | 2.169 |
| Functional Form | Linear | | Cubic | |

Notes: RD estimates are reported for both local linear regressions that are polynomials of degree one and three, respectively, in distance of birth date from the cutoff. Likewise, these sets of estimates are also reported using bandwidths equal to one and two time the calculated IK optimal bandwidth. Bootstrapped standard errors are in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

# Appendix C

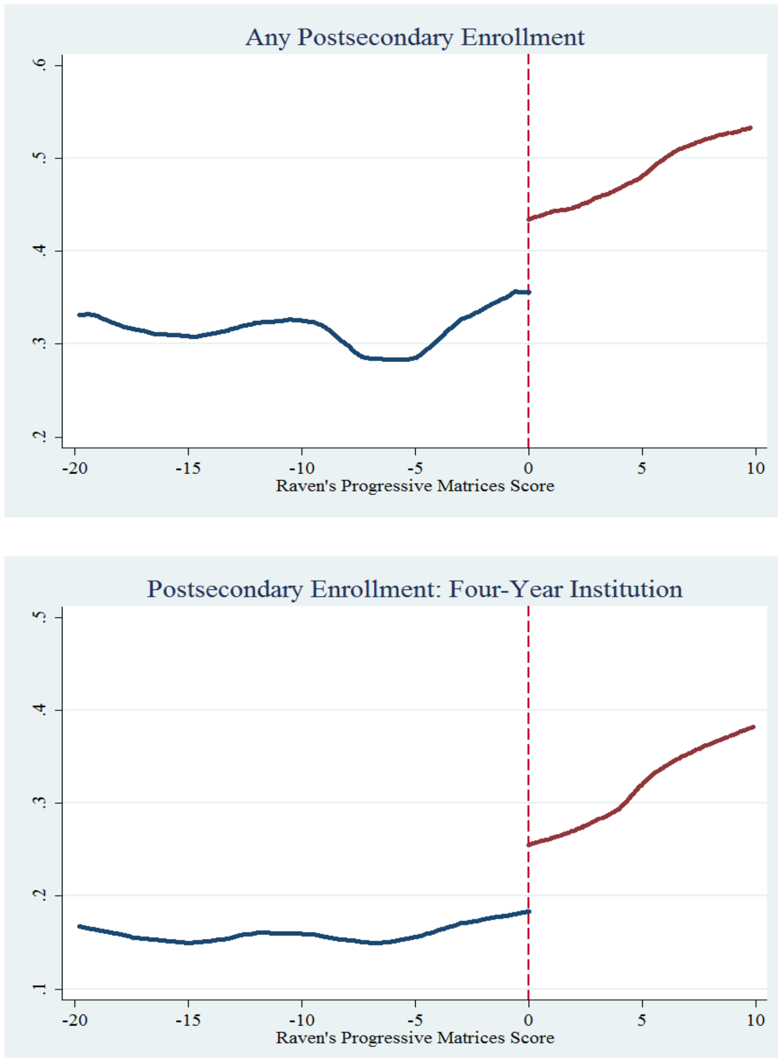
# Chapter 3 Supplemental Materials



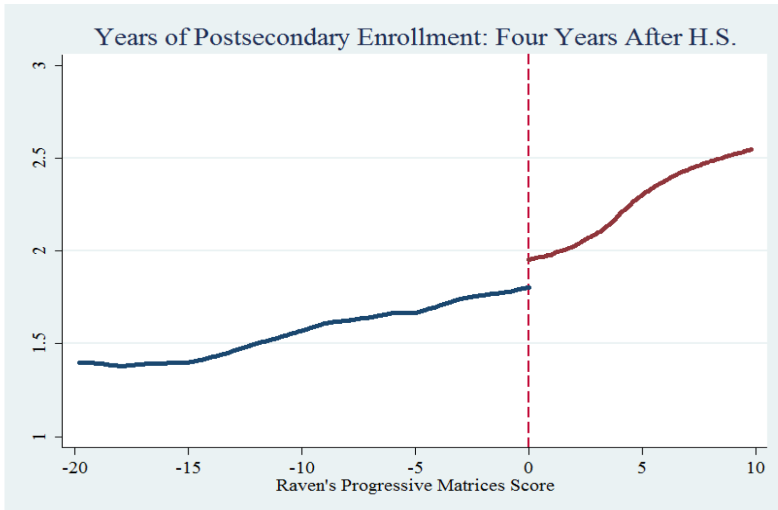

Notes: Figures display the proportion of observed high school entrants matriculating to grade 12 (top) and graduating high school (bottom) respectively, by Raven's score.Plots are of locally smoothed means weighted using a triangle kernel within the binwidths used in the analyses.

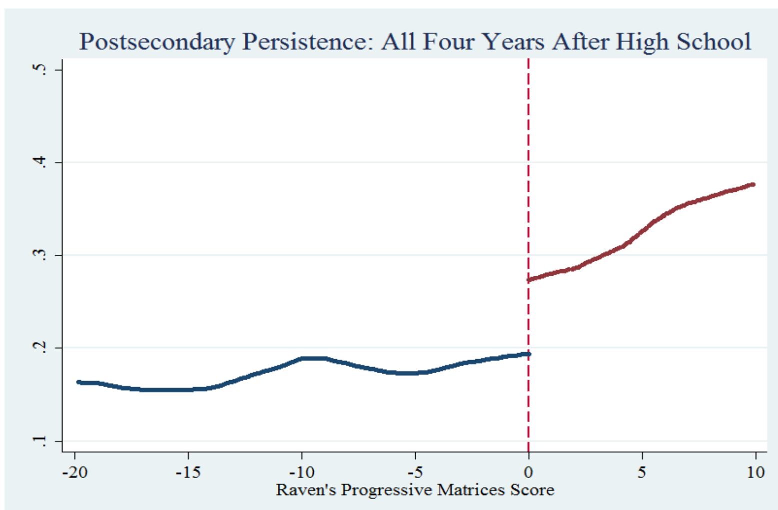
**FigureC.1**: Secondary Attainments by Raven's Score




Notes: Figures plot the proportion of observed high school students who enroll in any postsecondary (top) and a four-year postsecondary institution (bottom) by Raven's Progressive Matrices scores.

**FigureC.2**: Postsecondary Enrollment: First Year After High School




Notes: Figures plot the number of years during the first four years after expected high school graduation for which students are enrolled in any postsecondary institution (top) and the proportion of observed high school students who are enrolled in a postsecondary institution during all four years after high school (bottom).

**FigureC.3**: Postsecondary Persistence

**TableC.1**: Differential Attrition Between GATE Testing and High School

| | Observed High School Entry | | Observed Senior Year | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | *Bandwidth* $h = 3$ *points* | | | |
| Reduced Form | 0.031 | 0.014 | 0.010 | -0.004 |
| | (0.042) | (0.043) | (0.045) | (0.046) |
| First Stage | 0.851*** | 0.839*** | 0.851*** | 0.839*** |
| | (0.018) | (0.018) | (0.018) | (0.018) |
| **RD Estimate** | **0.036** | **0.017** | **0.012** | **-0.004** |
| | **(0.047)** | **(0.051)** | **(0.049)** | **(0.055)** |
| | *Bandwidth* $h = 4$ *points* | | | |
| Reduced Form | 0.013 | 0.001 | 0.004 | -0.003 |
| | (0.057) | (0.059) | (0.061) | (0.062) |
| First Stage | 0.852*** | 0.837*** | 0.852*** | 0.837*** |
| | (0.011) | (0.013) | (0.011) | (0.013) |
| **RD Estimate** | **0.016** | **0.002** | **0.005** | **-0.004** |
| | **(0.067)** | **(0.070)** | **(0.072)** | **(0.074)** |
| | *Bandwidth* $h = 5$ *points* | | | |
| Reduced Form | 0.022 | 0.009 | 0.013 | 0.003 |
| | (0.057) | (0.057) | (0.061) | (0.061) |
| First Stage | 0.851*** | 0.851*** | 0.851*** | 0.851*** |
| | (0.010) | (0.021) | (0.010) | (0.021) |
| **RD Estimate** | **0.026** | **0.011** | **0.016** | **0.003** |
| | **(0.067)** | **(0.067)** | **(0.071)** | **(0.071)** |
| | *Bandwidth* $h = 6$ *points* | | | |
| Reduced Form | 0.035 | 0.010 | 0.019 | -0.005 |
| | (0.048) | (0.047) | (0.051) | (0.050) |
| First Stage | 0.851*** | 0.845*** | 0.851*** | 0.845*** |
| | (0.009) | (0.018) | (0.009) | (0.018) |
| **RD Estimate** | **0.041** | **0.012** | **0.022** | **-0.005** |
| | **(0.056)** | **(0.056)** | **(0.060)** | **(0.059)** |
| Observations | 9,069 | 9,069 | 9,069 | 9,069 |
| Covariates | No | Yes | No | Yes |

Notes: The outcome variable is an indicator for the student being observed in the specified later years. Covariates include student age at testing date and a set of indicators for student gender, ethnicity and English learner status. Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

**TableC.2**: Postsecondary Attainment: Bandwidth Choice Robustness Checks

| | Any Postsecondary Degree Attainment | | Four-Year Degree Attainement | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | *Bandwidth $h = 3$ points* | | | |
| Reduced Form | 0.083** | 0.060** | 0.088*** | 0.059** |
| | (0.036) | (0.026) | (0.034) | (0.024) |
| First Stage | 0.893*** | 0.880*** | 0.893*** | 0.880*** |
| | (0.033) | (0.023) | (0.033) | (0.023) |
| **RD Estimate** | **0.093**** | **0.068**** | **0.099***** | **0.067**** |
| | **(0.040)** | **(0.029)** | **(0.038)** | **(0.027)** |
| | *Bandwidth $h = 4$ points* | | | |
| Reduced Form | 0.093*** | 0.073** | 0.091*** | 0.071** |
| | (0.031) | (0.032) | (0.028) | (0.030) |
| First Stage | 0.880*** | 0.882*** | 0.880*** | 0.882*** |
| | (0.012) | (0.025) | (0.012) | (0.025) |
| **RD Estimate** | **0.105***** | **0.083**** | **0.104***** | **0.080**** |
| | **(0.035)** | **(0.037)** | **(0.032)** | **(0.034)** |
| | *Bandwidth $h = 5$ points* | | | |
| Reduced Form | 0.090*** | 0.070** | 0.088*** | 0.067** |
| | (0.030) | (0.031) | (0.027) | (0.028) |
| First Stage | 0.877*** | 0.871*** | 0.877*** | 0.871*** |
| | (0.011) | (0.013) | (0.011) | (0.013) |
| **RD Estimate** | **0.103***** | **0.080**** | **0.101***** | **0.077**** |
| | **(0.035)** | **(0.035)** | **(0.031)** | **(0.032)** |
| | *Bandwidth $h = 6$ points* | | | |
| Reduced Form | 0.070*** | 0.046* | 0.068*** | 0.045* |
| | (0.026) | (0.026) | (0.023) | (0.023) |
| First Stage | 0.877*** | 0.873*** | 0.877*** | 0.873*** |
| | (0.010) | (0.012) | (0.010) | (0.012) |
| **RD Estimate** | **0.080***** | **0.053*** | **0.078***** | **0.051*** |
| | **(0.029)** | **(0.029)** | **(0.026)** | **(0.027)** |
| Observations | 6,501 | 6,501 | 6,501 | 6,501 |
| Covariates | No | Yes | No | Yes |

Notes: Local linear regression estimates of GATE on postsecondary attainment are reported for various bandwidths. Main results reported in the paper use a bandwidth of 4. See text for discussion of bandwidth choice. Bootstrapped standard errors are in parentheses.
(*** p<0.01, ** p<0.05, * p<0.1).

**TableC.3**: Postsecondary Attainment: Functional Form Specification Checks

| | Any Postsecondary Degree Attainment | | | Four-Year Degree Attainement | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Bandwidth* $h = 3$ *points* | | | | | |
| Reduced Form | 0.083** | 0.084 | 0.084 | 0.088*** | 0.077 | 0.100 |
| | (0.036) | (0.086) | (0.084) | (0.034) | (0.079) | (0.080) |
| First Stage | 0.893*** | 0.825*** | 0.826*** | 0.893*** | 0.842*** | 0.826*** |
| | (0.033) | (0.087) | (0.086) | (0.033) | (0.083) | (0.086) |
| **RD Estimate** | **0.093**** | **0.102** | **0.102** | **0.099***** | **0.092** | **0.121** |
| | **(0.040)** | **(0.104)** | **(0.102)** | **(0.038)** | **(0.094)** | **(0.097)** |
| | *Bandwidth* $h = 4$ *points* | | | | | |
| Reduced Form | 0.093*** | 0.087** | 0.088** | 0.091*** | 0.078** | 0.079** |
| | (0.031) | (0.038) | (0.037) | (0.028) | (0.034) | (0.033) |
| First Stage | 0.880*** | 0.881*** | 0.887*** | 0.880*** | 0.881*** | 0.887*** |
| | (0.012) | (0.028) | (0.024) | (0.012) | (0.028) | (0.024) |
| **RD Estimate** | **0.105***** | **0.099**** | **0.099**** | **0.104***** | **0.088**** | **0.089**** |
| | **(0.035)** | **(0.043)** | **(0.041)** | **(0.032)** | **(0.039)** | **(0.037)** |
| | *Bandwidth* $h = 5$ *points* | | | | | |
| Reduced Form | 0.090*** | 0.083*** | 0.091** | 0.088*** | 0.078*** | 0.095*** |
| | (0.030) | (0.031) | (0.035) | (0.027) | (0.028) | (0.033) |
| First Stage | 0.877*** | 0.888*** | 0.882*** | 0.877*** | 0.888*** | 0.882*** |
| | (0.011) | (0.026) | (0.031) | (0.011) | (0.026) | (0.031) |
| **RD Estimate** | **0.103***** | **0.094***** | **0.101***** | **0.101***** | **0.088***** | **0.103***** |
| | **(0.035)** | **(0.035)** | **(0.036)** | **(0.031)** | **(0.031)** | **(0.034)** |
| | *Bandwidth* $h = 6$ *points* | | | | | |
| Reduced Form | 0.070*** | 0.067** | 0.106*** | 0.068*** | 0.063*** | 0.102*** |
| | (0.026) | (0.027) | (0.037) | (0.023) | (0.024) | (0.033) |
| First Stage | 0.877*** | 0.886*** | 0.885*** | 0.877*** | 0.886*** | 0.885*** |
| | (0.010) | (0.023) | (0.020) | (0.010) | (0.023) | (0.020) |
| **RD Estimate** | **0.080***** | **0.075**** | **0.120***** | **0.078***** | **0.071***** | **0.115***** |
| | **(0.029)** | **(0.030)** | **(0.041)** | **(0.026)** | **(0.027)** | **(0.037)** |
| Observations | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 | 6,501 |
| Functional Form | Linear | Quadratic | Cubic | Linear | Quadratic | Cubic |

Notes: RD estimates of GATE on postsecondary attainment using polynomial specifications of degree one, two and three are reported for various bandwidths. Main results reported in the paper use local linear regressions (columns (1) and (4) ). Bootstrapped standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

# Bibliography


**Aaronson, D., L.Barrow, and W.Sander**, "Teachers and Student Achievement in the Chicago Public High Schools," *Journal of Labor Economics*, 2007, *25*, 95–135.

**Afridi, Farzana**, "The Impact of School Meals on School Participation: Evidence from Rural India," *Working Paper*, 2007.

**Altonji, J.G., T.E. Elder, and C.R. Taber**, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113*, 151–184.

**Angrist, JoshuaD. and AlanB. Krueger**, "Does Compulsory School Attendance Affect Schooling and Earnings?," *The Quarterly Journal of Economics*, 1991, *106* (4), pp. 979–1014.

**Bedard, Kelly and Elizabeth Dhuey**, "The Persistence of Early Maturity: International Evidence of Long-Run Age Effects," *The Quarterly Journal of Economics*, 2006, *121(4)*, 1437–7.

**Betts, JulianR., AndrewC. Zau, and LorienA. Rice**, "Determinants of Student Achievement: New Evidence from San Diego," *Public Policy Institute of California*, 2003.

**Bhatt, R.**, "The Impacts of Gifted and Talented Education," *Working Paper*, 2009.

**Bhattacharya, J., J.Currie, and S.Haider**, "Breakfast of Champions? The School Breakfast Program and the Nutrition of Children and Families," *Journal of Human Resources*, 2006, *41*, 445–466.

**Black, S.E., P.J. Devereux, and SalvanesK. G.**, "Too Young to Leave the Nest: The Effects of School Starting Age," *The Review of Economics and Statistics*, 2011, *93(2)*, 455–67.

**Bound, J. and DavidA. Jaeger**, "Do Compulsory Attendance laws Alone Explain the Association between Earnings and Quarter of Birth?," *Research in Labor Economics*, 2000, *19*, 83–108.

**Buckles, Kasey and DanielM. Hungerman**, "Season of Birth and Later Outcomes: Old Questions, New Answers," *NBER Working Paper*, 2008, *w14573*.

**Bui, S.A., S.G. Craig, and S.A. Imberman**, "Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Students," *NBER Working Paper 17089*, 2011.

**Cannon, J.S. and S.Lipscomb**, "Changing the Kindergarten Cutoff Date: Effects on California Students in Schools," *Public Policy Institute of California*, 2008.

**Card, D. and L.Giuliano**, "Does Gifted Education Work? For Whom?," *Working Paper*, 2013.

**Chetty, R., J.Friedman, and J.Rockoff**, "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood," *NBER Working Paper w17699*, 2011.

**Colasanti, Michael**, "Kindergarten Entrance Ages: A 30 Year Trend Analysis," *Mimeo*, 2007.

**Dahl, GordonB. and Lance Lochner**, "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit," *working paper*, 2010.

**Datar, Ashlesha**, "Does Delaying Kindergarten Entrance Give Children a Head Start?," *Economics of Education Review*, 2006, *25(1)*, 43–62.

**Davis, B., J.Engberg, D.N. Epple, H.Sieg, and R.Zimmer**, "Evaluating the Gifted Program of an Urban School District using a Modified Regression Discontinuity Design," *NBER Working Paper 16414*, 2010.

**Dhuey, Elizabeth and Stephen Lipscomb**, "Disabled or Young? Relative Age and Special Education Diagnoses in Schools," *Economics of Education Review*, 2010, *29(5)*, 857–72.

**Dobbie, Will**, "Teacher Characteristics and Student Achievement: Evidence from Teach For America," *Working Paper*, 2011.

**Dobkin, Carlos and Fernando Ferreira**, "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?," *Economics of Education Review*, 2010, *29*, 40–54.

**Elder, ToddE. and DarrenH. Lubotsky**, "Kindergarten Entrance Age and Childrens Achievement: Impacts of State Policies, Family Background, and Peers," *The Journal of Human Resources*, 2009, *44*, 641–83.

**Evans, WilliamN. and RobertM. Schwab**, "Finishing High School and Starting College: Do Catholic Schools Make a Difference?," *The Quarterly Journal of Economics*, 1995, *110(4)*, 941–74.

**Figlio, DavidN. and Joshua Winicki**, "The Effects of School Accountability Plans on School Nutrition," *Journal of Public Economics*, 2005, *89*, 381–394.

**Florence, M.D., M.Asbridge, and P.J. Veugelers**, "Diet quality and academic performance.," *Journal of School Health*, 2008, *78(4)*, 209–15.

**Frisvold, DavidE.**, "Nutrition and Cognitive Achievement: An Evaluatin of the School Breakfast Program," *working paper*, 2011.

**Gleason, P.M. and C.W. Suitor**, "Eating at School: How the National School Lunch Program Affects Children's Diets," *American Journal of Agricultural Economics*, 2003, *85*, 1047–1061.

**Granger, CliveW.J.**, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, 1969, *37*, 424–438.

**Hahn, J., P.Todd, and W.Van der Klaauw**, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 2001, *69(1)*, 201–09.

**Hirano, K. and G.W. Imbens**, "Estimation of Causal Effects using Propensity Score Weighting: An Application to Data on Right Heart Catheterization," *Health Services and Outcomes Research Methodology*, 2001, *2*, 259–278.

**_ , _ , and G.Ridder**, "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score," *Econometrica*, 2003, *71*, 1161–1189.

**Hofferth, S.L. and S.Curtin**, "Poverty, Food Programs, and Childhood Obesity," *Journal of Policy Analysis and Management*, 2005, *24*, 703–726.

**Imbens, G.W. and K.Kalyanaraman**, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *NBER Working Paper*, 2009, *w14726.*

**_ and T.Lemieux**, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 2008, *142(2)*, 615–35.

**Imbens, GuidoW. and JoshuaD. Angrist**, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), pp. 467–475.

**Imbens, G.W.**, "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *Review of Economics and Statistics*, 2004, *86*, 4–29.

_ **and J.M. Wooldridge**, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 2009, *47*, 5–86.

**Imberman, S.A. and A.D. Kugler**, "The Effect of Providing Breakfast on Student Performance: Evidence from an In-Class Breakfast Program," *NBER Working Paper*, 2012.

**Kleinman, R., S.Hall, and H.Green**, "Diet, breakfast and academic performance in children.," *Ann Nutr Metab*, 2002, *46(1)*, S24–S30.

**Lee, D.S. and D.Card**, "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 2008, *142(2)*, 655–74.

_ **and T.Lemieux**, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48*, 281–355.

**Long, S.K.**, "Do School Nutrition Programs Supplement Household Food Expenditures?," *Journal of Human Resources*, 1991, *26*, 654–678.

**Millimet, D.L., R.Tchernis, and M.Hussain**, "School Nutrition Programs and the Incidence of Childhood Obesity," *Journal of Human Resources*, 2010, *45(3)*, 640–654.

**Murphy, J.M., M.Pagano, and J.Nachmani**, "The Relationship of School Breakfast to Psychosocial and Academic Functioning: Cross-Sectional and Longitudinal Observations in an Inner-City School Sample.," *Arch Pediat Adol Med*, 1998, *152*, 899–907.

**Nichols, Austin**, "Causal Inference with Observational Data," *The Stata Journal*, 2007, *7(4)*, 507–41.

_ , "rd 2.0: Revised Stata Module for Regression Discontinuity Estimation," *The Stata Journal*, 2011.

**Nicklas, T.A., W.Bao, L.S. Webber, and G.S. Berenson**, "Breakfast Consumption Affects Adequacy of Total Daily Intake in Children," *J Am Diet Assoc*, 1993, *93*, 886–891.

**Puhani, PatrickA. and AndreaM. Weber**, "Persistence of the School Entry Age Effect in a System of Flexible Tracking," *Institute for the Study of Labor (IZA)*, 2007, *IZA Discussion Papers 2965.*

**Reichardt, RobertE.**, "The Cost of Class Size Reduction: Advice for Policymakers," *RAND Corporation*, 2000, *RGSD-156*, 38–71.

**Rivken, S.G., E.A. Hanushek, and J.F. Kain**, "Teachers, Schools, and Academic Achievement," *Econometrica*, 2005, *73*, 417–458.

**Rockoff, JonahE.**, "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *The American Economic Review*, 2004, *94*, 247–252.

**Rosenbaum, P.R. and D.B. Rubin**, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 1983, *70*, 41–55.

**Rubin, D.**, "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 1974, *66*, 688–701.

**Sampson, A.E., S.Dixit, A.F. Meyers, and R.Jr. Houser**, "The Nutritional Impact of Breakfast Consumption on the Diets of Inner-City African-American Elementary School Children," *J Natl Med Assoc*, 1995, *87*, 195–202.

**Schanzenbach, D.W.**, "Does the Federal School Lunch Program Contribute to Childhood Obesity?," *Journal of Human Resources*, 2009, *44(3)*, 684–709.

**Schoenthaler, S. and I.Bier**, "The effect of vitamin-mineral supplementation on juvenile delinquency among Amercian schoolchildren: A randomized, double-blind placebo controlled trial.," *Journal of Complementary and Alternative Medicine*, 2000, *6(1)*, 7–17.

**_ , _ , and K.Young**, "The effect of vitamin-mineral supplementation on the intelligence of Amercian schoolchildren: A randomized, double-blind placebo controlled trial.," *Journal of Complementary and Alternative Medicine*, 2000, *6(1)*, 19–29.

**Siega-Riz, A.M., B.M. Popkin, and T.Carson**, "Trends in Breakfast Consumption for Children in the United States from 1965-1991," *Am J Clin Nutr*, 1998, *67*, S748–56.

**Wesnes, K.A., C.Pincock, and D.Richardson**, "Breakfast reduces declines in attention and memory over the morning in schoolchildren.," *Appetite*, 2003, *41*, 329–331.