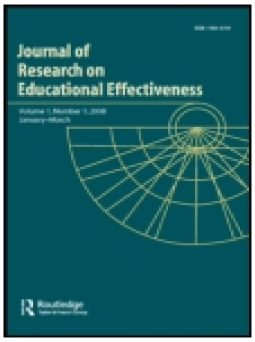

Journal of Research on Educational Effectiveness

ISSN: 1934-5747 (Print) 1934-5739 (Online) Journal homepage: https://www.tandfonline.com/loi/uree20

Expanding the Start of the College Pipeline: Ninth-Grade Findings From an Experimental Study of the Impact of the Early College High School Model

Julie A. Edmunds , Lawrence Bernstein , Fatih Unlu , Elizabeth Glennie , John Willse , Arthur Smith & Nina Arshavsky

Journal of Research on Educational Effectiveness, 5: 136–159, 2012

ISSN: 1934-5747 print / 1934-5739 online
DOI: 10.1080/19345747.2012.656182

Expanding the Start of the College Pipeline: Ninth-Grade Findings From an Experimental Study of the Impact of the Early College High School Model

Julie A. Edmunds
SERVE Center at UNCG, Durham, North Carolina, USA

Lawrence Bernstein
RTI International, Waltham, Massachusetts, USA

Fatih Unlu
Abt Associates Inc, Cambridge, Massachusetts, USA

Elizabeth Glennie
RTI International, Durham, North Carolina, USA

John Willse
University of North Carolina at Greensboro, Greensboro, North Carolina, USA

Arthur Smith
Abt Associates Inc, Cambridge, Massachusetts, USA

Nina Arshavsky
SERVE Center at UNCG, Durham, North Carolina, USA

Abstract: Early college high schools are a new and rapidly spreading model that merges the high school and college experiences and that is designed to increase the number of students who graduate from high school and enroll and succeed in postsecondary education. This article presents results from a federally funded experimental study of the impact of the early college model on Grade 9 outcomes. Results show that, as compared to control group students, a statistically significant and substantively higher proportion of treatment group students are taking core college preparatory courses and succeeding in them. Students in the treatment group also have statistically significantly higher attendance and lower suspension rates than students in the control group.

Keywords: High schools, experimental design, college readiness

INTRODUCTION

Across the United States, there has been an increasing drumbeat of dire news related to high school performance. Estimates indicate that only 70% of our nation's ninth graders

Address correspondence to Julie A. Edmunds, SERVE Center at UNCG, 2634 Durham-Chapel Hill Boulevard, Durham, NC 27707, USA. E-mail: jedmunds@serve.org

graduate within 4 years (Swanson, 2009). Furthermore, those who do graduate are often seen as underprepared for further education or the world of work. For example, 60% of employers rate students' basic skills as "fair" or "poor" (American Diploma Project, 2004), and more than one third of students graduate from high school unqualified or marginally qualified to go to college (National Center for Education Statistics, 2004). In response, the U.S. Department of Education has articulated a goal for middle and high schools that all students graduate on time from high school "prepared for at least one year of post-secondary education" (Martin, 2009).

To respond to these challenges, foundations, national organizations, and states have increased the attention paid to high school reform. Although there are many strategies being implemented across states (Edmunds & McColskey, 2007), one of the most visible has been the creation of new small schools, some of which are early college high schools. Early college high schools have been proposed as a way to increase both the number of students who graduate from high school and the number of students who are prepared for and go on to postsecondary education. Primarily located on college campuses, these high schools are designed to accelerate the academic progress of students while minimizing or even eliminating the barriers between high school and college. They are seen as particularly appropriate for students who may not have considered attending college. Students in early college high schools (we use the term "early colleges" as a shorthand) are expected to graduate in 4 to 5 years with a high school diploma and an associate's degree or 2 years of transferable college credit. Since 2002, more than 200 early colleges have been created under the auspices of the national Early College High School Initiative primarily funded by the Bill & Melinda Gates Foundation (Jobs for the Future, 2011), and many early colleges have also been created independently by districts or states.

Given their high expectations, early colleges represent a test of whether substantially redesigned high schools can realize the vision of all students graduating from high school prepared for college and work. This article looks at the extent to which early colleges are on track for achieving this goal by analyzing ninth-grade outcomes from the first large-scale experimental study of the impact of this rapidly spreading model.

THEORETICAL BACKGROUND

As articulated by the national Early College High School Initiative, early colleges should increase the number of students graduating from high school and prepared for attending college because

> encountering the rigor, depth, and intensity of college work at an earlier age inspires average, underachieving, and well-prepared high school students. In addition, the early college high school model helps reduce financial and admissions barriers faced by many low income students. (Jobs for the Future, 2005, p. 3)

Notwithstanding the rhetoric and the Initiative's rapid growth, little research has been completed on the effectiveness of the early college design (American Institutes for Research & SRI International, 2005; Jacobson, 2005). Some literature does exist on middle colleges, which have been in existence for longer and share some (but not all) of the features of early colleges. Early descriptive studies have suggested that middle colleges can increase the graduation rates and college attendance of low-performing students (Cullen, 1991; Houston, Byers, & Danner, 1992). However, an experimental study of the middle college model as implemented in Portland, Oregon, found that the model had no impact on graduation or dropout rates (Dynarski, Gleason, Rangarajan, & Wood, 1998).

The largest study completed on the early college model to date has been a 6-year national evaluation commissioned by the Bill & Melinda Gates Foundation and conducted by the American Institutes of Research and SRI International (2009). The evaluation was descriptive in nature and focused primarily on understanding the different features of implementation and the outcomes of students enrolled in the early college. This study found that early colleges were enrolling students who were underrepresented in higher education and that those students experienced success in the early college high schools. Students reported they were engaged in school and that they had a positive academic self-concept. The study also found that the early colleges' students outperformed the corresponding school district average on state assessments, although the study was not designed to account for students' entering achievement or motivation.

Given that students must apply to the early college and can thus be seen as potentially systematically different from the traditional high school population, it would be very challenging to conduct a quasi-experimental impact study that attempted to match early college students to traditional high school students. In fact, a descriptive study found that the early college populations in North Carolina had overall higher Grade 8 achievement scores and higher levels of motivation than average for students in their respective districts (Glennie & Purtell, 2008). This inherent difference highlights the critical need for an experimental study in examining the impact of a model like the early college. An experimental study, such as the one reported in this paper, can help eliminate any bias introduced by students self-selecting into the school and can help ensure the most accurate estimate of the impact of the model.

Although the Early College High School Initiative is a national intiative, the study reported in this article is examining the impact of schools only in North Carolina: schools that are part of the national initiave but are funded by the North Carolina General Assembly. The study focuses on North Carolina for three main reasons. First, with more than 70 early colleges in place—approximately one third of those established under the national initiative—North Carolina is supporting the lion's share of this reform. Second, North Carolina's model is among the best defined because these schools are all managed by the same organization, which has clearly articulated the components of the model and monitors implementation. Finally, all of the schools in the study are within the the same state, have the same student assessments, and submit the same data, which allows the focus to remain more on the model itself and not on other factors that might be influencing outcomes. The next section describes the model as implemented in North Carolina and as examined in this study.

THE EARLY COLLEGE HIGH SCHOOL MODEL

The core components of the Early College High School model may vary as it is implemented in different locations around the country. The schools that are part of North Carolina's initiative are all managed by the same entity, the North Carolina New Schools Project (NCNSP) that has conceptualized the model and provides intensive professional development centered on the core elements of the model.

North Carolina's early college model exhibits a set of specific organizational characteristics. The target population of early colleges is intended to be students who are underrepresented in college, including those who are low income, the first in their family to go to college, or a member of minority group underrepresented in college. The early colleges are autonomous schools managed by the local school district in partnership with a higher education partner, either a community college or a university. Almost all of the schools are physically located on the campus of their higher education partner, although a small number are considered “virtual” schools with their college courses being offered online. (None of the virtual schools are part of this study.) The early colleges’ maximum size is 400 students total, serving students in Grades 9 to 12 with some schools offering a 5th year or Grade 13. In most settings, students begin taking college courses in their freshman year of high school, and in all settings the expectation is that participating students will graduate from high school with 2 years of transferable college credit.

In addition to these specific organizational characteristics, early colleges are expected to implement a core set of design principles that are reflective of a high-quality high school. These six design principles as articulated by the NCNSP are as follows:

- **Ready for College:** NCNSP schools are characterized by the pervasive, transparent, and consistent understanding that the school exists for the purpose of preparing all students for college and work. They maintain a common set of high standards for every student to overcome the harmful consequences of tracking and sorting.
- **Require Powerful Teaching and Learning:** NCNSP schools are characterized by the presence of commonly held standards for high-quality instructional practice. Teachers in these schools design rigorous instruction that ensures the development of critical thinking, application, and problem-solving skills often neglected in traditional settings.
- **Personalization:** Staff in NCNSP schools understand that knowing students well is an essential condition of helping them achieve academically. These high schools ensure that adults leverage knowledge of students in order to improve student learning.
- **Redefine Professionalism:** Evident in NCNSP schools are the collaborative work orientation of staff, the shared responsibility for decision making, and the commitment to growing the capacity of staff and schools throughout the network.
- **Leadership:** Staff in NCNSP schools work to develop a shared mission for their school and work actively as agents of change, sharing leadership for improved student outcomes in a culture of high expectations for all students.
- **Purposeful Design:** NCNSP schools are designed to create the conditions that ensure the other five design principles: ready for college, powerful teaching and learning, personalization, leadership, and redefined professionalism. The organization of time, space, and the allocation of resources ensures that these best practices become common practice. (North Carolina New Schools Project, 2011)

From the organizational characteristics and design principles,[1] we created a logic model (Figure 1) that guided the overall design of our study, including the selection of implementation and outcome variables. In the next section, we describe the study’s methodology.

[1]Starting in 2010, NCNSP added the leadership design principle. Because this principle was added after the data for this study were collected, examining it was not part of our study design.

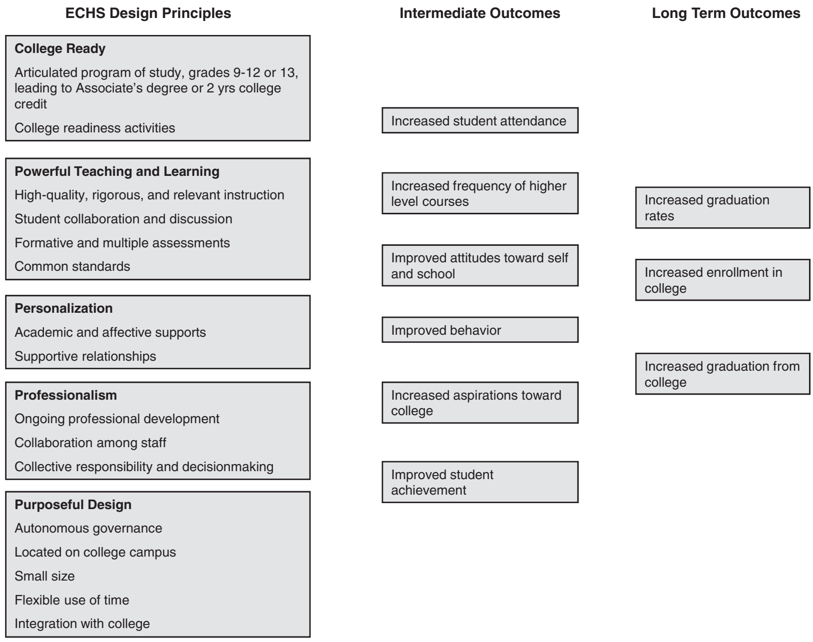

Figure 1. Logic model for North Carolina's early college high schools.

METHODOLOGY

Funded by the Institute of Education Sciences, the Study of the Efficacy of North Carolina's Early College High School model is a longitudinal experimental study examining the program's implementation and impact. The study is designed to accomplish three primary aims:

1. Determine the impact of the model on selected student outcomes,
2. Determine the extent to which impacts differ by student characteristics, and
3. Examine the implementation of the model and the extent to which specific model components are associated with positive outcomes.

We have reported ninth-grade findings from earlier, smaller samples on outcomes and implementation elsewhere (Edmunds, Bernstein, Glennie, Willse, Arshavsky et al., 2010; Edmunds, Bernstein, Unlu, Glennie, Smith et al., 2011). In this article, we report on the impact of the model on an expanded set of ninth-grade outcomes for a much larger sample of students. Specifically, we examine the following research question:

Do ninth-grade students who attend early college high schools perform significantly better than students in traditional high schools on coursetaking and course progression, attendance, behavior, and academic aspirations?

Sample

Schools participating in the study had more applicants than they had slots and agreed to use random assignment to select students for enrollment. In all cases, students were required to apply for the early college. Schools identified a pool of eligible applicants and provided that list to the research team. The research team assigned each student a randomly generated number and ordered the list from lowest to highest, creating a randomly ordered list with an embedded waitlist. Early colleges then offered students spots in the order in which they appeared on the list. In some cases, the research team conducted a stratified random lottery to allow the schools to overrepresent or equally represent certain targeted populations. For example, some schools had a local district requirement that they accept the same number of students from all traditional high school attendance zones. In all cases, the odds of acceptance into the early college for each student were recorded in the data set and accounted for in analyses via weights created based on these odds. Specifically, all analyses incorporated weights based on the inverse of these probabilities, so that students who were less likely to be admitted to the early college were given greater weight in the analyses.

Students who were on the initial waitlist and were offered spots according to the correct randomized order were included in the treatment group because their process of selection was random. If, however, a student on the waitlist was, through a nonrandom process, selected by the school to attend the early college, that student remained in the control group in terms of the study analyses. This was done because we used an intent-to-treat (ITT) framework, which is described in more depth in the analysis section.

In addition to the schools that followed the preceding process, two schools used random numbers to assign students prior to the beginning of the study. An examination of the characteristics of the treatment and control groups for these schools found no important differences between the two groups, except for a statistically significant difference in the proportion of students who were retained prior to Grade 8. The data for these schools were pooled with the data from the sample of schools included in the beginning of the formal study.

In this study, two sets of students were excluded from all analyses. The first set includes students who were in the original assignment sample but were ultimately retained in eighth grade. We exclude these students from all analyses because, if the lottery had been held later in the school year, they would not have been considered for the lottery process. Four students were excluded because they were retained in eighth grade. The second group included any students who were automatically admitted to the school for various reasons (e.g., sibling of a currently enrolled student, child of a staff member, etc.). These students were excluded from the original random assignment pool and were therefore also excluded from all analyses.

The analyses reported in this article include outcomes for a total of 1,607 Grade 9 students in 18 cohorts in 12 schools.[2] Table 1 provides the demographic characteristics of the treatment and control groups, weighted by students' probability of selection into the early college. The table shows that the only statistically significant difference between the two groups was in the pass rates of those students who took Algebra I in eighth grade; all other differences were not statistically significant. To account for these differences

[2]Some schools repeated the random assignment procedure in multiple years, and thus supplied multiple cohorts of ninth-grade students to the overall sample.

Table 1. Descriptive statistics—Grade 9 analysis sample

	Whole Sample[a] M	Treatment Group[b] M	Control Group[c] M	T-C Difference Diff.	Effect Size	p
Race & ethnicity						
American Indian	0.8%	0.8%	0.9%	–0.1%	–0.03	0.142
Asian	0.9%	1.0%	0.7%	0.3%	0.21	0.451
Black	26.8%	27.3%	26.1%	1.3%	0.04	0.328
Hispanic	8.2%	9.3%	6.6%	2.7%	0.22	0.223
Multiracial	3.1%	2.6%	3.9%	–1.4%	–0.27	0.305
White	60.2%	59.0%	61.8%	–2.8%	–0.07	0.114
Male	41.4%	41.0%	41.9%	–0.9%	–0.02	0.595
Age	15.35	15.34	15.37	–0.03	–0.07	0.093
Socioeconomic Background						
First generation College	40.8%	41.0%	40.5%	0.5%	0.01	0.922
Free or reduced-price Lunch	50.6%	51.3%	49.8%	1.6%	0.04	0.584
Exceptionality						
Disabled/Impaired	2.9%	2.5%	3.5%	–1.0%	–0.21	0.271
Gifted	11.8%	11.4%	12.4%	–1.0%	–0.06	0.825
Retained	3.7%	2.9%	4.8%	–1.9%	–0.32	0.053
Grade 8 achievement						
Math – scale (z score)	0.02	–0.01	0.06	–0.1	–0.06	0.234
Reading – scale (z score)	0.00	–0.01	0.01	0.0	–0.02	0.663
Math – pass	81.17%	82.81%	78.97%	3.8%	0.15	0.075
Reading – pass	80.09%	80.18%	79.98%	0.2%	0.01	0.906
Algebra 1 – take-up	23.00%	22.53%	23.62%	–1.1%	–0.04	0.476
Algebra 1 – pass	95.35%	98.04%	91.94%	6.1%	0.90	0.023*
Algebra 1 – scale (z score)	0.00	0.07	–0.09	0.2	0.16	0.198

Note. The proportions are weighted by students' probability of being selected into the early college.
[a]$N = 1{,}607$; this is the core analytic sample used for many outcomes and excludes students who could not be found in the ninth-grade administrative data. [b]$n = 919$. [c]$n = 688$.
*Statistically significant p values at the usual .05 level.

and increase the precision of our estimates, we include key baseline characteristics in our analyses; this is explained in more depth in the analysis section.

The sample for each analysis may vary slightly depending on the outcome analyzed. For many analyses, we exclude students who are missing a value on the outcome variable or students who are missing from the data set entirely. Students who are missing from the data set include students who are no longer enrolled in North Carolina public schools, such as students who moved, who left school but have not officially dropped out, or students who went to a private school. The specific sample is described for each outcome in the next section and is shown in Table 2.

Data Sources

To track progress toward the anticipated long-term outcomes of early colleges, which include increased graduation from high school and increased enrollment in and success in

Table 2. Sample size, by outcome analyses

Analytic Samples	Treat.	Con.	Analytic Samples	Treat.	Con.	Analytic Samples	Treat.	Con.
Baseline Sample[a]	952	736						
excluding students who could not be found in the administrative data	–33	–48						
Core analytic sample for coursetaking and success outcomes	919	688						
excluding students who dropped out	–1	–2	*excluding students in sites offering Integrated Math*	–181	–111			
			Analytic sample for math coursetaking outcomes	738	577			
Core analytic sample for behavior	918	686						
excluding students missing attendance data	–20	–30	*excluding students missing aspirations data*	–24	–48	*excluding students missing suspension data*	–39	–37
Analytic sample for attendance	898	656	Analytic sample for aspirations	894	638	Analytic sample for suspensions	879	649

Note. Treat = treatment group; Con = control group.
[a]This sample includes all students who applied to enroll in an early college and participated in the random assignment except those who were retained in the eighth grade ($n = 4$).

college, we identified a series of intermediate measures found to be previously associated with continued enrollment in high school and/or success in college.

This article includes data on the following outcome variables collected by the North Carolina Department of Public Instruction (NCDPI): coursetaking patterns and success, attendance, suspension, and aspirations. These data are linked to our study data in a longitudinal dataset. The specific measures are defined next.

College Preparatory Coursetaking and Success. Taking a core set of challenging academic courses has been connected to higher graduation rates (Lee & Burkham, 2003) and to persistence and success in college (Adelman, 2006). These courses tend to follow a standard trajectory, often defined as a "college preparatory course of study," that is frequently tied to the entrance requirements for state universities. Students who do not take an expected set of courses in Grade 9, including English I and Algebra I or higher level math courses, are less likely to graduate from high school with the courses required for college. For example, a study that looked at high school transcripts in California found that, out of the students who did *not* complete Algebra I by the end of Grade 9, only 6% had completed the courses necessary for college by the end of Grade 12 (Finkelstein & Fong, 2008). As a result, this paper looks at English I and Algebra I as core Grade 9 outcomes.

Algebra I is the first math course in a set of courses generally required for college, which includes Geometry, Algebra II, and one course at a level higher than Algebra II. Algebra I, Geometry, and Algebra II have end-of-course exams that count in a school's accountability ratings. Only Algebra I is required for graduation for this cohort of students. It is possible that students who are perceived as less capable may traditionally be steered away from taking Algebra I earlier in their high school career and from taking the other upper-level math courses at all. Given this situation, students' enrollment in higher level mathematics courses is likely a good indicator of the extent to which a school is serious about increasing the college preparedness of its students. As a result, this study also looks at the number of college preparatory math courses taken, defined as taking either Algebra I, Geometry, or Algebra II.

For each course-related outcome, we present three measures. The first is coursetaking—whether the student took the course or not—and serves as a measure of access. Given that North Carolina did not have transcript data for the period analyzed, scores on state-mandated End-of-Course (EOC) exams were used as proxies for course enrollment and success. As students were required to take the test when they took the course, a student was thus shown as taking the course if they had any score on the exam. Passing the test was used as a proxy for passing the course.[3] This may not represent an exact course pass rate given that there may be students who passed the test but did not pass the course or students who did not pass the test but did pass the course. On the other hand, the advantage of using the EOC exam as an indicator of passing the course was that it is a standardized statewide assessment, administered and scored consistently across all schools. Further,

[3]The use of mandated EOC exams as a proxy for coursetaking is appropriate for most of the schools in our sample. There are two treatment schools and one control school (affiliated with one of these two treatment schools), however, that offered Integrated Math I, II, and III instead of the traditional Algebra I, Geometry, and Algebra II sequence. Students taking Integrated Math take the Algebra I exam after the 2nd year and the Algebra II exam after the 3rd year; they do not take any Geometry exam. As a result, the math exams cannot be used as proxies for math coursetaking in these two sites. Therefore, the treatment and control students in these two sites are excluded from the math coursetaking analyses only. They are included for all other analyses.

EOC exams are curriculum-based tests focused on goals of the State Course of Study. The NCDPI periodically reviews and updates the State Course of Study, and the State Board of Education approves its objectives. Then, tests are developed and modified to measure these objectives. These exams thus provided an external check on the content students have learned in the course.

The second outcome is a traditional pass rate—the number of students who passed the test out of the number who took it. The final outcome, the one we see as the most important, is entitled successful completion and collapses the first two measures. Successful completion represents the percentage of students who took the course and passed the state-mandated test associated with the course. Compared to the traditional pass rate, this measure better captures the extent to which more students are on-track for college. It also does not penalize schools for expanding access to courses, because one way schools can artificially inflate their test scores is to restrict the types of students taking specific courses to only those who are most prepared.

The sample for these outcome measures includes students who were enrolled in school or had dropped out in the current or previous academic year(s) as well as those who were retained in the ninth grade. We include students who dropped out in this sample because the primary goal is to identify the proportion of students who are academically on-track for college—students who have dropped out can thus be seen as being off-track for college. Students who were missing in the ninth-grade data collection were excluded from these analyses (81 students did not have any ninth-grade data in the academic, school membership, or dropout files).

Attendance. Student attendance has been positively associated with progress in school (Lee & Burkham, 2003); changes in student attendance are therefore seen as a reliable indicator of students' likelihood of remaining in school. Each school reports the number of days students are absent from school to the NCDPI. Excluded from the analyses are any students with missing attendance data: students who were missing entirely from the ninth-grade data files, students who had dropped out (three students), and students who were enrolled but did not have attendance data (50 enrolled students had missing ninth-grade attendance data).

Student Behavior. Positive school behavior has been shown to be positively correlated with high school graduation (House, 1993; Lan & Lanthier, 2003; Lee & Burkham, 2003). The primary behavior-related outcome in this report is suspensions. Schools reported students who had been suspended out of school—either short term or long term. If students had been suspended more than once, each suspension was reported individually. For these analyses, we looked at the percentage of students who had been suspended at least once. The sample for these analyses excludes students who were missing all ninth-grade data or had dropped out in the analysis year. It also excludes a total of 76 students who applied to the early college for the 2005–2006 school year, a year for which we do not have suspension data.

Aspirations. Early colleges are supposed to encourage more students to consider the possibility of postsecondary education. Therefore, we would expect that treatment group students would have a higher level of aspiration toward college than control group students. As an indicator of students' college aspirations, we look at students' plans to attend a 4-year college as measured in a survey that accompanies each EOC test in North Carolina. The survey also asked if students were planning to attend a 2-year college; however, we did not include this in the postsecondary aspirations because the treatment students were already enrolled as community college students, resulting in a comparison that would be somewhat skewed. These analyses include the same subset of students as the attendance analysis, which excludes students who are missing and those who dropped out. It also excludes those students with missing values on the aspirations variable (72 enrolled students were missing aspirations data).

In addition to the outcome variables listed, the study used demographic data collected through student applications to the early colleges and by the NCDPI as covariates, including gender, race/ethnicity, free or reduced-price lunch status, disability status, English Language Learner status, and the educational level of the parents. These data were also used to identify students for the subgroup analyses; the subgroups are based on the initiative's target population and are listed next.

- *Underrepresented minority.* Students in this subgroup include those who are members of minority groups underrepresented in college. This includes students who identify themselves as African American/Black, Hispanic/Latino, and Native American/American Indian. Students who identify themselves as White, Asian, or Multiracial are considered to be non-minority because they are not underrepresented in college in North Carolina.
- *Low-income*. Students in this category are those students who were identified as being eligible for free or reduced-price lunch in eighth grade. Because high school students are less likely to sign up for free lunch than younger students (Riddle, 2011), we keep the eighth-grade low-income designation throughout a student's high school career.
- *First generation*. Students whose parents had only a high school diploma or less at the time the student applied to the early college were considered first generation. Any student who had at least one parent with some postsecondary education was not considered first generation.

With all subgroup analyses, sites with 1 or 0 students in either the treatment or control group were excluded from the analyses.

Analyses

This study examined whether students in the early college perform better than their control group peers on core outcomes including coursetaking and success, behavior, and aspirations. This paper reports experimental estimates of the average effect of lottery assignment to an early college (an ITT analysis) and an instrumental variables extension of these estimates to represent the average effect of attending an early college (local average treatment effect [LATE]). The primary impact estimates were obtained from the following multivariate linear regression model, which was customized for each outcome measure as needed:

$$Y_{ij} = \sum\nolimits_{b=1}^{B} \beta_b T_{ij} I_{ij}^{b} + \sum\nolimits_{k=1}^{K} \beta_{k+H} X_{ij}^{K} + \sum\nolimits_{b=1}^{B-1} \beta_{k+H+b} I_{ij}^{b} + \varepsilon_{ij}, \quad (1)$$

where

Y_{ij} = outcome of interest for student i in randomization block j,[4]

[4]We refer to the group of students who applied to enroll in an early college and were randomized to the treatment and control group as a "randomization block."

β_b = ITT impact estimate for the b^{th} randomization block or site,
T_{ij} = treatment status indicator which equals one if student i in block j was randomly assigned to enroll in the early college and zero otherwise,
I_{ij}^b = indicator variable for the b^{th} randomization block (b = 1, 2,. . .,B). It is set to one for student i if b = j (i.e., student i is in the b^{th} randomization block) and to zero otherwise.
β_{k+H} = association between the k^{th} student covariate and the outcome
X_{ij}^k = k^{th} (k = 1,2,. . ., K) student-level baseline covariate, such as gender, race/ethnicity, age, free or reduced price lunch status, and passing reading and math state tests in the eighth grade;
β_{K+H+b} = fixed effect for the b^{th} randomization block or site; and
ε_{ij} = usual error term for student i in randomization block j.

The analytic model in Equation 1 is designed to reflect the sampling and randomization scheme that was employed for this study. Specifically, applicants to each early college high school form a block and within each block, students were randomized to the treatment and control conditions. Each school (or block) is represented in the analysis via a block indicator (I_{ij}^b) in the model. Thus, the model accounts for school differences via the block indicators. Furthermore, within each randomization block, a treatment effect is estimated via the Treatment × Block interaction terms ($T_{ij} I_{ij}^b$). We have included the fixed Treatment × Block interaction terms rather than a random effect for the treatment indicator to reflect the purposive sampling of schools that were selected for the study. Note that if schools had been selected at random from a defined population and we were seeking to generalize the results of the study to such a broad population, we would have used a random treatment effect model, but with this purposive sample, the use of the fixed Treatment × Block interaction terms is appropriate (Raudenbush, Martinez, & Spybrook, 2007; Schochet, 2008a). This analytic strategy is consistent with those employed by large-scale Institute of Education Sciences-funded studies that utilized a similar randomization design (Bernstein, Dun Rappaport, Olsho, Hunt, & Levin, 2009; Constantine et al., 2009; Gleason, Clark, Tuttle, & Dwyer, 2010) as well as with a recently published study that examined the impact of New York City's small schools efforts (Bloom, Thompson, & Unterman, 2010).

In Equation 1, the treatment indicator in the Treatment × Block interactions (T_{ij}) captures the original random assignment status of students; thus, the resulting school- or block-specific effects (β_b) represents the ITT effect of the early college for students in the randomization block (or site) b, which is the primary effect of interest for this study. We calculate an overall ITT impact estimate by averaging these block-specific effects, weighting them proportionally to the total number of students (treatment and control) in each block. This ensures that the resulting impact estimate pertains to the average student who applied to enroll in an early college and went through the lottery. In addition, we conducted several sensitivity and specification tests including (a) using a logistic regression model instead of the linear probability model for binary outcome measures, (b) excluding sites with a severe (greater than 3:1) treatment-control imbalance, (c) using an alternate weighting scheme (weighting site-specific impact estimates by the inverse of the variance of the site-level impact), and (d) excluding sites with five or fewer students in either the treatment or control groups.[5] None of these tests yield results substantively different from those yielded by the primary analytic strategy just described.

[5]As all sites had 10 or more students in each group, this last sensitivity analysis was relevant only for specific outcomes and for the subgroup analyses. For brevity, results from these specification tests are not presented in this article, but they are available upon request.

As previously mentioned, this study primarily uses an ITT analysis coupled with a treatment-on-the-treated analysis. ITT is an analysis approach, used commonly in medical studies (Hollis & Campbell, 1999) and applied increasingly to education policy studies (Institute of Education Sciences, 2005), that keeps all study participants in the group to which they were originally assigned, regardless of whether participants actually received the entire intervention. This analysis preserves the integrity of the original random assignment (Hollis & Campbell, 1999). In this study, any students initially assigned to the early college were included in the treatment group, even if they changed their mind and did not go (no-shows) or if they later left the school (transfers) after being enrolled there for some time. In addition, students who were initially identified as being in the control group remained in the control group for analysis purposes, even if they later attended the early college for any reason (crossovers). If we were not doing an ITT analysis, no-shows would be counted in the control group and crossovers would be included in the treatment group. This would distort the balance of the original treatment and control groups produced by random assignment if these students were systematically different from those who remained in the original group to which they had been assigned (compliers).

Although the ITT approach provides the most policy-relevant impact estimate, it does ignore no-shows and crossovers and may consequently understate the impact of the early college on those who ended up participating in the intervention (Hollis & Campbell, 1999). To address this, we also calculated the treatment-on-the-treated or LATE estimates, which are calculated by dividing the ITT impact estimates by the factor (1-r-c) where r is the no-show rate and c is the crossover rate (Angrist, Imbens, & Rubin, 1996; Gennetian, Morris, Bos, & Bloom, 2005). With this approach, we operate under several assumptions: (a) that the impact of the early college is zero for the no-shows, (b) that the impact on crossovers is the same as if they had been originally assigned to the treatment group, and (c) that crossovers would have attended the early college if they had been assigned to it and that no-shows would not have attended if they were not assigned. The ITT estimate is adjusted as appropriate for each outcome. For example, for the non-math coursetaking outcomes, the proportion of no-shows[6] was .118, whereas the proportion of crossovers was .026. As a result, to obtain the LATE estimates for these outcomes, the ITT impact estimates are divided by the total compliance rate, which is $[1 - (.118 + .026)] = 0.856$ (see Table 3).

We adjust for multiple comparisons for all core outcomes using the Benjamini–Hochberg multiple comparisons correction (Benjamini & Hochberg, 1985). This correction is carried out separately within the two outcome domains (academic, and attitudinal and behavioral) and it is only applied to analyses conducted with the full analytic sample. We do not apply the multiple comparisons correction to the subgroup analyses as we did not have any a priori hypotheses regarding whether the effect of the early college was larger or smaller on a particular subgroup of students. Hence, we consider subgroup analyses as exploratory; these types of analyses are generally not subject to multiple comparisons considerations (Schochet, 2008b).

RESULTS

Results are divided into two primary categories: academic outcomes, including college preparatory coursetaking and success; and attitudinal and behavioral outcomes, which

[6]No-shows only include those students who were assigned to the early college and did not enroll. It does not include students who attended the early college for any length of time and then transferred to another school.

Table 3. Grade 9 outcomes: Adjusted impact estimates, LATE adjustment, and group means

	Adjusted Impact			Group Means[a]	
Outcomes	Estimate	p	LATE Adjustment	Early College	Control
Algebra I[b]					
% Take-up	9.7	<.001*	10.5	90.4	80.8
% Pass (takers)	–1.7	.38	–1.8	83.9	85.6
% Successful completion	5.5	.008*	5.9	74.6	69.1
College prep. math courses					
% At least one course take-up	9.7	<.001*	10.5	91.8	82.2
% At least two courses take-up	8.1	<.001*	8.8	34.2	26.1
% At least one course success	6.1	.002*	6.6	76.5	70.4
% At least two courses success	3.6	.012*	3.9	28.7	25.1
English I[c]					
% Take-up	1.8	.041	2.0	96.5	94.7
% Pass (takers)	1.9	.215	2.0	90.0	88.2
% Successful completion	3.1	.057	3.4	86.6	83.5
Attitudinal and behavioral outcomes[d]					
Absences (days)	−1.3	<.001*	–1.4	5.0	6.3
% Suspended at least once	−6.0	<.001*	–6.5	7.1	13.1
% Planning to attend 4-year college	2.3	.34	2.5	71.8	69.5

Note. LATE = local average treatment effect.

[a]The reported control group mean is the unadjusted average outcome in the control group while the reported treatment group mean equals the control mean plus the adjusted impact estimate.[b]Early college, $N = 738$; control, $N = 577$. [c]Early college, $N = 919$; control, $N = 688$. [d]Early college, $N = 918$; control, $N = 686$.

*Statistically significant estimates at the .05 level after the Benjamini–Hochberg multiple comparisons correction is applied.

include attendance, suspensions, and college aspirations. The estimated impact for each outcome was also analyzed by the early college target population subgroups of minority status, first-generation college attendee, and low-income students.

Academic Outcomes

An analysis of performance in the core Grade 9 courses of Algebra I and English I shows that the model had a 10 percentage point impact on coursetaking in Algebra I and a 5.5 percentage point impact on success in Algebra I. However, there was no statistically significant impact on pass rates. The table also shows similar size impacts on students' taking and progressing in at least one or two core college preparatory math courses, with all outcomes statistically significant. For all the math courses, the impact was greatest on the coursetaking outcomes. For English I, there were no differences that were significant after adjusting for multiple comparisons.

As described under the measures section, Table 3 shows three different course-related outcomes for the treatment and control groups. The first outcome is the percentage of students taking the course; this percentage is calculated based on the full sample of ninth graders (excluding missing students). The second outcome under each subject area is based on the percentage of students taking and passing the state-mandated standardized test that is required for the subject. The final outcome is what we call "successful completion." This percentage is calculated based on the number of students who took and passed the test from the entire sample of ninth graders (i.e., including those who did not take the test, and by definition could have not passed the test). This final mean—successful completion—encompasses both coursetaking and pass rates on the associated EOC exam. Statistically significant estimates at the .05 level after the Benjamini–Hochberg multiple comparisons correction (Benjamini & Hochberg, 1985) are shown with an asterisk. The third column includes the LATE estimates. We also present the group means—unadjusted means for the control group and adjusted means for the treatment group (i.e., the control group means plus the adjusted impact estimates for the treatment group)—for each outcome as those reflect the proportion of each group attaining the given outcome. Unadjusted means for the treatment and control groups are presented in the appendix.

Attitudinal and Behavioral Outcomes

As we have reported elsewhere (Edmunds, Willse, Arshavsky & Dallas, 2012), results show that the early college reduced the suspension rate by 6 percentage points and resulted in a drop in absences of 1.3 days, both statistically significant. There was no statistically significant difference, however, in students' college-going aspirations. Included in Table 3 are the adjusted impacts and adjusted means for three outcomes that occurred in Grade 9: the number of absences, the percentage of students in each group who had been suspended at least once during Grade 9, and the percentage who were planning to attend a 4-year college. The unadjusted means for these outcomes are presented in the appendix.

Subgroup Analyses

To test whether these results differ by subgroup, we conducted exploratory analyses for each of the subgroups following the same procedures as for the entire study sample. Table 4 shows the adjusted impacts for members of the core target populations of the early college: Students who are members of minority groups underrepresented in college (African American, Native American, and Hispanic), students who are first-generation college-goers, and students who are eligible for free or reduced-price lunch. For most outcomes, the adjusted impacts are positive, with many of these outcomes statistically significant. In particular, the highest number of statistically significant impacts was found for first-generation college-goers. The one negative impact is in Algebra I pass rates, although this difference is not statistically significant.

Bloom and Michalopoulos (2010) argued that estimating the impact of an intervention or program for a specific subgroup (e.g., minority students) is not sufficient; they recommended that such an impact should be compared to the estimated impact for the corresponding subgroup (e.g., nonminority students) to test for differential impacts. We thus conducted pairwise comparisons between the estimated impacts previously summarized for the three subgroup characteristics. In particular, for each outcome measure, we calculated the difference in the estimated impacts for underrepresented versus not underrepresented, first-generation college-bound versus not first-generation college bound, and free or reduced-price lunch eligible versus ineligible students and tested whether these

Table 4. Core outcomes: Adjusted impact estimates, by subgroup

	Underrepresented Minority[a]		First Generation College-Goers[b]		Free and Reduced-Price Eligible[c]	
Outcomes	Impact	*p*	Impact	*p*	Impact	*p*
Algebra I						
% Take-up	7.9	.010*	11.0	<.001*	12.3	<.001*
% Pass (takers)	–4.3	.259	–1.9	.577	–1.9	.549
% Successful completion	3.0	.443	4.3	.211	6.8	.034*
College prep. math courses						
% At least one course take-up	7.8	.008*	11.0	<.001*	11.4	<.001*
% At least two courses take-up	6.7	.001*	15.4	<.001*	12.3	<.001*
% At least one course successful completion	3.6	.346	5.2	.12	6.9	.027*
% At least two courses successful completion	1.2	.491	9.1	<.001*	6.8	<.001*
English I						
% Take-up	2.1	.201	4.0	.016*	2.6	.037*
% Pass (takers)	5.3	.136	2.8	.322	2.2	.392
% Successful completion	6.9	.054	5.4	.077	4.0	.136
Additional outcomes						
Absences (days)	–1.1	.08	–1.2	.071	–1.8	.004*
Suspensions (% suspended at least once)	–5.8	.034*	–7.4	.001*	–5.7	.015*
% Planning to attend 4-year college	5.0	.204	1.0	.805	5.0	.176

[a]$n = 567$. [b]$n = 637$. [c]$n = 772$.
*Statistically significant *p* values at the usual .05 level.

differences were statistically significant.[7] Table 5 presents the corresponding results, very few of which were statistically significant. The statistically significant findings were in math-related outcomes for first-generation and for low-income students. This suggests that the model had a greater impact on math outcomes for these populations than for students who were not members of these populations.

DISCUSSION

Results from this experimental study show that the Early College High School model is expanding the college preparatory pipeline, in terms of increasing the number of students on track for college. For purposes of this study, being on-track for college in Grade 9 can be seen as taking and successfully completing English I and a mathematics course at the level of Algebra I or higher. Early college students were statistically significantly more likely to have taken and successfully completed core college preparatory mathematics courses. There was no statistically significant difference in English.

[7]It is important to note that such tests tend to have low statistical power.

Table 5. Core outcomes—Comparison of subgroup impacts—Grade 9

	Differences in the Impact Estimates for . . .					
	Minority vs. Nonminority		First-Gen. College vs. Not First-Gen. College		Free or Reduced-Price Lunch vs. Not Free or Reduced-Price Lunch	
	Diff.	*p*	Diff.	*p*	Diff.	*p*
English I						
% Take-up	0.3	.884	4.0	.039	2.4	.138
% Pass (takers)	4.8	.213	2.0	.552	0.4	.909
% Progress	5.1	.199	4.6	.197	2.3	.486
Algebra I						
% Take-up	–3.2	.366	3.3	.302	4.7	.133
% Pass (takers)	–3.6	.405	–2.2	.597	–0.5	.891
% Progress	–4.1	.368	–2.7	.538	2.1	.612
College prep. math coursetaking						
% At least one course take-up	–2.9	.388	3.3%	.278	3.1	.304
% At least two courses take-up	–2.1	.439	11.6*	<.001	6.5*	.012
% At least one course progress	–3.8	.397	–2.1%	.629	1.6	.693
% At least two courses progress	–3.7	.152	8.6*	.004	5	.054
Attitudinal and behavioral outcomes						
Absences (days)	0.2	.792	0.2	.828	–0.8	.229
Suspensions (% suspended at least once)	0.6	.854	–2.2	.465	0.2	.941
% Planning to attend 4-year college	4.5	.367	–2.9	.563	5.3	.288

Note. Gen = generation; Diff = difference.
*Statistically significant *p* values at the usual .05 level are noted with an asterisk.

When a student takes and completes a course such as Algebra I in ninth grade, that student is more likely to graduate from high school with the courses needed for college entry. For example, an analysis of California's coursetaking patterns found that 40% of students who took Algebra I (or higher) in ninth grade completed the courses needed for college; this was compared to only 6% of the students who had not taken Algebra I in ninth grade (Finkelstein & Fong, 2008).

The model's larger impact in mathematics, as compared to English, appears to be resulting from the difference in coursetaking expectations in the early college and the standard coursetaking expectations in North Carolina public schools. In North Carolina, 96% of ninth graders take English I, whereas only 70% take Algebra I or higher. In the early college, there are very limited course options with almost all students required to follow an honors-level college preparatory course of study. This is borne out by the over 90% college preparatory math coursetaking rate in the early colleges (only one early college in our sample had any enrolled ninth graders not taking Algebra I or higher). In contrast, the high school mathematics curriculum in traditional high schools is often seen as being used to sort students into different tracks of study, such as general math, prealgebra, and college preparatory math including Algebra I and Geometry (Gamoran & Hannigan, 2000).

It is possible that traditional high schools could get the same impact on coursetaking results if they enrolled more students in Algebra I courses. Our data show, however, that this is not happening currently. In our control sample, approximately one fifth of the students had not taken any college preparatory math course by the end of Grade 9, compared to about 6% of the early college students (see the appendix for the unadjusted means). Although this study is not collecting data on why traditional high schools might not offer Algebra I for more students, we do have data from another study we have conducted on the impact of high school reform models on students' mathematics and science coursetaking. These data indicate that the traditional high schools tend to place students into ninth-grade math courses other than Algebra I because they believe that these students would benefit from taking an additional mathematics course prior to taking Algebra I in 10th grade or later (Arshavsky, Edmunds, Miller, & Corritore, 2011; Edmunds & Arshavsky, 2011).

Consistent with studies that have suggested that all students benefit from exposure to more rigorous courses (Burris, Welner, Wiley, & Murphy, 2008; Gamoran & Hannigan, 2000), there have been efforts to mandate core college preparatory courses for all students. Recent studies have indicated that these efforts have met with mixed success. For example, a study of Chicago's efforts to mandate Algebra I and English I for all ninth graders (Allensworth, Nomi, Montgomery, & Lee, 2009) found an increase in coursetaking and credits earned in both subjects. There were no adverse consequences for increased English I enrollment; however, increased Algebra I enrollment was accompanied by an increase in course failures and a decline in attendance. The policy had no impact on standardized achievement test scores in either reading or math and no perceived impact on long-term outcomes.

Similar to the Chicago study, our study found some positive impacts on coursetaking and successful course completion. Our study's 5.5 percentage point impact on Algebra I course completion was higher than in the Chicago study, which found impacts on Algebra I credits earned of 1.0 percentage points for average performers and –0.4 percentage points for high performers. Our English findings were lower, on the other hand. The Chicago study found impacts of 11.7 and –3.2 percentage points on English I credits earned for average and high performers respectively. These differences in results appear to be driven primarily by the baseline curriculum expectations between the two settings. As we previously noted, almost all ninth graders in North Carolina take English I, compared to only 64% of Chicago's ninth graders prior to the policy change (Allensworth et al., 2009).

Contrary to the Chicago findings, we actually found no difference in pass rates in either English I or Algebra I. As a reminder, our study uses pass rates on state-mandated and standardized EOC exams as a proxy for passing the test, which allows our findings to be unaffected by any teacher bias in grading. In both Algebra I and English I, the EOC exam pass rates for early college students were not significantly different from the pass rates for the control group.[8]

[8]This finding differs from earlier analyses we have reported where the treatment group had significantly higher enrollment in math courses accompanied by lower pass rates on the exams associated with those courses (citation pulled for review). These earlier analyses were conducted on a much smaller sample of only two schools and were primarily driven by one or two early college teachers with low performance in their classes. Because early colleges are very small, they often only have one teacher for each subject area; as a result, an individual teacher's performance can drive overall school results in a particular subject. As we have added more schools to our sample, the influence of any individual teacher on the overall outcomes of the model has been reduced. We believe the findings reported in this paper are thus more reflective of the impact of the model.

Why do our results differ from the Chicago findings when both interventions essentially mandated a core college preparatory curriculum? One possible explanation is that the Chicago intervention changed only the coursetaking requirements for students, whereas early colleges changed the coursetaking requirements as part of a more comprehensive school reform effort focused on preparing all students for college. As the Chicago study authors argue in their conclusion, "Getting the content and structure of courses right is just the first step. Real improvements in learning require strategies that get students feeling excited about learning, attending school and classes regularly, and working hard in whatever course they are taking" (Allensworth et al., 2009, p. 384). Their recommendation is consistent with other studies which have suggested that, for students to be successful in more rigorous courses, they need additional academic and social support (Lee & Smith, 1999; Swanson, Mehan, & Hubbard, 1995).

Early colleges thus changed the coursetaking expectations for students at the same time as they were attempting to implement the model's other design principles, such as changing instructional strategies, improving relationships, providing more academic support, and implementing a more collegial and professional working environment for the staff. As part of our study, we examined whether early colleges did in fact reflect a different school environment for students. As described in more depth elsewhere (Edmunds, Willse, Arshavsky & Dallas, 2012), we administered surveys on students' attitudes and experiences to both treatment and control students and analyzed the results using the same analytic procedures described in this article. Early college ninth graders reported more positive school experiences than control students on all of the dimensions we examined. In particular, early college students reported more rigorous instruction (impact[9] = 0.53 *SD*s), more relevant instruction (impact = 0.42 *SD*s), higher academic expectations (impact = 0.68 *SD*s), better relationships (impact = 0.37 *SD*s), and more frequent and varied types of support (impact = 1.12 *SD*s) than students in the control group. All of these effects were statistically significant at $p \leq .001$. We also conducted site visits to all participating early colleges.[10] In interviews, early college students reported that teachers worked hard to ensure that every student understood the content. Similarly, staff members commented that they saw it as their job to ensure that all students will be ready for college, regardless of where the students started. Students also reported being in an environment where academic work was valued and hard work was expected. We theorize that the different environment of the early college is contributing to some of the differences in impacts we find between our work and the work of Chicago.

Although coursetaking and student achievement are critical outcomes, behavioral outcomes are equally important, with suspensions and absences both associated with increases in students' dropping out of school (Arcia, 2006; Finn, 1989). Our study has shown that the early college has led to students being absent 1.3 fewer days of school. This is in contrast to the Chicago study, where the change in coursetaking requirements resulted in an average increase in absences of 1.6 days. In addition, the early college suspension rate was cut approximately in half. As previously postulated, we believe that these positive impacts occur because of other aspects of the early college design. Data collected from

[9]The effect sizes were Glass's Δ, which is calculated by dividing the adjusted mean difference between groups by the standard deviation of the control group, an approach recommended when the variance for the treatment group differs from that of the control group (Lipsey & Wilson, 2001).

[10]Site visits to the control schools were beyond the scope of this study, especially given the fact that the control students were spread throughout all other high schools in the county in which the early college was located.

interviews suggest some specific mechanisms through which the early college model might be obtaining these behavioral outcomes. In particular, schools used their small size to create intentional structures designed to build relationships with students. Students frequently commented on how the teachers know a tremendous amount about students' lives; students even joked that the teachers know too much about them. Teachers also commented on how they were aware of what was happening in students' lives and knew when something was not going well. In an environment such as this, a teacher may be more willing to understand why a student is being more disruptive than normal. In contrast, in larger schools, where teachers have responsibility for more than 100 students daily, teachers may not have the time or emotional energy to determine why a student is misbehaving in class; it might be easier to simply suspend the student. The potential influence of positive relationships is consistent with research that found that schools with high expectations and/or supportive environments all had lower suspension rates than schools that had low expectations and low support. For example, schools with supportive environments had a suspension rate for African American students of 9 percentage points lower than schools with low expectations and support (Gregory, Cornell, & Fan, 2011).

One unexpected finding is that there was no statistically significant difference in students' aspirations to attend a 4-year university. This may actually not be surprising given that other research has suggested that most high school students believe they are going to college (Venezia, Kirst, & Antonio, 2003). Another point to remember here is that early college students have the potential to earn a 2-year degree while still in high school. For some students, this may be sufficient postsecondary training for their career choice and thus they may not feel the need for additional postsecondary education.

Our analyses also show that the early college model is having a positive impact on students, with the treatment group outperforming the control group for all subgroups and virtually all outcomes; many of these impacts were statistically significant. When comparing outcomes for each subgroup (e.g., the impacts for minority vs. the impacts for nonminority), the results are not as clear. Some impacts appear to be higher for the targeted populations (as compared to students who are not members of the target population), whereas other impacts appear to be lower for the targeted populations. We hope that future analyses will provide a clearer picture.

LIMITATIONS

Despite the advantages of a randomized controlled trial, this study does have limitations that need to be considered. The first limitation is one of external validity, that is, that our sample may not be representative of all early college high schools in that schools had to meet two conditions to participate: (a) they had to have more applicants than slots, which ruled out most schools in very small districts, and (b) they had to be willing to use a lottery. Our site visits to schools participating in the study and to other schools around the state suggest that the schools in our study do not differ systematically in any way, but this has not been confirmed with empirical data.

Second, students had to apply to the early college. Therefore, the study's results only apply to students who had sufficient interest, motivation, or family support to apply. These schools are thus similar to magnet or charter schools, which also involve student choice to attend. It is worth noting, however, that the experimental design ensures that the control group consists of students who were similarly motivated to apply. Nevertheless, the fact that it is a self-selected group of students could have accounted for some of the differences that we find between our study and the Chicago study.

The final limitation we identify has to do with the way in which the early college model was implemented in North Carolina. The early colleges in North Carolina were required to incorporate relatively well-defined program components. In addition, the NCNSP provided extensive professional development and support to these schools, which has resulted in overall high levels of implementation across the vast majority of schools in our study. In other settings, early colleges can be structured very differently, and the implementation may be different, which may lead to different results. A significant implication of this limitation is that the many groups seeking to replicate this promising model should carefully consider their program's components and how they will support those schools in implementing the components at a high level of fidelity.

CONCLUSION AND NEXT STEPS

The early results from the Study of the Efficacy of North Carolina's Early College High School model show that this model is making substantial progress toward creating an environment where all students graduate from high school prepared for college and work. The data show that these schools are succeeding at significantly increasing the number of students who are on track for college. The schools are also creating environments that result in better attendance and fewer suspensions. Some of these effects are quite large.

As the study moves forward, we will be collecting information on the same outcomes for the upper grade levels. We will also be adding more schools and more students to the sample. Our future research should provide insight into strategies and approaches that schools and districts can use as they work toward realizing a vision of all students graduating from high school prepared for college and work.

ACKNOWLEDGMENTS

This material is based upon work supported by the Institute of Education Sciences under grant number #R305R060022. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Institute of Education Sciences.

REFERENCES

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.

Allensworth, E., Nomi, T., Montgomery, N., & Lee, V. E. (2009). College preparatory curriculum for all: Academic consequences of requiring Algebra and English I for ninth graders in Chicago. *Educational Evaluation and Policy Analysis, 31*, 367–391

American Diploma Project. (2004). *Ready or not: Creating a high school diploma that counts*. Retrieved from Achieve website: www.achieve.org/dstore.nsf/lookup/ADPreport/$file/ADPreport.pdf

American Institutes of Research & SRI International. (2005). *Early College High School Initiative evaluation year-end report: 2003–2004*. Washington, DC and Arlington, VA: Authors.

American Institutes of Research & SRI International. (2009). *Six years and counting: The ECHSI matures*. Washington, DC and Arlington, VA: Authors.

Angrist, J., Imbens, G., & Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444–455.

Arcia, E. (2006). Achievement and enrollment status of suspended students. *Education and Urban Society*, *38*, 359–369.

Arshavsky, N., Edmunds, J. A., Miller, L. C., & Corritore, M. (2011). *Success in the college preparatory mathematics pipeline: The role of policies and practices employed by three high school reform models*. Greensboro: SERVE Center at the University of North Carolina at Greensboro, Working Paper.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, *57*(1), 289–300.

Bernstein, L., Dun Rappaport, C., Olsho, L., Hunt, D., & Levin, M. (2009). *Impact Evaluation of the U.S. Department of Education's Mentoring Program*. Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Bloom, H. S., & Michalopoulos, C. (2010). *When is the story in the subgroups? Strategies for interpreting and reporting intervention effects for subgroups*. New York, NY: MDRC.

Bloom, H. S., Thompson, S. L., & Unterman, R. (2010). *Transforming the high school experience*. New York, NY: MDRC.

Burris, C. C., Welner, K. G., Wiley, E. W., & Murphy, J. (2008). Accountability, rigor, and detracking: Achievement effects of embracing a challenging curriculum as a universal good for all students. *Teachers College Record*, *110*, 571–607.

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification, Final report*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Cullen, C. (1991). Membership and engagement at Middle College High School. *Urban Education*, *26*, 83–93.

Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998). *Impacts of dropout prevention programs*. Princeton, NJ: Mathematica Policy Research.

Edmunds, J. A., & Arshavsky, N. (2011). *The importance of intentionality: How schools implement policies and practices that influence the mathematics and science pipeline* (Working paper). Greensboro: SERVE Center at University of North Carolina at Greensboro.

Edmunds, J. A., Bernstein, L., Glennie, E., Willse, J., Arshavsky, N., Unlu, F., Bartz, D., Silberman, T., Scales, W.D., & Dallas, A. (2010). Preparing students for college: the implementation and impact of the Early College High School model. *Peabody Journal of Education*, *85*, 348–364.

Edmunds, J. A., Bernstein, L., Unlu, F., Glennie, E., Smith, A., Arshavsky, N. (2011). *The impact of the Early College High School Model on core 9th and 10th grade outcomes*. Paper presented at the Annual Meeting of the Society for Research on Educational Effectiveness, Washington, DC.

Edmunds, J.A., Willse, J., Arshavsky, N., & Dallas, A. (2012). *Mandated engagement: the impact of early college high schools.*

Edmunds, J. A., & McColskey, W. (2007). *Levers for change: Southeast Region state initiatives to improve high schools. (Issues & Answers Report, REL 2007–No. 024)*. Washington DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.

Finkelstein, N. D., & Fong, A. B. (2008). *Coursetaking patterns and preparation for postsecondary education in California's public university systems among minority youth. (Issues & Answers Report, REL 2008–No. 035)*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, *59*, 117–142.

Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis*, *22*, 241–254.

Gennetian, L. A., Morris, P. A., Bos, J. S., & Bloom, H. S. (2005). Constructing instrumental variables from experimental data to explore how treatments produce effects. In H. S. Bloom (Ed.), *Learning more from social experiments: Evolving analytic approaches* (pp. 75–114). New York, NY: Russell Sage Foundation.

Gleason, P., Clark, M., Tuttle, C. C., & Dwyer, E. (2010). *The evaluation of charter school impacts: Final report (NCEE 2010–4029).* Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Glennie, E., & Purtell, K. (2008, April). *Who's being served? Do North Carolina Early College High Schools serve their target population?*Paper presented at the Annual Meeting of the American Educational Research Association, New York.

Gregory, A., Cornell, D., & Fan, X. (2011). The relationship of school structure and support to suspension rates for Black and White high school students. *American Educational Research Journal*, *48*, 904–934. doi:10.3102/0002831211398531

Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analyses? Survey of published randomised controlled trials. *British Medical Journal*, *319*, 670–674.

House, J. (1993). The relationship between academic self-concept and school withdrawal. *Journal of Social Psychology*, *133*, 125–127.

Houston, A., Byers, S., & Danner, D. (1992). A successful alternative to traditional education: Seattle Middle College High School at Seattle Central Community College. *The Journal of Negro Education*, *61*, 463–470.

Institute of Education Sciences. (2005). *Key items to get right when conducting a randomized controlled trial in education*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/guide_RCT.pdf

Jacobson, J. (2005). The early college experiment. *Chronicle of Higher Education*, *51*(27), 36–38.

Jobs for the Future. (2005). *Early College High School Initiative: Core principles*. Boston, MA: Author.

Jobs for the Future. (2011). *Welcome to Early College High School*. Retrieved from http://www.earlycolleges.org

Lan, W., & Lanthier, R., L. (2003). Changes in students' academic performance and perceptions of school and self before dropping out of schools. *Journal of Education for Students Placed at Risk*, *40*, 309–332.

Lee, V. E., & Burkham, D. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal*, *40*, 353–393.

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: the role of school academic press. *American Educational Research Journal*, *36*, 907–945.

Lipsey, M. W., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Martin, C. (2009, July 28). *U.S. Department of Education: Vision and Initiatives* [PowerPoint slides]. Retrieved from New Hampshire Department of Education website: http://www.education.nh.gov/data/documents/us_doe_data_sys.ppt

National Center for Education Statistics. (2004). *The high school transcript study*. Washington, DC: U.S. Department of Education.

North Carolina New Schools Project. (2011). *Design principles*. Retrieved from http://newschoolsproject.org/our-strategy/design-principles

Raudenbush, S. W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, *29*, 5–29.

Riddle, W. (2011, June). *Title I and high schools: Addressing the needs of disadvantaged students at all grade levels* [Policy Brief]. Washington, DC: Alliance for Excellent Education.

Schochet, P. Z. (2008a). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics*, *33*, 62–87.

Schochet, P. Z. (2008b). *Technical methods report: Guidelines for multiple testing in impact evaluations* (NCEE 2008–4018). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
Swanson, C. B. (2009, June 11). Gauging graduation, pinpointing progress. *Education Week*. Retrieved from http://www.edweek.org
Swanson, M. C., Mehan, H., & Hubbard, L. (1995). The AVID classroom: academic and social support for low-achieving students. *Creating New Educational Communities. Ninety-fourth Yearbook of the National Society for the Study of Education. Part I* (pp. 53–69). Chicago, IL: University of Chicago Press.
Venezia, A., Kirst, M. W., & Antonio, A. L. (2003). *Betraying the college dream: How disconnected K-12 and postsecondary education systems undermine student aspirations*. Stanford, CA: Stanford Institute for Higher Education Research.

APPENDIX

Table A1. Grade 9 outcomes: Unadjusted group means

	Unadjusted Means	
Outcomes	Treatment	Control
Algebra I[a]		
% Take-up	93.9	80.8
% Pass (takers)	85.2	85.6
% Successful completion	80.1	69.1
College prep. math courses		
%At least one course take-up	95.0	82.2
%At least two courses take-up	35.0	26.1
% At least one course success	81.8	70.4
% At least two courses success	29.6	25.1
English I[b]		
% Take-up	97.2	94.7
% Pass (takers)	90.5	88.2
% Successful completion	87.9	83.5
Attitudinal and behavioral outcomes[c]		
Absences (days)	4.7	6.3
% Suspended at least once	6.5	13.1
% Planning to attend 4-year college	73.4	69.5

[a]Treatment $n = 738$, control $n = 577$. [b]Treatment $n = 919$, control $n = 688$. [c]Treatment $n = 918$, Control $n = 686$.