Paper 2: Targeting Summer Credit Recovery

Jared Eno and Jessica Heppen, American Institutes for Research
jeno@air.org, jheppen@air.org

Problem / Background / Context:
Algebra is considered a key gatekeeper for higher-level mathematics course-taking in high school and for college enrollment (Adelman, 2006; Gamoran & Hannigan, 2000). Yet, algebra pass rates are consistently low in many places (Higgins, 2008; Ham & Walker, 1999; Helfand, 2006), including Chicago Public Schools (CPS). This is of particular concern because academic performance in core courses during the first year of high school is the strongest predictor of eventual graduation (Allensworth & Easton, 2005).

Offering credit recovery options is one strategy to deal with high failure rates. Recently, online learning has emerged as a promising and increasingly popular strategy for credit recovery (Picciano & Seaman, 2009). Credit recovery is one of the most common applications of online courses (Greaves & Hayes, 2008). However, no rigorous evidence currently exists about the efficacy of online credit recovery courses.

To address this research gap, American Institutes for Research (AIR) and the Chicago Consortium for School Research (CCSR) used a student-level random controlled trial to test the impact of online Algebra I for credit recovery against the standard face-to-face (f2f) version of the course. As described in the introduction to Paper 1, the experiment was implemented twice with two consecutive cohorts, one in summer 2011 (cohort 1) and one in summer 2012 (cohort 2). Results suggest that overall, the online format had a significant, negative impact on credit recovery rates in cohort 2 (but not cohort 1); negative impacts on an algebra posttest in cohort 2 (but not cohort 1); no significant impacts on algebra PLAN scores; and no significant impacts on later math coursetaking.

These average treatment effects may mask heterogeneity in treatment effects for students with different levels of incoming risk or background characteristics. Predicting the types of students for which online credit recovery might be most effective is difficult because there is little rigorous evidence on the relative advantages and disadvantages of online and f2f instruction. However, a recent experimental evaluation of online versus f2f instruction in a higher education setting found that while there were no significant treatment effects overall, there were positive and significant differences favoring f2f instruction for Hispanic students, male students, and lower-achieving students (Figlio, Yin, and Rush, 2010). The authors speculated that language-minority students might find online lectures more difficult to understand, or that male and lower-achieving students tended to procrastinate viewing online lectures. If so, these tendencies may hold for online credit recovery as well.

Alternatively, the promise of online courses for credit recovery may lie in features that make them seem new to students or different from the f2f course they failed. For example, online courses can use technology to engage students in content with animations, simulations, video, and other interactive content (Archambault et al., 2010; Blackboard K–12, 2009; U.S. Department of Education, 2009). If so, online courses may be most effective for those students who are the least motivated and focused on their coursework. This hypothesis is supported by evidence suggesting that students' attendance and work effort are stronger predictors of ninth grade course failure than student background characteristics such as prior test scores, gender, race, and SES (Allensworth & Easton, 2007).

An additional possibility is that the strength of online instruction may be in the immediacy of feedback on activities and assessments, and the fact that pacing of course content can be individualized more than in a class with many different students (Archambault et al., 2010; Blackboard K–12, 2009; U.S. Department of Education, 2009). If so, then online courses may be most effective for those students whose ability lies at the edges of the distribution of their peers'. Students who would be too behind or too advanced for instruction directed to the average student in the class might benefit most from the individualized setting on online instruction.

Purpose / Objective / Research Question / Focus of Research:
Understanding patterns of treatment effects may provide clues to the relative strengths and weaknesses of online and f2f learning. A related policy question is whether district and school administrators should target online learning to certain students. This paper investigates these questions by exploring heterogeneity in the treatment effects of online algebra credit recovery.

Furthermore, this paper places the average and differential treatment effects in a broader context by examining the gap between study students (overall and by subgroup) and other CPS students who are "on track" and "off track." The policy question is whether students who fail Algebra I and attempt credit recovery ever resemble students who passed the course. That is, what does it mean to get "back on track" for different types of students?

Improvement Initiative / Intervention / Program / Practice:
The online credit recovery course used for this experiment was an algebra course offered by Aventa/K12. Aventa/K12 operates online courses in every U.S. state and their Algebra I course had been implemented widely for credit recovery—in an estimated 500 schools around the country in addition to a recent expansion in CPS. The course includes both an online teacher, who communicates individually with students through email and class message boards, and an in-class mentor, who provides in-person assistance and support. Students who have questions as they proceed through the activities have access to two sources of help.

Aventa also targets its instruction for at-risk Algebra I students by allowing them to demonstrate mastery of concepts that they had previously mastered in the course that they failed. Students can then spend more time on the topics they need to master and receive a potential boost in self-confidence as they realize they are not starting "from scratch." The Aventa Algebra I course has several types of instructional supports for at-risk students, such as lowered reading level of the content, shorter topics, an audio "read aloud" function, targeted vocabulary instruction, and formative and summative assessments. The reading support increases the likelihood that students will comprehend the material and therefore be able to progress through the course. Small content "chunks" increase students' retention and expand assessment opportunities. The assessments allow students to get quick feedback on their learning.

Setting:
The setting for the study is high schools in the Chicago Public Schools, as described in the abstract for Paper 1.

Population / Participants / Subjects:
The study sample is also described in the abstract for Paper 1; they are the students who failed second semester Algebra I in the spring of freshman year, and attempted credit recovery as part of the study in summer 2011 or 2012. The characteristics of the study students are shown in Tables 3 and 4 in Appendix B. In this paper, we examine differential treatment effects by these student characteristics. In addition, we place the sample in a broader context by examining their backgrounds and short- and longer-term outcomes to two groups of interest: (1) students who, as first-time freshmen, failed Algebra I but did not attempt credit recovery the following summer (i.e. "off track" students), and (2) students who passed Algebra I in their freshman year (i.e. "on track" students).

Research Design:
The study design is described as part of Paper 1. In brief, the core design is a randomized trial with student-level assignment to condition (online or f2f). To place the results in broader context, the randomized study sample is examined descriptively in relation to "off track" and "on track" students in the same schools. These analyses will examine the extent to which students who took online and f2f Algebra I for credit recovery "close the gap" with students who passed Algebra I in ninth grade.

Data Collection and Analysis:
Initial subgroup analyses focused on the critical outcome of algebra credit recovery. Summer grades were collected from CPS in the years after the study. Students without summer grade data are assumed to have not passed the course (i.e., they dropped the course), with the exception of a small number of students for whom data could not be requested because of missing ID numbers (one in cohort 1 and five in cohort 2).

Summer credit recovery is an important outcome for students, but it may be only partially related to math achievement. Therefore, we also examined whether online credit recovery had differential impacts on an algebra posttest administered at the end of the summer in both online and f2f classrooms. Test items were taken from NAEP. IRT analysis was used to produce scale scores for each student. Although every effort was made to obtain posttest data for all students in the study, students who dropped the course or were otherwise absent on all potential testing days of the test had missing data. Overall, 66 percent of cohort 1 students and 70 percent of cohort 2 students had posttest data.

Impact models regressed each outcome on an indicator for the online condition, controlling for a variety of student characteristics (gender, race/ethnicity, SES, special education status, an indicator for any suspensions in the previous academic year, number of absences in the previous academic year, an indicator for having passed Algebra 1A), as well as fixed effects for schools and summer sessions. Subgroup impacts were examined by including terms for both the covariate of interest as well as an interaction between the covariate and the indicator for online condition. Logistic regression was used for the binary outcome of summer credit recovery; ordinary least squares regression was used for the continuous posttest scale score.

Following the theoretical reasoning and previous evidence outlined above, we examined whether there were significantly different treatment effects by Latino background, special education status, gender, whether students' previous math performance was close to the class average (as defined by more than half the interquartile range away from the classroom mean), suspensions in the previous year, and a high number of absences in the previous year (20 or more). In additional analyses scheduled for completion in spring 2014, we will examine other moderators including prior reading achievement and failure rates across all courses).

To conduct the gap analyses between study students and students who are "on track" and "off track," we must first define the standard to which students attempting credit recovery should strive. We define this standard as the average outcome at each time point for students in study schools, who, as first-time freshmen, had passed Algebra I. While we do not have study-administered posttest scores for the broader population of "off track" and "on track" students, additional analyses will examine all of the other outcomes for students in both conditions, relative to those for these groups. For example, the gap for the control group is the average on-track student outcome score minus the average score for the control group. The gap indicates to what extent the average student in the online and f2f conditions lags behind the average student in the on-track student population. We will conduct these descriptive analyses—scheduled for completion by or before July 2014—overall and by student subgroup.

Findings / Outcomes:
Preliminary analyses did not find strong evidence that online credit recovery had differential effects on credit recovery across subgroups. Estimated differences in impacts on summer credit recovery were not statistically significant for any subgroup except for students with a higher number of absences. In cohort 1, students with less than 20 absences in the previous year had significantly lower credit recovery rates in the online course relative to f2f instruction, whereas there was no significant difference for students with more than 20 absences; this difference in effects was statistically significant. However, in cohort 2, there were significant, negative impacts of online credit recovery for students in both groups.

Preliminary analyses also did not find evidence of differential impacts on algebra posttest scores. No significant impacts on posttest scores were found for any subgroup. A substantively large difference in the treatment effect for students with more versus less than 20 absences in cohort 1 was again not found in cohort 2.

Conclusions:
Overall, findings do not suggest that online credit recovery had different effects relative to face to face instruction for several subgroups examined thus far. In particular, we do not find evidence that student with differing gender, race/ethnicity, or previous math performance had systematically different credit recovery rates or algebra posttest scores, as found by Figlio et al. (2010). Students did not appear to fare better or worse than their peers of lower or higher prior math achievement, as we might expect if online courses' relative advantage comes from increased ability to individualize instruction. Though there was some evidence that impacts varied by student motivation, for which students' previous absences (but not suspensions) may be a proxy, this effect was not consistent across the two cohorts.

These findings may suggest that, in this case, contextual factors may have outweighed student-level moderators. There were differences in average treatment effects on credit recovery and posttests across the two cohorts, with no significant impacts in cohort 1 and negative impacts in cohort 2. The differences in subgroup characteristics examined here were limited, except for the percentage of students who had been suspended in the previous year, which was 10 percentage points higher in cohort 1 than cohort 2. Other differences include technology challenges at the start of one of the two summer sessions of cohort 2, and a slightly different set of schools participating in each year of the study. Further analyses will attempt to tease apart the separate contributions of student background, implementation, and implementation context.

The lack of heterogeneous treatment effects may also be partially explained by a lack of power for subgroup analyses. In addition, analyses of impacts across multiple subgroups for multiple outcomes may increase the risk of Type I error (Sun et al., 2009). For these reasons, these analyses and findings should be considered exploratory. The additional gap analyses will add further context to these results, as well as the average effect findings presented in Paper 1.