Journal of Research on Educational Effectiveness

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

The Struggle to Pass Algebra: Online vs. Face-to-Face Credit Recovery for At-Risk Urban Students

Jessica B. Heppen, Nicholas Sorensen, Elaine Allensworth, Kirk Walters, Jordan Rickles, Suzanne Stachel Taylor & Valerie Michelman

To cite this article: Jessica B. Heppen, Nicholas Sorensen, Elaine Allensworth, Kirk Walters, Jordan Rickles, Suzanne Stachel Taylor & Valerie Michelman (2017) The Struggle to Pass Algebra: Online vs. Face-to-Face Credit Recovery for At-Risk Urban Students, Journal of Research on Educational Effectiveness, 10:2, 272-296, DOI: 10.1080/19345747.2016.1168500

To link to this article: http://dx.doi.org/10.1080/19345747.2016.1168500

JOURNAL OF RESEARCH ON EDUCATIONAL EFFECTIVENESS
2017, VOL. 10, NO. 2, 272–296
http://dx.doi.org/10.1080/19345747.2016.1168500

The Struggle to Pass Algebra: Online vs. Face-to-Face Credit Recovery for At-Risk Urban Students

Jessica B. Heppen[a], Nicholas Sorensen[a], Elaine Allensworth[b], Kirk Walters[a], Jordan Rickles[a], Suzanne Stachel Taylor[a], and Valerie Michelman[b]

ABSTRACT

Students who fail algebra are significantly less likely to graduate on time, and algebra failure rates are consistently high in urban districts. Identifying effective credit recovery strategies is critical for getting students back on track. Online courses are now widely used for credit recovery, yet there is no rigorous evidence about the relative efficacy of online versus face-to-face credit recovery courses. To address this gap, this study randomly assigned 1,224 ninth graders who failed algebra in 17 Chicago public high schools to take an online or face-to-face algebra credit recovery course. Compared to students in face-to-face credit recovery, students in online credit recovery reported that the course was more difficult, were less likely to recover credit, and scored lower on an algebra posttest. There were no statistically significant differences by condition on any outcomes measured during the second year of high school (standardized mathematics test and algebra subtest scores, likelihood of passing subsequent math classes, cumulative math credits, or on-track rates). The benefits and challenges of online learning for credit recovery are discussed in light of the findings to date.

KEYWORDS
online learning
randomized controlled trial
credit recovery
algebra
at-risk students

Failing core academic courses during the first year of high school is a strong signal of trouble to come. More students fail courses in ninth grade than in any other grade, and a disproportionate number of these students subsequently drop out (Herlihy, 2007). Research indicates that academic performance in core courses during the first year of high school is the strongest predictor of eventual graduation (Allensworth & Easton, 2005, 2007).

Algebra failure is of particular concern in high schools across the country. Pass rates are consistently low, particularly in urban districts (Ham & Walker, 1999; Helfand, 2006), and students who fail Algebra I are dramatically less likely to graduate on time than students who pass. Students who fail key gateway courses such as Algebra I need opportunities to recover content that they have not yet mastered and to recover credits required for graduation. Thus, many schools offer credit recovery programs to give students an opportunity to retake failed classes to get them back on track and keep them in school (Watson & Gemin, 2008).

CONTACT Jessica B. Heppen jheppen@air.org American Institutes for Research, 1000 Thomas Jefferson Street NW, Washington, DC 20007, USA.
[a]American Institutes for Research, Washington, DC, USA
[b]University of Chicago Consortium on School Research, Chicago, Illinois, USA

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uree.
Supplemental data for this article can be accessed on the publisher's website at http://dx.doi.org/10.1080/19345747.2016.1168500

In recent years, online learning has emerged as a practical and popular strategy for credit recovery, signaling general agreement among district and school practitioners that expanding credit recovery options through online courses may help more students get back on track toward graduation (e.g., Atkins, Brown, & Hammond, 2007). States, districts, and schools are investing significant resources into building the infrastructure to offer online credit recovery. However, no rigorous evidence currently exists on the efficacy of online credit recovery in high school. Research on online learning has predominantly focused on postsecondary students (Means, Toyama, Murphy, & Bakia, 2013) and advanced students ready for acceleration (Heppen et al., 2012)—very different populations than students who fail algebra in ninth grade. Schools and districts must choose whether to use online providers for credit recovery, or find teachers for traditional face-to-face (f2f) classes, without evidence about their relative effects for this population of at-risk students.

In this study, we partnered with the Chicago Public Schools (CPS) to compare online versus f2f credit recovery options for students who failed algebra in their first year of high school. This study used an experimental design in which half of the students who failed Algebra I in the spring and subsequently enrolled in summer school were randomly assigned to an online credit recovery course offered by a widely used online course provider, and the other half were assigned to a traditional f2f summer course. In this paper, we report on the impact of online versus f2f summer credit recovery in Algebra I on outcomes through the second year of high school for a sample of 1,224 students who participated in the summer of 2011 or 2012. The study asks: What is the relative impact of online and f2f Algebra I for credit recovery on students' (a) experiences in the class? (e.g., perceived class difficulty, teacher expectations); (b) math skills and mindsets? (e.g., end-of-course algebra test and standardized math and algebra assessment scores, reported liking of and confidence in math); (c) grades and likelihood of successfully recovering Algebra I credit? and (d) subsequent math course-taking performance and credit accumulation?

The sections that follow provide background and the rationale for the study, describe the theory of action highlighting the ways in which an online or f2f Algebra I credit recovery course may impact short and longer term outcomes, provide an overview of the study design, and detail the methods and results. The discussion considers the implications of the findings for practice and further research.

Background and Context for the Study

The high school dropout problem continues to be a national crisis: almost one in five public high school students either leaves before graduating or does not earn a regular high school diploma within four years (Stetser & Stillwell, 2014). Research is clear that ninth grade is a critical transition year, with student behavior and performance strongly predicting the likelihood of on-time graduation (Allensworth & Easton, 2005; Herlihy, 2007; Neild & Balfanz, 2006). The relationship between credit attainment and graduation is so strong that for CPS students, each semester course failure in ninth grade is associated with a 15-percentage-point decline in four-year graduation rates (Allensworth & Easton, 2007).

Algebra failure, especially among ninth graders, continues to be a particularly intractable problem in districts across the country. For example, six years after the implementation of an initiative to increase access to Algebra I, failure rates for freshmen in Milwaukee were 47% (Ham & Walker, 1999). Similarly, Helfand (2006) reported that 44% of ninth graders failed Algebra I in Los Angeles. In CPS, the location of this study, 37% of first-time freshmen failed one or both semesters of Algebra in 2009–10 (the year prior to the start of the study).

Algebra I failure rates have been closely tracked in CPS following implementation of a districtwide policy in 1997 that all high school students enroll in a college-preparatory curriculum. The policy raised graduation requirements and eliminated previously available remedial courses, so that all ninth-graders took Algebra I or a higher course in the mathematics sequence (geometry, Algebra II) (Lee & Ready, 2009). As a result of the policy, more ninth graders enrolled in Algebra I, but their failure rates increased, and they were no more likely to take advanced mathematics courses (Allensworth, Nomi, Montgomery, & Lee, 2009). In subsequent years, the district implemented a policy to reduce failure rates by requiring students entering ninth grade with below-average math scores to take two periods of algebra. Although students' test scores improved with the policy, their pass rates did not (Nomi & Allensworth, 2009). Thus, ninth-grade algebra failure remains a key concern in many CPS high schools.

Typically in CPS, students who fail one or both semesters of Algebra I still enroll in the next mathematics course in the sequence (geometry or Algebra II) the following year, but to earn a diploma they must eventually recover the Algebra I credit at some point during high school. However, Algebra I credit recovery rates are low; in 2009–10, only 13% of CPS ninth graders who failed second-semester Algebra I (when the content is considered to be more challenging and failure rates are historically higher than in the first half of the course) recovered the credit during the summer after ninth grade. Identifying effective ways to broaden opportunities for these students to recover credits early in high school is of critical importance in this district and in others across the country.

Online Courses in K–12 Settings

Online learning is expanding rapidly in U.S. secondary schools. Seventy-five percent of U.S. districts offer some online courses (Watson, Murin, Vashaw, Gemin, & Rapp, 2013) and the number of Grade K–12 students enrolled in online courses has been projected to be five million by 2016 (Picciano, Seaman, Shea, & Swan, 2012). Christensen, Horn, and Johnson (2008) predicted that by 2019, half of all U.S. high school enrollments will be online. Credit recovery is one of the fastest growing areas of K–12 online education (Greaves & Hayes, 2008; Picciano & Seaman, 2010), and credit recovery is one of the most common purposes that school districts use online courses (Clements, Stafford, Pazzaglia, & Jacobs, 2015; Murin, Powell, Roberts, & Patrick, 2015; Queen & Lewis, 2011), particularly larger districts (Watson, Pape, Murin, Gemin, & Vashaw, 2014).

In general, online courses deliver content and instruction over the Internet (Murin et al., 2015; Watson & Gemin, 2008). Online courses have a variety of formats and features that can vary according to subject matter and provider. Some are completely online and self-paced; others are models that combine online learning with f2f teacher support for students (Murin et al., 2015; Picciano & Seaman, 2009; Tucker, 2007; Watson & Ryan, 2006).

The promise of online courses for credit recovery may lie in their features and format that make them seem new to students or different from the f2f course they failed (Archambault et al., 2010; Murin et al., 2015). For example, online courses can use technology to engage students with animations, simulations, video, and other interactive content (Murin et al., 2015). Students also receive immediate feedback on activities and assessments, and the pacing of course content can be flexible and individualized (Archambault et al., 2010; Bakia et al., 2013; U.S. Department of Education, 2010). Online courses can present students with rigorous content and some require students to demonstrate proficiency as they move through them and thus skip over material they have already mastered.[1]

States and districts offer online credit recovery courses in a variety of ways, including through state virtual schools, full-time credit recovery charter schools, and district-based programs through which students can take courses at their school, before or after school, during a regular class period, or during the summer (Archambault et al., 2010; Murin et al., 2015). District-based programs can provide convenience, efficiency, and flexibility to schools, who may have many students in need of recovering credit. (See U.S. Department of Education, 2012, for a review of cost–benefit issues related to online learning.)

All indications are that states and districts will continue to make significant investments in infrastructure to provide online courses to students in K–12 settings, yet rigorous evidence of the impact of online credit recovery on student learning and later academic outcomes is distinctly lacking (Davis, 2015). Evidence to inform educators' decisions about whether to offer online credit recovery options to their most at-risk and already disengaged students is especially scarce.

Most of the research on online learning is based at the college level, or on high-achieving students. A recent meta-analysis by Means et al. (2013) reviewed 45 studies of online learning and found that, on average, online instruction yields positive effects relative to f2f instruction (Means et al., 2013). However, this finding was mostly based on postsecondary research; only five articles included in the meta-analysis focused on K–12 education, which together produced a total of seven online/blended vs. f2f comparisons (one experimental, six quasi-experimental). None examined online learning for credit recovery (i.e., among students who had previously failed the course being studied). The average effect across comparisons in the meta-analysis was positive in favor of online/blended learning, but was not statistically significant. Five comparisons favored the online/blended condition, including those comparing an f2f condition to online history lessons for U.S. eighth graders, a science lab program for fifth graders in Taiwan, a writing support program for elementary students in special education classrooms, and, relevant to the present study, an online Algebra I program. This study, conducted by O'Dwyer, Carey, and Kleiman (2007), compared the online program to business as usual f2f algebra instruction and found that students in the online course had higher scores on an end-of-course posttest than students in the f2f classes.

Not included in the meta-analysis, Heppen et al. (2012) conducted the first experimental study of online Algebra I and found that offering Algebra I as an online course was an effective way to broaden access to eighth graders with limited access to Algebra I in mostly rural middle schools. Students who took online Algebra I in Grade 8 instead of a standard eighth-grade mathematics course learned more algebra and were more likely to take and succeed in advanced math courses in high school, even though students who took the standard mathematics courses had substantial exposure to algebraic content (see also Heppen, Clements, & Walters, 2015).

Thus, a small body of literature on the impact of K–12 online learning is emerging and suggests that under certain conditions, online courses can benefit students relative to f2f

[1] There are many models and types of online courses, and some have been criticized for having low or no teacher involvement and few requirements for students to demonstrate proficiency (Murin et al., 2015).

alternatives. However, no studies have yet rigorously examined online credit recovery. School and district administrators who need to provide an opportunity for students to recover credits must decide whether to offer the class in an online format or a traditional f2f format, without any research on which is most likely to improve achievement for students who have previously struggled in the course. The present study was designed to address this gap, with a particular focus on Algebra I.

This study compared outcomes for students randomly assigned to take an online Algebra I course for credit recovery to those for students assigned to a traditional f2f course. Both types of courses were offered in participating CPS high schools during summer 2011 and 2012 to ninth graders who had failed second-semester Algebra I that spring. Although we studied a particular course offered by a single online course provider, the online course represented a common model for online credit recovery (see Bakia et al., 2013 for a review of online Algebra I courses offered by six widely used providers). Both the online and the f2f classes were offered at school, during summer school. The f2f classes were taught by licensed CPS teachers, as usual in the district. Thus, this comparison represents a typical choice that would be made by school administrators implementing credit recovery.

Different Features of Online and F2F Credit Recovery Courses

In establishing a comparison between online and f2f credit recovery courses in this study, it is important to acknowledge that they were expected to differ in multiple ways. That is, each is composed of an "instructional bundle" of features that would be expected to vary by type of course. Most noticeably, the two courses differ in their *mode of delivery*, where the online option has a different format from that of the failed course and the f2f option has the same format as the failed course. The *presentation of material* can be interactive and graphically rich in the online course (but is not necessarily so); in the f2f course the presentation of material is up to the teacher, but would not likely be enhanced by technology to the same extent.

Even though the course title is the same (in the case of this study, second-semester Algebra I), the *content* and *sequencing* may differ. The online course content is standardized with a clear, organized ordering of topics, and in the f2f course, teachers have flexibility in determining course content and sequencing. The *pacing* in the online course is flexible; students can move through at their own pace, while the pacing would be assumed to be generally the same for all students within an f2f class, and varied across f2f classes with different teachers.

The courses also differ in *staffing intensity* with two adults involved with the online course (an online teacher and an in-class mentor, whose roles are defined in the "Online Course Used in the Study" section) and one f2f teacher in the traditional class. With these staffing configurations, the online course can cost more than the f2f course. Online course costs include rates of pay for the in-class mentors, which may be the same as f2f teachers, plus the district- or school-specific rates with a given provider for student "seats" in the course, which include costs for the online teacher as well as the courseware.

Communication between students and online teachers in many online course models is mostly asynchronous, meaning that students and teachers are not necessarily online at the same time and communicate via messages within the course system, with periodic opportunities for synchronous communication for small groups or whole sections of students. In the f2f course, virtually all communication is synchronous. The two types of courses also differ in the ways that *feedback* is provided to students: in the online course, feedback on assignments and assessments is immediate, while in the f2f course, feedback on assessments is likely to be delayed (e.g., at least 24 hours). *Grading* policies may also vary; in-class mentors translate online students' test scores into final course grades, and f2f teachers determine grades in their typical fashion. In general, along many of these features, the online course would be expected to be more uniform across sections than the f2f course, which is subject to classroom teachers' decisions about what and how to teach.

Theory of Action

Figure 1 shows different educational outcomes that could vary for students who take an online or f2f course for Algebra I credit recovery. It begins with a choice point, reflecting a school's decision to offer algebra credit recovery as an online course or in traditional f2f form. At this choice point, the two types of courses are assumed to have the features described in the section above: the online course includes standardized curricular content and an online teacher certified to teach algebra, taken in a brick-and-mortar school with the support of a mentor to provide on-site support. The f2f course is assumed to cover content aligned to the district curriculum taught by a certified teacher with substantial discretion in what is covered.

As depicted in Figure 1, students' classroom experiences, including engagement, classroom personalism, perceptions of academic press (teacher expectations, course difficulty), and class clarity, may vary as a function of taking an online versus an f2f credit recovery course. Advocates for online learning and some of the prior literature (reviewed earlier) suggest that the online course may provide a better experience, due in part to its features of interactivity, flexible pacing, immediate feedback, plus personal support and monitoring from on-site mentors. However, it is also reasonable to postulate that students would benefit more from the f2f class, with an in-person teacher who is able to adjust the content to the needs of the class and address earlier gaps in understanding.

In turn, students' experiences in their credit recovery class will affect whether their math skills improve, as measured by assessments, and also affect their confidence in themselves as math students and their perceptions of math relevance. Those skills and attitudes then, in theory, affect their decisions to take subsequent math classes, and their performance in those classes.

Students' experiences in their algebra credit recovery class will also affect their likelihood of passing and successfully recovering the credit, both directly and indirectly, through improved math skills and mindsets. Successfully recovering credit may indirectly facilitate progress toward graduation through greater motivation and skills to take and succeed in subsequent math courses; recovering credit also directly contributes to progress toward graduation by providing students an additional course credit toward the total math credits required for eventually graduating. Thus, choices about which version of algebra credit recovery to offer could affect students' eventual progress to graduation through (though not limited to) the multiple mechanisms shown in Figure 1.

The Online Course Used in the Study

Prior to this study CPS had piloted different online course providers and was particularly interested in further studying the efficacy of the credit recovery course model offered by Aventa Learning. Aventa's credit recovery course was meant to be taken mainly at school in a supervised setting, as opposed to an "anytime, anyplace" model and included both an

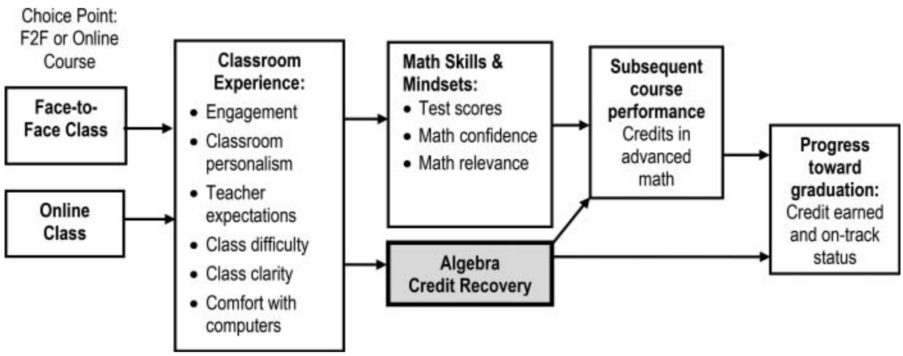

Figure 1. Theory of action for how different credit recovery options influence classroom experiences and subsequent outcomes. Figure shows different outcomes that could vary for students who take an online or f2f course for Algebra I credit recovery.

online teacher and an in-class mentor. The online teachers' role included giving assignments, monitoring student progress, and communicating with students through individual messages and class message boards. The in-class mentors' role included monitoring attendance and classroom behavior, providing technical support, and proctoring assessments; the in-class mentors could but were not required to provide math instructional support directly to students.

The Aventa Algebra I course had multiple strategies and instructional supports intended to meet the individual learning needs of at-risk students. For example, it included an audio "read-aloud" function that targeted vocabulary instruction and comprehension. Small content "chunks" were integrated throughout the course to increase students' retention. Frequent formative assessments offered students feedback on their learning. In addition, lessons included avatars, flash technology, animations, and interactive games to promote student engagement with the content. The course's learning management system allowed students to upload assignments and monitor their own progress.

Aventa (now FuelEducation) operates online courses in every state, and its Algebra I credit recovery course was used in an estimated 500 schools around the country when this study began in 2011. Given its widespread use, a delivery model that appeared particularly appropriate for the purpose of credit recovery, and documented feasibility and satisfaction in CPS, Aventa's online second-semester Algebra I course, was selected for this efficacy study.

Methods

Overview of Study Design

The study design used within-school randomization of first-year high school students who failed second-semester Algebra I (Algebra IB) to either the online or f2f Algebra IB class. The study team invited CPS high schools to participate in the study and provided funding for two or more sections of Algebra IB summer credit recovery. At each of the 17 participating schools, students who had failed Algebra IB in the spring were encouraged to enroll in summer school. Students who enrolled and showed up were randomly assigned to one of the courses. The analyses compare students assigned to the online course to students assigned to the traditional f2f course on their reports of their classroom experiences, measures of math skills and mindsets (algebra posttest scores, reported math liking and confidence), Algebra IB credit recovery success rates, and their subsequent course performance, including total math credits earned by the end of the second year of high school. The next sections describe the sample, recruitment, and random assignment procedures.

Study Sample

Participating Schools

The participating 17 high schools included 11 schools that participated in both 2011 and 2012, four schools that participated in 2011 only, and two schools that participated only in 2012 ($n = 15$ high schools in 2011; $n = 13$ high schools in 2012). Schools were selected and recruited for the study because they had large numbers of students who failed Algebra IB in the prior school year and they were open for summer school. Table 1 details the characteristics of all students (Grades 9–12) in participating high schools as of 2011, alongside those in all CPS high schools during those years. Relative to nonstudy schools, participating schools tended to be larger, had higher algebra failure rates than nonstudy schools, and served a significantly higher proportion of Hispanic students and students for whom English was not the primary home language. In other ways, they were similar to other schools in the district.

Participating Students

All students who failed Algebra IB in participating schools were eligible to participate in the study. The study team supported participating schools to attract students to summer school with outreach activities including meetings with students at schools, mailings from school leaders about the study, and calls to students' homes.

Table 1. Characteristics of participating CPS high schools.

	CPS study high schools ($N = 17$)		All CPS high schools ($N = 86$)
Characteristics	Average *N*	Average percent	Average percent
Female	884	50	51
Race/Ethnicity			
White	153	7	6
African American	607	40	52
Hispanic	931	49	37
Asian	51	2	3
Native American	7	<1	<1
Other Race	21	1	1
Eligible for free or reduced-price lunch	1,490	85	85
Home language not English	934	49	37
Eligible for special education services	248	15	15
Failed second-semester Algebra I[a]	91	23	17
		Mean (*SD*)	Mean (*SD*)
Enrollment (total)		$N = 1{,}770$ ($n = 653$)	$N = 1{,}022$ ($n = 762$)
Enrollment (first-time freshmen)		$N = 339$ ($n = 158$)	$N = 249$ ($n = 193$)

Notes. District-provided records. The study school averages are based on students considered "active" per district records for the participating schools in fall 2011. Eighty-six CPS high schools are included in the district average. District averages include noncharter, noncontract schools that served Grade 9 students in School Year (SY) 2011–12.

[a]The second-semester algebra failure rate is for first-time freshmen in spring 2011 (SY 2010–11), the semester before the study began in summer 2011.

Students who showed up on the first or second day of summer session to enroll in Algebra IB were part of the study sample. The sample totaled 1,224 first-time freshmen across the two summers (2011 and 2012).[2] Participating students were 38% female, 57% Hispanic, 33% African American, 8% White, 2% other races/ethnicities; 86% were eligible for free or reduced-price lunch; 12% were eligible for special education services, and 47% spoke Spanish as their home or native language. In addition to having failed Algebra IB in the prior year, study students generally had weak academic records overall; on average, they failed 4.5 semester courses, and their prior mathematics scores were 0.29 standard deviations below the district average. During the prior year, 40% were suspended and 5% changed schools; students had missed, on average, 30 days of school.

Random Assignment. To protect against the threat to internal validity of "no-shows," only those students who attended the first or second day of summer school were randomly assigned to one of the conditions. Study team members conducted the random assignment procedure on site at each school. Students were blocked for random assignment by gender and whether they passed or failed Algebra IA, based on information collected from each school prior to the first day of the summer session. In total, 613 students were assigned to the online condition and 611 students were assigned to the f2f condition. (See top portion of the CONSORT Diagram included as Figure S1 in the supplementary online materials.)

Baseline Equivalence. To examine potential imbalances at baseline between students in the online and f2f conditions, we calculated standardized mean differences for each characteristic, including student demographic characteristics and ninth-grade reading and mathematics test scores (possible range 1–25) from the EXPLORE assessment. (CPS administers to students the ACT battery of assessments that includes the EXPLORE in Grade 9, PLAN in Grade 10, and ACT in Grade 11.[3]) As shown in Table 2, all of the standardized mean differences between the treatment and control groups were under 0.10 standard deviations, and none of the differences was statistically different from zero at a 95% confidence level, suggesting the random assignment procedure produced two groups that did not differ statistically at baseline.

Participating F2F Teachers, Mentors, and Online Teachers

Study schools identified staff to serve as the f2f algebra teachers and in-class mentors. In-class mentors were certified teachers, but were not required to be licensed in mathematics. Across both cohorts, 63 school staff participated in the study (34 as f2f algebra teachers and 30 as in-class mentors[4]). Eleven teachers and in-class mentors supported more than one course section over the course of the study. Aventa selected a total of six online teachers for the study. All f2f algebra teachers and Aventa online teachers were certified to teach mathematics, compared to 53% of the in-class mentors. On average, the f2f teachers had 14 years

[2] First-time freshmen who failed Algebra IB were our target sample, but students who were not first-time freshmen were not excluded from participating in the summer credit recovery courses. In total, 159 non-first-time freshmen enrolled in a summer credit recovery course offered as part of the study. They were randomly assigned to an online or f2f course, but they are excluded from the analytic sample.

[3] See https://www.act.org/epas/ for information about ACT's Educational Planning and Assessment System (EPAS).

[4] One teacher participated as an f2f algebra teacher in one summer session and as an in-class mentor in another.

Table 2. Baseline covariates by condition.

Covariate	Online			F2F			Standardized mean difference	P
	N	Mean	SD	N	Mean	SD		
Percent Female	613	0.38	0.49	611	0.37	0.48	0.03	0.964
Percent Passed Algebra 1A	613	0.40	0.49	611	0.41	0.49	−0.02	0.470
Percent Unknown Algebra 1A	613	0.20	0.40	611	0.21	0.41	−0.03	0.927
Percent 2012 Cohort	613	0.56	0.50	611	0.58	0.49	−0.03	0.879
Percent Summer Session 2	613	0.68	0.47	611	0.69	0.46	−0.04	0.507
Mean Explore Math Score	532	13.57	2.87	525	13.55	2.93	0.01	0.932
Mean Concentrated Poverty[a]	612	0.05	0.79	609	0.05	0.74	−0.01	0.915
Mean Social Status[b]	612	−0.44	0.87	609	−0.46	0.86	0.02	0.737
Percent Special Education	613	0.12	0.33	611	0.12	0.33	0.00	0.493
Percent Latino	613	0.56	0.50	611	0.58	0.49	−0.04	0.826
Percent Other Race	613	0.09	0.29	611	0.11	0.31	−0.06	0.063
Percent Native Spanish Speaker	613	0.46	0.50	611	0.48	0.50	−0.04	0.714
Percent Suspended	603	0.39	0.49	601	0.41	0.49	−0.04	0.512
Percent Moved Schools	603	0.05	0.22	601	0.05	0.23	−0.02	0.626
Percent Absent 5–9 Days	603	0.14	0.34	601	0.11	0.31	0.09	0.303
Percent Absent 10–14 Days	603	0.10	0.30	601	0.08	0.28	0.06	0.285
Percent Absent 15–19 Days	603	0.12	0.32	601	0.13	0.33	−0.03	0.525
Percent Absent 20–29 Days	603	0.17	0.38	601	0.17	0.38	−0.01	0.569
Percent Absent 30+ Days	603	0.39	0.49	601	0.42	0.49	−0.06	0.554

Notes. Sample includes 17 schools and a total student sample of 1,224 first-time freshmen. Values represent unadjusted means. *P* values are based on an unconditional ordinary least squares regression for continuous covariates and an unconditional logistic regression for dichotomous covariates.

[a]Concentrated poverty is a standardized measure of poverty for the census block group in which the student lives. A large positive number indicates a high level of poverty concentration; a large negative number indicates a low level of poverty concentration. This measure is calculated from census data (the percent of adult males employed and the percent of families with incomes above the poverty line), and is standardized such that a "0" value is the mean value for census block groups in Chicago.

[b]Social status is a standardized measure of educational attainment/employment status for the census block group in which the student lives. A large positive number indicates a high social status; a large negative number indicates a low social status. This measure is calculated from census data (mean level of education of adults and the percentage of employed persons who work as managers or professionals), and is standardized such that a "0" value is the mean value for census block groups in Chicago.

of teaching experience and the in-class mentors had 13 years; in 2011, the Aventa online teachers had five years of teaching experience.[5]

Implementation of the Credit Recovery Courses

The courses were offered as full sections, with all students in the class taking second-semester algebra. The study team randomly assigned students to a total of 76 sections: 18 online and 18 f2f sections in summer 2011 ($N = 36$), and 20 online and 20 f2f sections in summer 2012 ($N = 40$). The course was offered during one or both of two 3- to 4-week summer sessions at each participating school, scheduled for 60 hours for a one-semester course. As described above, the online and traditional f2f courses were expected to vary along a number of features. This section summarizes implementation findings related to these features.

With respect to *mode of delivery* and *presentation of material*, as planned, the online course was delivered via computers within the participating schools, mainly in computer labs. The

[5] Teachers were asked to complete a background survey, including characteristics, qualifications, perceptions of student engagement in their summer Algebra IB class, and grading criteria. Ninety-four percent of the teachers and 93% of the mentors completed the survey. We did not collect background information from Aventa online teachers in 2012 because there were only two involved in the course by the end of the second summer session.

technological delivery of the online course went smoothly except in summer 2012, when the online course provider conducted a system migration just prior to the second three-week session start date. All schools experienced a period of system instability that lasted from a few days to a week, and the problems appeared to be more severe in some schools than others.[6] The f2f course was delivered in traditional classrooms. Because the courses were delivered as planned, the online students experienced a digital curriculum that was more interactive and graphically rich than the f2f students, who, based on our review of f2f classroom materials, experienced a variety of teacher-created and published print materials, including textbooks.

In terms of *content, sequencing*, and *pacing*, Aventa's Algebra IB credit recovery course consisted entirely of content considered typical of second-semester algebra. Organized into five units, the topics included systems of equations, polynomials, quadratics and radicals, rational expressions, and exponentials. Students had some flexibility in how they progressed through topics but they were strongly encouraged to move sequentially. On average, students completed about 2.5 of the five units of the online course, and on average, passed 1.7 of the five unit exams.[7] These figures were consistent across both years and sessions within years.[8] By collecting classroom materials from the f2f teachers in summers 2011 and 2012, we observed that the f2f classes covered a mix of pre-algebra and first- and second-semester algebra topics that were not presented in a uniform sequence across the sections with different degrees of coherence.[9] Our analyses of classroom materials across both summers indicated that the f2f classes had, on average, a 50% focus on second-semester algebra content and a 50% focus on first-semester algebra and pre-algebra topics. We also observed that the topics in 28% of the f2f classes were not organized in a coherent sequence. Because student-level content exposure in the f2f classes cannot be closely tracked as it can be in the online course, it is not clear how much of the content covered in their classes was completed by each student.

With respect to *staffing*, both the online and f2f classes were taught by certified math teachers—an online teacher hired by Aventa for the online course and licensed CPS instructors for the f2f classes. As planned, all online classes also had a site-based mentor to support and monitor students. Prior to summer school, in-class mentors attended a training session on how to use the online course, monitor student progress, and communicate with the

[6] Problems cited by students in schools included trouble logging in, being kicked out of the system, not being able to save quiz responses, and not being able to select answers when taking a test. In-class mentors reported not being able to access student assessments or grades. By the beginning of the second week of Summer Session 2, the system problems had subsided for most, but not all, schools. By the middle of the second week, all problems were resolved.

[7] The study team collected archival data from the Aventa course system, including the amount of time students were logged in, number of quiz attempts and grades per quiz attempt (percent correct), unit exam grades (percent correct), and cumulative and final grades.

[8] Based on this observation and examination of other finer grained archived course data, the study team determined that the technology problems in Summer Session 2 of summer 2012 did not detectably hinder student progress relative to the other, smoother implementations of the course during the other summer sessions in the study.

[9] We collected classroom materials from the f2f teachers to describe the proportion of time spent on specific algebra topics in these classes. The materials included annotated tables of contents from the algebra textbook, detailed syllabi, and collections of materials assembled or generated by teachers. We used the topics in Aventa's second-semester course to frame our analyses of the proportion of pre-algebra, first-semester algebra, and second-semester algebra content in the f2f classroom materials. Because the types of classroom materials varied across teachers, we used the most fine-grained unit available within each set of classroom materials to calculate the relative content emphasis. For example, in cases where teachers annotated each lesson within chapters of a textbook, we analyzed the content at the lesson level, unless otherwise noted. In cases where teachers submitted every handout they used in the course, we analyzed the content at the handout level. We defined coherent sequencing as topic progression that followed what might be seen in a typical textbook or district pacing guide, where pre-algebra, Algebra IA, and Algebra IB topics are presented sequentially. Materials for which topics did not follow this order—e.g., we were not able to determine a logical sequencing of how the pre-algebra, Algebra IA, and Algebra IB topics were assembled—were coded as having incoherent sequencing.

online teachers, who received ongoing professional development and supervision from Aventa. Both the in-class mentors and f2f teachers were paid their regular teaching rates, and the online course provider was paid for student "seats" in the course, therefore the relative costs were higher for the online than the f2f classes.[10]

Communication in the online course was mostly asynchronous. Online teachers communicated with students through the learning management system, online chats, and online "whiteboard" demonstrations. The online course also had an *Elluminate Live!* platform that enabled synchronous teacher-to-student and student-to-student communication. Communication between the f2f teachers and students was synchronous and in-person, as expected. In terms of in-class mentors' actual roles, analysis of their daily logs indicates that they spent, on average, 28% of the 60-hour course proctoring exams, transferring student grades, handling scheduling or other administrative issues, and addressing student behavioral problems; 16% of their time providing or contacting technology support; 2% of their time communicating with the online teacher; and though not required of them, 17% of their time providing math instructional support to students.

Finally, with respect to *feedback* and *grading*, the online course gave students immediate feedback on practice items and assessments, while in the f2f course, the timing of feedback varied by teacher but was not immediate. To progress through the online course, students had to pass each quiz and unit test with a score of 70% or higher. The system housed both grade-to-date and cumulative grade information that was accessible to the in-class mentors (as well as the online teacher and students themselves). The mentors translated these course data into final grades. When asked how they did this, mentors, on average, reported that students' test and quiz performance counted for about 60% of their final grade; the other 40% was based on other factors including effort, completion of classwork, and behavior. F2f teachers, on average, reported that quizzes and tests counted for just over half (53%) of students' grades; the rest was based on other factors including effort, work completion, and behavior.

These observations of how the two types of courses were delivered provide the context for the impact analyses comparing outcome for students by condition. These analyses are presented following a description of data collection, measures, and the analytic strategy.

Data Collection

To describe the implementation of the credit recovery courses reported in the previous section, we collected extant usage data from Aventa reflecting students' progress through the online course, in-class mentor logs, classroom materials from f2f teachers, and end-of-course surveys for f2f and online (Aventa) teachers and in-class mentors.

The main source of outcome data was extant data provided by the district, which included credit recovery course grades and pass rates for all students in the study, math scores on the PLAN assessment taken by students in the fall of their second year of high school (if they were considered tenth graders), and course grades in the second year of high school for students who remained in school within the district. The study team also administered an algebra posttest and end-of-course survey to all students who persisted to the end of the course. Working in collaboration with the teachers in participating schools, we obtained consent forms from 88% of students; of those, 90% indicated agreement to participate in the survey and posttest (78% of the

[10]This is not always the case; see U.S. Department of Education (2012) for a discussion of cost–benefit issues in online learning.

Table 3. Student-level data collection summary overall and by condition.

	Overall ($N = 1{,}224$)		Online ($N = 613$)		F2F ($N = 611$)		
Data collected	Number	Percent	Number	Percent	Number	Percent	p
Consent Returned	1,071	88	543	89	528	86	0.253
Consent Affirmed	959	78	506	83	453	74	<0.001
Student Posttest and Survey	838	68	441	72	397	65	0.009

Note. Study records.

full sample of study participants). Some of those students, however, did not persist to the end of the course, and so they did not complete the survey or posttest. Study staff administered the posttest (on paper) to students in the online and f2f course sections on the penultimate day of the summer session. Students had 50 minutes to complete the posttest, after which students were asked to complete the survey. Thus, response rates for the survey and the study posttest were identical—68% across conditions. As shown in Table 3, these rates differed by condition (72% for online; 65% for f2f). Our approach to handling missing data is described in the "Analytic Strategy" section. (See CONSORT diagram in Figure S1 in the supplementary online materials, for information about data collection rates for the randomized sample.)

Measures

Table 4 presents student survey measures, Rasch person reliability, and example items.

Table 4 Survey measures.

Construct	N items	Person reliability	Example items
Engagement	4	0.78	"The topics we studied were interesting and challenging." "I worked hard to do my best in this class."
Class Difficulty	4	0.81	"I found the work challenging." "I had to work hard to do well in this class."
Teacher Expectations	4	0.60	"The teacher expected me to do my best all the time." "The teacher expected everyone to work hard."
Classroom Personalism	7	0.86	"The teacher really listened to what I had to say." "The teacher was willing to give extra help on work if I needed it." "The teacher gave me specific suggestions about how I could improve my work in this class."
Class Clarity	4	0.78	"It was clear what I needed to do to get a good grade." "The work we did in class was good preparation for the quizzes and tests in this class."
Usefulness of Mathematics	5	0.84	"I think learning math will help me in my daily life." "I need to do well in mathematics to get into the college or university of my choice."
Liking/Confidence in Mathematics	7	0.90	"Mathematics is easier for me than for many of my classmates." "I enjoy learning mathematics."
Comfort With Computers	6	0.76	"I feel comfortable working with a computer." "I enjoy lessons on the computer."

Notes. For Teacher Expectations and Classroom Personalism, students in the online course were asked each item with respect to both their in-class mentor and their online teacher. For each item, the higher response of the two was the one taken to calculate the composite scale score, under the assumption that students' instructional experience would be most influenced by the adults whom students perceived to have higher expectations or greater classroom personalism.

Classroom Experience Measures

Students indicated their agreement on a 4-point scale (0 = *Strongly Disagree;* 3 = *Strongly Agree*) with items about their experience in their credit recovery class. Responses to items were Rasch-scaled for analysis. These measures—Engagement, Class Difficulty, Teacher Expectations, Classroom Personalism, and Class Clarity—are regularly used in surveys of CPS students.[11] One measure—Comfort with Computers—was adapted from a study of online algebra in middle school (Heppen et al., 2012).

Math Mindsets Measures

Two survey measures using the same 4-point scale—Usefulness of Mathematics and Liking/Confidence in Mathematics—were similarly adapted from Heppen and colleagues (2012).

Table 5 provides a brief description of all other measures used as baseline covariates or outcomes, including details about how each measure was constructed.

Baseline Covariates

We collected student characteristics and scores on the EXPLORE mathematics assessment administered by schools in Grade 9 for use as baseline covariates. See Table 2 for a list of covariates and descriptive statistics.

Math Skills Measures

We measured students' math skills at the end of the summer credit recovery courses with an algebra posttest administered by the study team. The study posttest included 28 items from the National Assessment of Educational Progress (NAEP) that addressed pre-algebra and first- and second-semester algebra topics. The posttest was for study purposes only and was not factored into students' course grades. We measured students' math skills in the second year of high school by collecting their scores on the PLAN mathematics assessment administered by the district in the fall of Grade 10. The PLAN is part of the battery of ACT's standardized, nationally normed "college and career readiness" assessments. The algebra subtest has 22 items that address pre-algebra and first- and second-semester algebra topics. The composite math score combines scores from the algebra subtest and a geometry subtest, for a total of 40 items.

Algebra Credit Recovery

Of primary interest, we examined whether students successfully recovered credit in the summer course, determined by a course grade of D or higher. This outcome is considered primary both because of its important role in the theory of action (Figure 1) and because this measure was based on extant data for all participating students, with no missingness due to course completion or other factors (see "Missing Data Strategy" section).

Subsequent Course Performance

We coded whether students earned course credit in a geometry or a higher mathematics course in their second year of high school, separately for the first and second semesters. In CPS, students received credit if they earned a D or higher.

[11] For more information on these surveys and measures, see "Survey Documentation," http://ccsr.uchicago.edu/surveys/documentation

Table 5. Baseline covariates and academic outcome measures.

Measure	Description	Measurement construction
Student Characteristics	Student background characteristics were collected from CPS records and converted to dichotomous variables for analyses. Concentrated poverty and social status are standardized measures of poverty and educational attainment/employment status, respectively.	Ethnicity indicators; native Spanish speaker = 1; receives special education services = 1; and in Grade 9: ever suspended = 1, moved schools = 1, and attendance dummy indicators.
Baseline EXPLORE Mathematics Assessment	CPS administered the EXPLORE mathematics assessment in the fall of ninth grade as part of the ACT battery of assessments (EXPLORE, PLAN, ACT).	Mathematics scaled scores can range from 1–25.
End-of-Course Posttest[a]	Students completed a 28-item algebra posttest at the end of the summer course. The scaled scores were standardized to the mean and standard deviation for algebra NAEP items in Chicago ($M = 276$; $SD = 35$).	Students' item-level accuracy was Rasch-scaled, separately by cohort, to produce an overall scale score. (Person Reliability = 0.70)
PLAN Mathematics and Algebra Assessment	CPS administered the PLAN assessment in the fall of tenth grade as part of the ACT battery of assessments (EXPLORE, PLAN, ACT).	Composite mathematics scaled scores can range from 1–32. Algebra subtest scaled scores can range from 1–16. (Estimated reliability = 0.80 for both math and algebra; ACT, 2013)
Credit Recovery	Credit recovery was based on the grades students earned in their summer Algebra IB course. Students recovered their credit if they earned a grade of A, B, C, or D, and did not recover credit if they earned an F or dropped the course.	0 = did not recover credit; 1 = recovered credit
Subsequent Course Performance	In the second year of high school, students received credit in geometry or a higher mathematics course if they earned a D or higher in the course, and did not receive credit if they earned an F or did not take geometry or a higher course that semester. If students failed their first-semester course and retook the course alongside their second-semester math course, we used the higher of the two grades for analysis. If students were missing course grade data and they had a leave code indicating they transferred outside the district or to a charter school, their course data were treated as missing. Any students who were unconfirmed transfers were assumed to have not earned credit in geometry or higher course.	We coded whether students earned course credit in geometry or a higher mathematics course separately for the first and second semesters. Students who earned a D or higher recovered credit (0 = did not earn credit; 1 = earned credit).
Cumulative Math Credits Earned	We calculated cumulative credits earned in mathematics courses by the end of the following school year (2011–12 for Cohort 1; 2012–13 for Cohort 2).	Total number of semester credits earned in mathematics courses (range 0 to 4 credits)
CPS "On-Track" Indicator	The on-track indicator was developed by CCSR and is used on CPS school report cards. Students are categorized as on track for graduation if they earned at least 11 full-year course credits (22 semester credits) during their first two years of high school and had no more than one semester course failure in their second year of high school.	0 = off track; 1 = on track

Note. [a]The end-of-course posttest was administered by the study team for the purpose of the study. All other measures in the table were collected via district administrative records.

Table 6. Number and percent of nonmissing observations for student outcomes based on extant data: Overall and by condition.

	Overall		Online		F2F		
Outcome	*N*	Percent	*n*	Percent	*n*	Percent	*P*
Summer Course Grade	1,224	100	613	100	611	100	—
Algebra Credit Recovery	1,224	100	613	100	611	100	—
PLAN Math	878	72	442	72	436	71	0.772
PLAN Algebra	878	72	442	72	436	71	0.772
Credit in Geometry or Higher Course (Semester 1)	1,120	92	563	92	557	91	0.669
Credit in Geometry Higher Course (Semester 2)	1,056	86	530	86	526	86	0.850
Cumulative Math Credits	1,015	83	512	84	503	82	0.577
On-Track Indicator	1,015	83	512	84	503	82	0.577

Notes. Sample includes 17 schools and a total student sample of 1,224 first-time freshmen (n = 613 online; n = 611 f2f). *P* values are based on an unconditional logistic regression. The primary reasons for missing data differed by outcome. For the PLAN assessments, students typically only took the test if they were officially designated as a 10th grader in CPS by their second year of high school. For credit in geometry or higher, students were classified as missing if they were not enrolled in CPS during that semester. For cumulative math credit and the on-track indicator, students were classified as missing if they transferred to, or transferred from, a non-CPS school (including charter schools) at any point during the first two years of high school.

Progress Toward Graduation

Following the theory of action, we examined cumulative math credits earned and students' "on-track" status (using Chicago's indicator; see Table 6) at the end of their second year of high school.

Analytic Strategy

To test the impact of taking an online versus f2f Algebra IB credit recovery course, we modeled schools, cohort (2011 or 2012) and summer session (1 or 2) as fixed effects to account for the unique effects of schools, cohorts, and summer school sessions. The model also included the student-level characteristics shown in Table 2 to improve precision. All predictors, with the exception of the treatment indicator, were centered around their grand mean. For continuous outcomes, the impact model is as follows:

$$\begin{aligned} Y_i = \beta_0 + \beta_1(Treatment)_i + \beta_2{}^*(Student\ Characteristics)_i \\ + \beta_3{}^*(Prior\ Achievement)_i + \beta_4{}^*(School)_i + \beta_5{}^*(Cohort)_i \\ + \beta_6{}^*(Summer\ Session)_i + e_i \end{aligned} \tag{1}$$

where Y_i is the outcome (e.g., posttest score) for student i, β_0 is the average student outcome, β_1 is the impact of the online course relative to the f2f course, $StudentCharacteristics_i$ is a vector of demographic characteristics (e.g., ethnicity, gender, special education status), $PriorAchievement_i$ is Grade 9 math achievement (EXPLORE) for student i, $School_i$ is a vector of binary indicators representing the fixed effects of each school for student i, $Cohort_i$ is a binary indicator for whether student i participated in summer 2012 (rather than 2011), $SummerSession_i$ is a binary indicator for whether student i participated in Summer Session 2 (rather than Summer Session 1), and e_i is the random error associated with the estimated outcome for student i. For binary outcomes (e.g., credit recovery) and ordinal outcomes (e.g., credits earned), we used logistic and ordered logistic regression models, respectively, with the same covariates.

Missing Data Strategy

For cases with missing covariate data, we included missing data indicators in the impact model. Less than 2% of students were missing covariate data, with the exception of prior achievement (EXPLORE scores), where 14% of students had missing values. Because treatment assignment was randomized, missing covariate data was not associated with treatment condition (e.g., the percent of students missing prior achievement data was almost identical in the treatment and control groups) and the missing data should not bias impact estimates (Jones, 1996). For cases with missing outcome data, we applied inverse probability weights (IPW) to the impact models (Ridgeway, McCaffrey, Morral, Griffin, & Burgette, 2013; Wooldridge, 2007) to account for potential nonresponse bias and improve our ability to generalize results to the target study population. The IPW were based on students' predicted probabilities of having data for a given outcome, where predicted probabilities were estimated with a generalized boosted regression (McCaffrey, Ridgeway, & Morral, 2004).[12] A separate boosted regression was run for each student cohort and data collection source (student survey, end-of-course assessment, PLAN assessment, and mathematics course completion in the year following the summer credit recovery course). We did not apply IPW to outcome models in which less than 1% of students had missing outcome data (e.g., credit recovery).[13]

Missing data rates for student outcomes were 32% for student survey and posttest outcomes (see Table 3) and ranged from 0–28% for student outcomes based on extant data (see Table 6). Differences in missing data rates between treatment and control groups were less than two percentage points for the outcomes based on extant data, and seven percentage points for the student survey and posttest outcomes. Tables S1–S3 in the supplementary online materials show characteristics of students included in each of the outcome analyses (i.e., had nonmissing data for the impact analysis for that outcome), by condition. Across all the observed baseline covariates, treatment and control students included in the impact analysis for a given outcome had similar baseline characteristics, on average. Standardized mean differences for most covariates were less than 0.05 standard deviations (with the largest difference being 0.13 standard deviations), and none of the differences were statistically significant. The inclusion of the baseline covariates in the impact model provides additional assurance that observed baseline differences in the treatment groups will not bias the impact estimates.

Results

Table 7 shows the impact of online versus f2f algebra credit recovery on students' class experiences, math skills and mindsets, credit recovery course outcomes, subsequent math course performance, and progress toward graduation. Results for each student outcome are summarized in the sections that follow. This section follows the order of the proximal and distal

[12] This method iteratively tests various combinations of student background covariates (see Table 2), including school fixed effects and treatment group indicator, to predict the probability of response, searching for the combination that minimizes the differences in measured characteristics between students with valid data and those with missing data. We used the *twang* package in the *R* statistical program to execute the generalized boosted regression. Following the recommendations set forth by the package authors (Ridgeway et al., 2013), we set the interaction depth to 4, shrinkage to 0.0005, and bagging to 0.50.

[13] To examine whether the results were robust to our missing data approach, we also ran all impact models using multiple imputation. The results using the multiply imputed data sets were consistent with the main analyses using IPW presented in the next section.

Table 7. Impact of online versus f2f summer course on student survey measures, summer course student outcomes, and second-year student outcomes.

Outcome	N	Online mean[a]	F2F mean[b]	Effect estimate[c]	Standard error[d]	p	Effect size[e]
Class Experience Measures[f]							
Engagement	832	0.34	0.60	−0.26	0.18	0.150	−0.10
Teacher Personalism	836	3.14	3.04	0.09	0.22	0.673	0.03
Academic Press: Teacher Expectations	836	5.95	6.04	−0.09	0.35	0.806	−0.02
Academic Press: Class Difficulty	829	2.07	0.37	1.69	0.24	0.000	0.51
Class Clarity	838	1.53	3.60	−2.07	0.23	0.000	−0.64
Comfort With Computers	813	0.43	−0.16	0.59	0.11	0.000	0.35
Math Skills Measures[f]							
End-of-Course Algebra Posttest	838	272.97	279.62	−6.64	2.09	0.002	−0.19
PLAN Composite Math Score	878	14.16	13.94	0.23	0.18	0.213	0.08
PLAN Algebra Subtest Score	878	5.42	5.27	0.15	0.13	0.240	0.07
Math Mindsets Measures							
Usefulness of Mathematics	824	1.32	1.34	−0.02	0.20	0.921	−0.01
Liking/Confidence in Mathematics	829	−0.76	−0.22	−0.54	0.20	0.007	−0.18
Algebra Credit Recovery	1,224	66%	78%	0.56	0.08	0.000	−0.35
Subsequent Math Course Performance							
Credit in Geometry or Higher (Summer Semester 1)	1,120	53%	54%	0.96	0.13	0.772	−0.02
Credit in Geometry or Higher (Summer Semester 2)	1,056	47%	48%	0.96	0.13	0.783	−0.02
Progress Toward Graduation							
Cumulative Math Credits[g]	1,015	2.39	2.51	−0.12	0.07	0.089	−0.10
On-Track Indicator	1,015	28%	25%	1.14	0.18	0.403	0.08

[a]Means reported for students in the online course are observed means.
[b]Means reported for students in the f2f condition are model-adjusted, calculated by subtracting the effect estimate from the observed online group mean. See Table S4 in the supplemental online appendix for observed means and standard deviations for both the online and f2f conditions.
[c]Odds ratios from logistic regression models are reported for binary outcomes.
[d]Standard errors are robust standard errors due to inverse probability weights.
[e]Effect sizes are presented as Cohen's *d*, calculated using the pooled standard deviation for continuous measures; a Cox index (odds ratio/1.65) is reported for binary outcomes.
[f]Means presented represent Rasch-scaled scores, not observed responses on the 4-point scale.
[g]Results are based on an ordinary least square regression model for ease of interpretation but also were tested using an ordered logit regression model, given that semester credits are not continuous; results did not differ.

outcomes shown in the theory of action (Figure 1), although as noted in the "Measures" section, successful recovery of second-semester Algebra I credit is considered the primary outcome for the study.

Classroom Experiences

Students in the online and f2f courses did not differ significantly at the end of the course in either their self-reported engagement or the extent to which they felt they received personalized support. (See Table 7 for impact results based on Rasch-scaled scores.) They also did not differ in perceived teacher expectations; however, students in the online course perceived the course to be significantly more difficult than students in the f2f course ($d = 0.51$). Students in the online course also reported experiencing significantly less clarity about what they needed to do to succeed in the class than students in the f2f course ($d = -0.64$). The one area in which students in the online section reported a more positive experience was comfort with computers; students in the online course reported more comfort and enjoyment when using computers in a classroom at the end of the summer session than their f2f counterparts ($d = 0.35$).

Math Skills and Mindsets

Posttest

Students in the online course had significantly lower scores on the end-of-course posttest relative to their f2f counterparts ($d = -0.19$; see Table 7). The scaled scores presented in Table 7 translate to an accuracy rate of 38% for students in the online course and 40% for students in the f2f course. Although we do not have scores on this specific combination of items from a broader population of students to contextualize the scores of students in the study, these accuracy rates suggest that students in both conditions generally performed poorly on the posttest. Given the differences in course content between the online and f2f classes, we also examined whether students in the online and f2f conditions performed differently on the pre-algebra, Algebra IA, and Algebra IB items on the posttest. The results were consistent with the overall effect: online students' scores were significantly lower than f2f students' scores on the pre-algebra ($\beta = -0.27$, $SE = 0.11$, $p = 0.013$, $d = -0.16$); Algebra IA ($\beta = -0.20$, $SE = 0.07$, $p = 0.003$, $d = -0.19$); and Algebra IB ($\beta = -0.23$, $SE = 0.09$, $p = 0.014$, $d = -0.16$) posttest items.

PLAN

Despite observing significant differences favoring f2f students on the algebra posttest administered only months earlier, we found no statistically significant differences by condition in performance on the PLAN mathematics assessment or algebra subtest (see Table 7). Scores for both groups of students were low. (Means on the algebra subtest, on which scores can range from 1 to 16, were 5.42 and 5.27 for the online and f2f groups, respectively. Means for composite math, which can range from 1 to 32, were 14.16 and 13.94 for online and f2f, respectively.)

Math Mindsets

Students in the online and f2f courses did not differ in their perceptions of the usefulness of mathematics for their future or daily life activities. However, online students reported significantly lower liking of math, and confidence in their mathematical skills, than students in the f2f course ($d = -0.18$; see Table 7).

Algebra Credit

Of primary interest to this study is whether students assigned to an online vs. f2f course would have different credit recovery rates. Reflecting students' self-reports that the online course was more difficult than the f2f class, we found that students randomly assigned to the online class were less likely to successfully recover the credit than students assigned to the f2f class. Across conditions, more than two thirds (71%) of study participants successfully recovered credit. However, there was a 12-percentage-point difference in credit recovery rates by condition: 66% of students in the online course in contrast to 78% of students in the f2f course successfully recovered credit as part of the study ($d = -0.35$; see Table 7). Because this was a primary outcome for the study, we further examined whether students in the online course whose mentors were certified math teachers were more likely to successfully recover credit than students whose mentors were not. We found that the credit recovery rates for students in sections with and without a math-certified mentor were similar—68% and 65%, respectively.

As noted in the "Measures" section, the basis for whether students recovered credits or not was the grade they earned in the course—students recovered credit with a grade of D or higher. Figure 2 graphs the observed percentage of students earning different grades in the

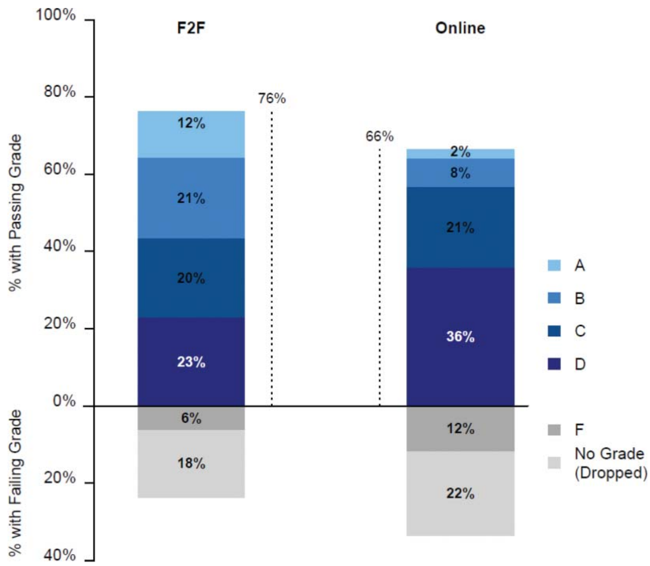

Figure 2. Summer session course grades and credit recovery rates by condition. Figure shows the percent of students in each condition who earned grades of A, B, C, D, F, or no grade in the credit recovery course to which they were randomly assigned as part of the study. Students who received no grade had dropped the course before completion. Group percentages are observed, not model-adjusted. Sample sizes are $n = 611$ f2f students; $n = 613$ online students.

course and descriptively shows that the grades in the f2f class were generally higher than grades in the online course. For example, only 31% of students in the online course, in contrast to 53% of students in the f2f course, earned an A, B, or C.

Subsequent Course Performance

Although we found that students in the f2f course had higher posttest scores and were more likely to recover Algebra IB credit than students in the online course, there were no statistically significant differences by condition in the rates at which students earned credit in geometry or a higher mathematics course in either semester of their second year of high school (see Table 7). About half of the study students in both conditions earned credit in geometry or a higher mathematics course the year after they participated in the summer credit recovery course.

Progress Toward Graduation

Cumulative Math Credits

Because we found that students in the f2f course were significantly more likely to recover Algebra IB credit than students in the online course, we might expect a difference by condition, in favor of f2f students, in total math credits earned toward graduation by the end of the second year of high school. On average, students in the online credit recovery course had 2.39 math credits by the end of their second year of high school, while students in the f2f credit recovery course had 2.51 math credits ($d = -0.10$, see Table 7). This difference was not statistically significant by conventional standards ($p = 0.09$).

On-Track Indicator

Finally, we found no difference by condition in students' likelihood of being on track for graduation at the end of the second year of high school—a measure that takes into account course performance and credits earned in all courses, not just math. Twenty-six percent of students in the online course and 23% of students in the f2f course were on track for graduation at the end of their second year of high school (see Table 7).

Discussion

Online courses are increasingly used for credit recovery in districts across the country, yet there has been little research about the effectiveness of these courses for students who have previously failed. Despite great optimism about the potential for online learning to provide more flexibility and interactive instruction to students, this study suggests that online credit recovery can yield worse outcomes for students than standard, f2f summer classes, at least in the short-term. Relative to f2f, the online option had clear disadvantages with respect to key proximal outcomes—including students' recovery rates and some measures of math skills and mindsets. Students who took the online Algebra IB course had lower credit recovery rates, lower scores on an end-of-course algebra assessment, and less confidence in their mathematical skills than students who took an f2f credit recovery class. Although this study alone cannot determine whether another online course provider's model would yield different results, this model was chosen because it was in wide use and had similar features to other widely used courses.

The online credit recovery course was also selected because it was perceived, based on prior experience in the district, to provide appropriate supports for struggling students. These supports did not seem sufficient for students to perform as well as students in classes that provided f2f instruction. The f2f teachers had discretion over what and how to teach, including the content, and our analyses showed that the f2f teachers spent time on remedial topics and first-semester Algebra I topics, even though the course title was second-semester Algebra I. By contrast, the online course covered exclusively second-semester Algebra I content. For students who had failed their prior algebra class, many of whom had weak mathematical skills, the online class may have simply been too hard. By including content from first-semester algebra and pre-algebra, the f2f teachers may have provided instruction on topics for which students were ready and could understand. Moreover, the f2f classes may have been more likely to help establish foundational algebraic knowledge, which many of the students in the study needed, given the high rates of failure of first-semester algebra (50%).

These findings underscore a tension between ensuring that credit recovery courses meet rigorous content standards and the need to remediate students' prior knowledge and skills. The online course tested in this study had rigorous content but lacked the flexibility to address gaps in students' prior skills and understandings, and thus may not have met students' individual needs. More adaptive online course models designed to assess prior knowledge and scaffold learning accordingly may hold promise for supporting struggling students in need of credit recovery. Research on Carnegie's Cognitive Tutor, for example, suggests that an adaptive online algebra program, combined with f2f instruction, can support student learning (e.g., Pane, Griffin, McCaffrey, & Karam, 2014), but this approach has not yet been rigorously tested for credit recovery. Given the limited time frame in which many credit recovery courses are implemented, any effort to remediate students' prior knowledge may shortchange the content covered—in either an f2f or an adaptive, more personalized online credit recovery course. Future research is needed to address this question and to examine if and how online courses can address the diverse learning needs of struggling students, both to prevent failure and provide the supports needed to help students get back on track.

The struggles of students in the credit recovery classes to master second-semester algebra content highlights a substantial problem for remediating students with weak mathematical skills. Even among students who showed up every day for credit recovery and successfully passed and received credit, scores on the end-of-course algebra assessment were low. On average and across conditions, students answered just more than one third of the questions correctly. On the PLAN mathematics assessment taken in the fall of the following year, these students scored an average of 14, which put them five points behind the ACT assessment system's benchmark score of 19—a difference that is equivalent to about five years' worth of growth.[14] Perhaps end-of-course achievement would have been higher if the online and f2f courses had been designed to reflect the latest research on algebra teaching and learning, which suggests that students should have ongoing opportunities to analyze solved problems and to choose from alternative algebraic strategies (Star et al., 2015). Neither the online or f2f course seemed to emphasize these types of activities. But even if either course had been designed to reflect this research, it might nevertheless seem an impossible task for a three- to four-week summer class to catch students up to the standards that are currently used to judge whether students are making sufficient progress.

Nevertheless, the majority of students in both conditions successfully recovered their Algebra IB credit, and despite the differences in favor of f2f on proximal outcomes, we detected no statistically significant differences in outcomes measured during the following academic year. Scores on the PLAN mathematics assessment and PLAN algebra subtest, administered within months of summer school, were not significantly different by condition. The discrepancy between these results and those for the study-administered algebra posttest may be the result of differences in content on the PLAN assessment and the study posttest, which was composed of released NAEP items on algebraic skills and concepts. However, both the study posttest and the PLAN algebra subtest covered pre-algebra, Algebra IA, and Algebra IB topics, and the general alignment between these two assessments is assumed to be high.[15] Still, it is possible that the study posttest was more sensitive to differences between the online and f2f student's algebra knowledge. It is also possible that the differences by condition that were detectable at the end of the summer course were attenuated soon after.

Analyses of other second-year outcomes also showed no differences by condition in students' likelihood of earning credit in geometry, being on-track for graduation, or in total math credits earned during the first two years of high school. Thus, the second-year findings may suggest that the initial negative effects of online relative to f2f credit recovery in Algebra I dissipate. However, this conclusion may not be correct. Other than PLAN assessment scores, the second year outcomes are rather coarse measures that may not be as sensitive to the differences in algebra knowledge between online and f2f credit recovery as the proximal measures (end-of-course algebra scores and whether students recovered the course credit). In addition, effects may resurface later. For example, although there was not a statistically significant difference by condition in total math credits at the end of the second year of high school, the difference was marginally significant in favor of f2f students and this gap may widen over time. Moreover, it is possible that some of the

[14] On average, students gain about one point per year on the EPAS assessments.

[15] It is impossible, however, to calculate the exact degree of alignment because the PLAN is a secure test and item-level detail is not provided.

attitudinal differences observed at the end of the summer credit recovery courses—particularly in students' liking of and confidence in mathematics—could persist into subsequent years and affect students' later outcomes.

If in further analyses we similarly observe no difference by condition on later (third and fourth year) academic outcomes, the overall findings of this study may suggest that offering online courses is a viable option for credit recovery, essentially no worse than an f2f option in the long run. This may be an appealing interpretation to some education decision makers, for whom online courses undoubtedly offer convenience and provide increased flexibility. When offering online courses, schools do not need to know how many students will enroll ahead of time, completely fill a section with students, or find qualified teachers for each course that students need to recover. However, further analysis of student outcomes related to course performance, credit accrual, and eventual graduation are needed to confidently conclude that this study provides evidence of initial negative effects of online relative to f2f credit recovery, with no longer term differences. Subsequent papers on this study will report these findings.

At this time, the interim findings provide some important cautions about providing online courses to at-risk, generally low-achieving students, particularly courses with relatively rigorous content and, within the time constraints of summer credit recovery, few opportunities for remediation. Over time the results of this study will build and continue to shed light on important questions regarding the long-term impact of online versus f2f credit recovery for students who fail algebra. Until future data become available for analysis, the first two years of this study offer important initial findings and emerging lessons regarding the choice between online and f2f credit recovery courses faced by school districts and high schools around the country as they work to help struggling students get back on track.

Funding

This research was supported by Grant R305A110149 from the Institute of Education Sciences, U.S. Department of Education to the American Institutes for Research. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

ARTICLE HISTORY

Received 22 May 2015
Revised 10 February 2016
Accepted 15 March 2016

References

Allensworth, E. M., & Easton, J. (2005). *The on-track indicator as a predictor of high school graduation*. Chicago, IL: Consortium on Chicago School Research.

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago Public High Schools: A close look at course grades, failures and attendance in the freshman year*. Chicago, IL: Consortium on Chicago School Research.

Allensworth, E. M., 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.

Archambault, L., Diamond, D., Coffey, M., Foures-Aalbu, D., Richardson, J., Zygouris-Coe, & Cavanaugh, C. (2010). *INACOL Research Committee Issues Brief: An exploration of at-risk learners and online education*. Vienna, VA: International Association for K–12 Online Learning.
Atkins, D., Brown, J. S., & Hammond, A. (2007). *A review of the Open Educational Resources (OER) movement: Achievement, challenges and new opportunities* (Report for the William and Flora Hewlett Foundation). Retrieved from http://www.hewlett.org/uploads/files/ReviewoftheOERMovement.pdf
Bakia, M., Mislevy, J., Heying, E., Patton, C., Singleton, C., Krumm, A. (2013). *Supporting K–12 students in online learning: A review of online Algebra I courses*. Menlo Park, CA: SRI International.
Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disrupting innovation will change the way the world learns*. New York, NY: McGraw-Hill.
Clements, M., Stafford, E., Pazzaglia, A. M., & Jacobs, P. (2015). *Online course use in Iowa and Wisconsin public high schools: The results of two statewide surveys* (REL 2015–065). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies.ed.gov/ncee/edlabs.
Davis, M. R. (2015, September). Online credit recovery, in need of improvement, study says. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2015/09/23/online-credit-recovery-in-need-of-improvement.html.
Greaves, T. W., & Hayes, J. (2008). *America's digital schools 2008: Six trends to watch*. Encintas, CA: The Greaves Group.
Ham, S., & Walker, E. (1999). *Getting to the right algebra: The Equity 2000 initiative in Milwaukee Public Schools*. New York, NY: MDRC. Retrieved from http://www.mdrc.org/publications/98/workpaper.html
Helfand, D. (2006, January). A formula for failure in L.A. schools. *Los Angeles Times*. Retrieved from http://www.latimes.com/news/local/la-me-dropout30jan30,0,1678653.story
Heppen, J. B., Clements, P., & Walters, K. (2015). Turning to online courses to expand access: A rigorous study of the impact of online Algebra I for eighth graders. In J. Middleton, J. Cai, & S. Hwang (Eds.), *Large-scale studies in mathematics education* (pp. 95–132). New York, NY: Springer Publishers.
Heppen, J. B., Walters, K., Clements, M., Faria, A., Tobey, C., Sorensen, N., & Culp, K. (2012). *Access to Algebra I: The effects of online mathematics for Grade 8 students* (NCEE 2012–4021). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Educational Evaluation and Regional Assistance.
Herlihy, C. (2007). *State- and district-level supports for successful transition into high school*. Washington, DC: American Institutes for Research, National High School Center.
Jones, M. P. (1996). Indicator and stratification methods for missing explanatory variables in multiple linear regression. *Journal of the American Statistical Association, 91*, 222–230.
Lee, V. E., & Ready, D. D. (2009). U.S. high school curriculum: Three phases of contemporary research and reform. *The Future of Children*, *19*, 135–156.
McCaffrey, D. F., Ridgeway, G., Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*, 403–425.
Means, B., Toyama, Y., Murphy, R., & Bakia, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, *115*, 1–47.
Murin, A., Powell, A., Roberts, V., & Patrick, S. (2015). *Using online learning for credit recovery: Getting back on track to graduation*. Vienna, VA: International Association for K–12 Online Learning.
Neild, R., & Balfanz, R. (2006). *Unfulfilled promise: The dimensions and characteristics of Philadelphia's dropout crisis, 2000–2005*. Baltimore, MD: Johns Hopkins University, Center for Social Organization of Schools.
Nomi, T. & Allensworth, E. (2009). "Double-dose" algebra as an alternative strategy to remediation: Effects on students' academic outcomes. *Journal of Research on Educational Effectiveness*, *2*, 111–148.
O'Dwyer, L. M., Carey, R., & Kleiman, G. (2007). A study of the effectiveness of the Louisiana Algebra I online course. *Journal of Research on Technology in Education*, *39*, 289–306.

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at scale. *Educational Evaluation and Policy Analysis*, *36*, 127–144.
Picciano, A., & Seaman, J. (2009). *K–12 online learning: A 2008 follow-up survey of the U.S. school district administrators*. Needham, MA: The Sloan Consortium. Retrieved from http://www.sloan-c.org/publications/survey/pdf/k-12_online_learning_2008.pdf
Picciano, A. G., & Seaman, J. (2010). *Class connections: High school reform and the role of online learning*. Babson Park, MA: Babson Survey Research Group. Retrieved from http://www3.babson.edu/ESHIP/researchpublications/upload/Class_connections.pdf.
Picciano, A. G., Seaman, J., Shea, P., & Swan, K. (2012). Examining the extent and nature of online learning in American K–12 education: The research initiatives of the Alfred P. Sloan Foundation. *The Internet and Higher Education*, *15*, 127–135. Retrieved from http://www.sciencedirect.com/science/article/pii/S1096751611000509
Queen, B., & Lewis, L. (2011). *Distance education courses for public elementary and secondary school students: 2009–10* (NCES 2012–008). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
Ridgeway, G., McCaffrey, D., Morral, A., Griffin, B.A., & Burgette, L. (2013). *Twang: Toolkit for weighting and analysis of nonequivalent groups* (R package version 1.3–20). Retrieved from http://CRAN.R-project.org/package=twang
Star, J. R., Caronongan, P., Foegen, A., Furgeson, J., Keating, B., Larson, M. R., … Zbiek, R. M. (2015). *Teaching strategies for improving algebra knowledge in middle and high school students* (NCEE 2014–4333). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
Stetser, M., & Stillwell, R. (2014). *Public high school four-year on-time graduation rates and event dropout rates: School years 2010–11 and 2011–12* (First Look, NCES 2014–391). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
Tucker, B. (2007). *Laboratories of reform: Virtual high schools and innovation in public education*. Washington, DC: Education Sector.
U.S. Department of Education. (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: Office of Planning, Evaluation, and Policy Development.
U.S. Department of Education. (2012). *Understanding the implications of online learning for educational productivity*. Washington, DC: Author.
Watson, J., & Gemin, B. (2008). *Promising practices in online learning: Using online learning for at-risk students and credit recovery*. Vienna, VA: North American Council for Online Learning.
Watson, J., Murin, A., Vashaw, L., Gemin, B., & Rapp, C. (2013). *Keeping pace with K–12 online & blended learning: An annual review of policy and practice*. Evergreen, CO: Evergreen Education Group.
Watson, J., Pape, L., Murin, A., Gemin, B., & Vashaw, L. (2014). *Keeping pace with K–12 digital learning. An annual review of policy and practice*. Evergreen, CO: Evergreen Education Group.
Watson, J., & Ryan, J. (2006). *Keeping pace with K–12 online learning: A review of state-level policy and practice*. Evergreen, CO: Evergreen Consulting Associates. Retrieved from http://evergreenassoc.com/online_education.html
Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics*, *141*, 1281–1301.