Assessing the first two years' effectiveness of Statway®: A multilevel model with propensity score matching

Hiroyuki Yamada and Anthony S. Bryk

Carnegie Foundation for the Advancement of Teaching, USA

Author Note

Hiroyuki Yamada, Director of Analytics, Carnegie Foundation for the Advancement of Teaching.

Anthony S. Bryk, President, Carnegie Foundation for the Advancement of Teaching.

This program of work is supported by the William and Flora Hewlett Foundation, the Bill & Melinda Gates Foundation, the Lumina Foundation, the Kresge Foundation, the Carnegie Corporation of New York, the Great Lakes Higher Education Corporation, and the National Science Foundation's grant DUE-1322844.

Correspondence concerning this article should be addressed to Hiroyuki Yamada, Carnegie Foundation for the Advancement of Teaching, 51 Vista Lane, Stanford, California, 94305, USA.

E-mail: yamada@carnegiefoundation.org

Abstract

Objective: Statway is a community college pathways initiative developed by the Carnegie Foundation for the Advancement of Teaching designed to accelerate students' progress through their developmental math sequence to acquiring college math credit in statistics. Statway is a multifaceted change initiative designed to address the complex problems that impede student success. Specifically, it is a one-year pathway program through which students acquire college math credit. Instructors use research-based learning principles to improve the content and pedagogy for student learning and incorporate social-psychological interventions to sustain student engagement and persistence. In addition, language supports for students' accessibility to mathematics learning are integrated into the curriculum. Professional development resources assist faculty as they teach new content utilizing unfamiliar pedagogies. Statway is organized as a networked improvement community intending to accelerate educators' efforts to continuously improve. This study was aimed to assess the effectiveness of Statway during its first two years of implementation.

Method: We applied a multilevel model with propensity score matching to control for possible selection bias and increase the validity of causal inference.

Results: We found large effects of Statway on students attaining college math credit with persisting effects into the following year as Statway students also accumulated more college-level credits. These improved outcomes emerged for each gender and race/ethnic groups and for students with different math placement levels.

Conclusion: This study provided robust evidence that Statway increases student success in acquiring college math credit and enhances equity in student outcomes. Directions for future work are suggested.

Keywords: causal inference, propensity score matching, multilevel modeling, community college mathematics

Assessing the first two years' effectiveness of Statway: A multilevel model with propensity score matching

Community colleges aim to provide educational opportunities that prepare citizens to lead more productive personal, civic, and work lives. Approximately 60% of incoming community college students are referred to at least one developmental math course, and 80% of these students will not have earned college-level math credit even after three years (Bailey, Jeong, & Cho, 2010). Minority students tend to be assigned to greater numbers of developmental math courses and are less likely than White students to progress through their developmental math requirements to achieve college math credit (Bailey et al., 2010). Absent such credit, students cannot transfer into four-year degree programs or qualify for entry into preparation programs in a wide range of occupational-technical specialties. As a result, hundreds of thousands of students each year find themselves unable to progress toward their educational, career, and life goals. This is one of the most significant social equity problems of our time (Cullinane & Treisman, 2010).

Researchers have identified several factors that impede student success. The current developmental math education system consists of a long, multi-course sequence (Hodara, 2013). The sequence typically begins with pre-algebra followed by elementary and intermediate algebra, all as pre-requisites to taking a college-level course. Based on their scores on a placement test, students may be required to take one, two, three, or in some cases even more of these developmental courses (Cullinane & Treisman, 2010). This structure offers a serious impediment to student success. Even when students successfully complete one of these courses, many fail to continue through the sequence (Bailey et al., 2010).

In addition, classroom instruction often does not use research-based instructional activities and pedagogic practices that can foster deeper student learning (National Research Council, 2002). Traditional math curricula do relatively little to engage students' interest and demonstrate the relevance of mathematical concepts to everyday life (Carnevale & Desrochers, 2003). Many students have had negative prior math experiences leading them to believe that they are not a math person (believing that math ability is fixed or innate), which often triggers anxiety when they are faced with difficult or confusing math problems (Blackwell, Trzesniewski, & Dweck, 2007; Haynes, Perry, Stupinsky, & Daniels, 2009). This is compounded for some students (e.g., women, African Americans) who identify as part of a group that has been stereotyped as not being good at math (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). Other research indicates that students' difficulty in developmental math frequently relates to the language and literacy demands of mathematics rather than their mathematical ability (Gomez, Rodela, Lozano, & Mancevice, 2013). More specifically, students struggle to use the language of mathematics effectively, understand problem situations that require mathematical reasoning, and communicate their learning with others orally and in writing.

Statway's Theory of Improvement

To address these challenges, the Carnegie Foundation for the Advancement of Teaching developed and implemented the Statway program. The program aims to accelerate students' progress through developmental math and to acquire credit in college statistics in just one year. Six key drivers—accelerated pathway through college level math, learning principles for curriculum and instruction, productive persistence, language and literacy supports, faculty development for advancing quality teaching, and networked improvement community—

organize the program's working theory of improvement (Bryk, Gomez, Grunow, & LeMahieu, 2015).

First, as depicted in Figure 1, students starting in elementary algebra within the traditional math sequence would need at least one and a half to two years to earn college math credit. For some students it takes even longer as they may need to repeat a course or find themselves unable to register for the next course in a time slot that accommodates their family and work schedules (Bailey et al., 2010; Cullinane & Treisman, 2010). To reduce these structural barriers, Statway is designed as an intensive, integrated one-year experience involving a cohort of students working with a faculty member to achieve college level math credit. It combines college-level statistics with the concepts and skills from elementary and intermediate algebra that students need to successfully complete college statistics. The program is intended to meet the requirements for introductory college-level math in liberal arts or general education in community colleges and four-year universities and provides quantitative preparation suitable for students pursuing a non-STEM major (Cullinane & Treisman, 2010; Hodara, 2013).

[Insert Figure 1 About Here]

Second, the Statway instructional system is anchored in three research-based principles known to enhance student learning. The first principle is productive struggle through which students are more likely to retain what they learn when they expend effort solving problems that are within reach and grappling with key mathematical ideas that are comprehensible but not yet well-formed (Schmidt & Bjork, 1992). Thus, each new concept is introduced with a rich problem that engages students' thinking and encourages this struggle to understand (Hiebert & Grouws, 2007). The next one is explicit connections to concepts. Mathematics instruction sometimes focuses on procedural competence at the cost of advancing real conceptual understanding.

Research suggests making explicit connections between mathematical or statistical facts, ideas, and procedures can improve both conceptual and procedural understanding (Boaler, 1998; Hiebert & Grouws, 2007). The last one is deliberate practice in which classroom and homework tasks are designed to overcome gaps in understanding, apply what is learned, and deepen facility with key concepts (Ericsson, 2008; Ericsson, Krampe, & Tescher-Römer, 1993). Deliberate practice eschews rote repetition for carefully sequenced problems developed to guide students toward deeper understanding of core concepts (Pashler, Rohrer, Cepeda, & Carpenter, 2007).

Third, Statway incorporates an evidence-based package of student activities and faculty actions that promote productive persistence. These interventions focus on promoting students' belief that they can learn math (i.e. the growth mindset beliefs countering the fixed mindset beliefs; Dweck, 2006), reducing their anxiety (Jamieson, Mendes, Blackstock, & Schmaeder, 2010), and increasing their sense of belonging (Walton & Cohen, 2011). Specific activities focus on developing the skills needed to be effective students and the flexible mindsets necessary to utilize those skills (Dweck, Walton, & Cohen, 2011; Yeager & Walton, 2011). These aims are advanced through targeted student interventions, guidance to help faculty create more engaging classroom environments, and a lesson structure that encourages active student engagement.

Fourth, efforts are also made to reduce language and literacy barriers that can hinder student learning of mathematics, led by Gomez and her colleagues (Gomez et al., 2013, Gomez et al., 2015). Language and literacy supports for students are interwoven in instructional materials and classroom activities so that learning is accessible to all. Language and literacy tools have been developed to help students extract important vocabulary or concepts, allow them to highlight key concepts in problem situations and readings, and assist them in synthesizing information in context and strengthening reasoning skills.

Fifth, it was hypothesized that many community college faculty might find teaching Statway to be novel or even challenging because their past teaching experiences relied heavily on lectures and a teacher-centered pedagogy emphasizing the transmission of math content, facts, and procedural knowledge (Edwards, Sandoval, & McNamara, 2015; Grubb, 1999; Grubb & Grabiner, 2013). To address these possible concerns, Statway incorporates faculty professional development consisting of ongoing mentoring, online activities and resources, and in-person meetings and workshops.

Sixth and finally, the Statway program is organized as a networked improvement community (NIC) aiming to accelerate educators' efforts to continuously improve (Bryk, Gomez, & Grunow, 2011, Dolle, Gomez, Russell, & Bryk, 2013). The NIC is a scientific learning community distinguished by four essential characteristics. It is (a) focused on a well specified common aim, (b) guided by a deep understanding of the problem and the system that produces it, (c) disciplined by the rigor of improvement science, and (d) networked to accelerate the development, testing, and refinement of interventions and their effective integration into varied educational contexts. The Statway NIC joins community college faculty and administrators with improvement specialists and educational researchers from other institutions. They collaboratively engage in disciplined inquires using common conceptual frameworks, measures, and inquiry protocols to advance measureable improvements in teaching and learning in Statway (Bryk et al., 2015).

As summarized in Figure 2, the Statway initiative is organized around six key drivers: (a) structural arrangement as an accelerated year-long pathway through college level math, (b) research-based principles for curriculum and instruction, (c) strengthening the skills and mindsets that students need to succeed in an academic setting (productive persistence), (c) language and literacy supports to make mathematics learning accessible to more students, (d) support of faculty professional development for advancing quality teaching, and (e) the social arrangement of a networked improvement community to accelerate learning to continuously improve. In this regard Statway can be considered as a multifaceted change initiative, addressing complex problems simultaneously in developmental math education and providing a solution for students, faculty, and colleges.

[Insert Figure 2 About Here]

The objective of the current study was to assess the first two years' efficacy of this change initiative. We used a propensity score matching technique to statistically reduce possible selection bias (where certain kinds of students may have been more likely to enroll in Statway, leading to more positive outcomes than there otherwise would have been) and accordingly increase the validity of causal inference (Rosenbaum & Rubin, 1983). Given the hierarchical nature of our data (i.e., students nested within colleges), we employed a hierarchical linear modeling (HLM) approach (Raudenbush & Bryk, 2002) to obtain propensity scores (Hong & Raudenbush, 2005, 2006). We then compared college math performance between Statway and the non-Statway matched comparison groups. A second objective was to track the academic outcomes of students one year after their enrollment in Statway. For this purpose, we compared college-level course credit accumulation between the two matched groups in the subsequent year. This comparative analysis was intended to determine whether Statway students continue to demonstrate success even after their Statway experience. All analyses were conducted separately for Years 1 and 2 cohorts.

Method

Participants

Statway was first launched during the 2011-2012 academic year. The first cohort of students began Statway in the fall of 2011. This initial cohort of students spanned 19 community colleges across five states (i.e., California, Connecticut, Florida, Texas, and Washington; for the participating community colleges, see Appendix). In total, 50 faculty members taught 55 sections of Statway with 1133 students enrolled (Strother, Van Campen, & Grunow, 2013). The second cohort included a total of 1553 students enrolled in 77 sections of Statway taught by 67 faculty members. Of the 19 community colleges that participated in Statway, all but one offered Statway in both Years 1 and 2 (Van Campen, Sowers, & Strother, 2013). The vast majority of students placed at least two levels below a college-level math course, and almost half were also required to take at least one developmental reading course (Strother et al., 2013; Van Campen et al., 2013). Approximately 60% of the students were female, and less than one-third were raised in families where the mother held either a two or four-year college degree. Well over half of the students were minorities.

Data Collection

Institutional researchers from participating colleges provided background data on student characteristics, course enrollment and performance. Two colleges were not included in the Year 1 analyses. One college discontinued the program partway through the year, because its district mandated an alternative developmental math program. The second college implemented Statway as a further accelerated one-semester course. In so doing, they substantially changed the course content, and hence, their implementation was not as comparable as the remaining colleges'. Consequently, the Year 1 analytic sample consisted of 928 Statway students from 17 community colleges. Four colleges were not included in the Year 2 analyses. Two colleges made major changes in the curriculum, how it was offered, and the data they were willing to collect and share. The institutional research offices in two other colleges failed to provide the data necessary for conducting adequate propensity score matching. Thus, the Year 2 analyses were based on 771 Statway students from 15 community colleges. All 15 of these colleges also offered Statway in Year 1.

Study Design

Figure 3 delineates the basic study design used in this research. The first objective in this study was to identify a group of students most comparable with Statway students. Defining an appropriate comparison group in this instance was a little more complex than typically the case. As noted earlier, Statway is designed as an intensive course-of-study intended to assist developmental math students to achieve college-level credit in statistics within one academic year of continuous enrollment. In contrast, students following the traditional developmental math sequence and starting two or more levels behind college-level math cannot typically achieve college-level math credit in one year. They would need to be enrolled for one and a half to two years to meet the same benchmark (Bailey et al., 2010; Cullinane & Treisman, 2010). This led us to draw a comparison group from students who began taking their developmental math course one year before their Statway counterparts and then compare both groups' course outcomes at the end of the Statway year. Thus, comparison students had two years to achieve the same outcomes that Statway students accomplished in one year. As illustrated in the left panel of Figure 3, the comparison group for students who began Statway in Fall 2011 consisted of students who began developmental math in Fall 2010. These two groups were then compared at the end of the Spring 2012. Our goal was to be conservative in forming the comparison group by giving these students twice as much time to reach the same success benchmark as Statway students. Data on 58034 and 48383 potential comparison group students were available in the

Years 1 and 2 analyses, respectively. We also compared descriptive data and course completion outcomes for comparison group students from Years 1 and 2 and found them very similar college-by-college. Thus, there is little reason to believe that this matched strategy biases results because the comparison and treatment groups were not strictly contemporaneous.

[Insert Figure 3 About Here]

To obtain propensity scores, we formulated a two-level HLM model with a total of 44 student-level covariates including student background characteristics, course taking and performance during the two years prior to Fall 2010/2011 for the Year 1 cohort and Fall 2011/2012 for the Year 2 cohort. We selected covariates based on prior research findings and advice from institutional researchers in the participating colleges. The list includes standard student background data such as gender, race/ethnicity, and placement levels. All of these characteristics have been shown to differentiate students' progress in the developmental math sequence (Bailey et al., 2010). We also matched on detailed data about students' prior course taking and success because this has been found to be a more reliable indicator of their educational and career goals than a declared program of study (Jenkins & Cho, 2012).

Table 1 presents all of the covariates used in the propensity score matching and their descriptive statistics for Years 1 and 2. Information on some variables for some students was not recorded in institutional records. We defined these data instances as "unknown" and included this as a separate matching category. For instance, there are a substantial number of unknown records for student placement levels because the information on student placement levels was missing and/or students did not take a placement test. Missing GPAs correspond to students who had not yet taken any college-level courses or received grades that do not have an effect on their GPAs (i.e., W [Withdrawals] and I [Incompletes]). To factor these cases with missing GPA scores into the propensity model, we formulated a dummy variable and coded missing GPAs as 1, otherwise 0.

[Insert Table 1 About Here]

Overall, Statway and traditional developmental math students look quite similar across the covariate set. Year 1 Statway students were somewhat less likely to be in their first year of study. The racial ethnic composition was also a bit different with slightly more Black students and fewer Hispanic students enrolled in Statway. Statway students were also more likely to have placed exactly two levels below college math. The latter is not surprising in that this was the target group of students for whom Statway was specifically recommended. Year 2 data appears similar with the exception of somewhat fewer full time students enrolled in the traditional developmental math sequence.

We conducted propensity score matching separately for each college by applying a nearest neighbor matching algorithm (Rosenbaum & Rubin, 1985). This algorithm was appropriate for our study because we wanted to retain all Statway students and had a large pool of non-Statway students available for creating matches. We attempted to find up to five matches per Statway student (5:1 ratio matching) to maximize the best matches from the non-Statway student group while still maintaining precision (Ming & Rosenbaum, 2000). We also specified a caliper distance of up to 0.2 to reduce the risk of bad nearest neighbor matches based on recommendations in the literature (Austin, 2011; Rosenbaum & Rubin, 1985).

Next, as illustrated in the middle panel of Figure 3, we estimated the effectiveness of Statway by comparing success rates of Statway students with their matched comparisons using a three-level HLM model with a binary outcome. Success was defined as a grade of C or higher for Statway students at the end of the year. For the matched comparisons, success was defined as a C or higher on any college-level math course as defined in local institutional records.[1] Statway students (level 1) were nested within their faculty member classrooms (level 2), and faculty in turn were nested within colleges (level 3). Since matched comparisons were created for each Statway student, their respective comparison students were also assigned the corresponding Statway faculty ID. In essence each faculty member's classrooms now formed as a mini experiment where the mean outcomes for their students could be compared to those for students just like them who had pursued the more conventional course of study. Key for our analytic purposes, this strategy permitted estimation of the variability in effects among faculty within colleges. Finally, as depicted in the right panel of Figure 3, we also examined college-level course performance between Statway and their matched comparisons in the subsequent calendar year including a summer term where applicable. These follow-up data were available only for the Year 1 cohort. We defined student performance in the follow-up year in terms of the college course credits accumulated with a grade of C or higher. We used HLM 7 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) for all of the HLM analyses.

Results

Propensity Score Matching

To obtain propensity scores, we formulated a two-level Bernoulli model and estimated its model parameters using maximum likelihood via adaptive Gaussian quadrature. ϕ_{ij} is the probability of student *i* enrolling in Statway in college *j*. Accordingly, η_{ij} is the log-odds of this incident and formally expressed as:

Level-1 Model (Student)

[1] A grade of C- or higher was used for six colleges that employ a +/– grading system to define college math success. The same strategy was applied to the following analysis to define college-level units earned in the subsequent year.

$$\text{Prob}(SW_{ij}=1|\beta_j) = \phi_{ij},$$

$$\log[\phi_{ij}/(1 - \phi_{ij})] = \eta_{ij},$$

$$\eta_{ij} = \beta_{0j} + \beta_{1j}*(COV1_{ij}) + \ldots + \beta_{44j}*(COV44_{ij}),$$

Level-2 Model (College)

$$\beta_{0j} = \gamma_{00} + u_{0j},$$

$$\beta_{1j} = \gamma_{10},$$

$$\ldots,$$

$$\beta_{43j} = \gamma_{430},$$

$\beta_{44j} = \gamma_{440} + u_{44j}$, where *SW* is a dummy variable indicating whether a given student was enrolled in Statway (coded as 1) or not (coded as 0), *COV1…COV44* are the set of propensity score covariates, and *i* and *j* denote student and college, respectively. We estimated one random slope, β_{44j}, for a dummy variable indicating placement two levels below college math. Preliminary analyses identified significant heterogeneity among colleges in this relationship**.** Consequently, the propensity score matching in each college was based on their local site specific relationship for this one variable. We matched a total of 4549 comparison students to 928 Statway students for Year 1 and a total of 3583 comparison students matched to 771 Statway students for Year 2.[2] Table 1 compares the descriptive statistics on each covariate before and after matching to the Statway group. Table 2 documents the balance in propensity score college-by-college for Years 1 and 2. For both cohorts, there were no significant differences in mean propensity score between

---

[2] Unlike Year 1, we found that a model with a fixed effect for math placement two levels below college fit the data better than the random effect model deployed in Year 1. Hence, in Year 2 we used propensity scores from the fixed slope model for purposes of matching.

the Statway and matched students in any of the colleges (see *t*-values). Taken together, this provides strong evidence that comparability of the groups was achieved on the measured covariates.

[Insert Table 2 About Here]

Estimating Statway Effects

To estimate differences in success rate, we formulated a three-level Bernoulli model[3] and estimated its model parameters using maximum likelihood via adaptive Gaussian quadrature. ϕ_{ijk} represents the probability that student *i* associated with faculty member *j*'s class in college *k* successfully achieved college math credit. Correspondingly, η_{ijk} is the corresponding log-odds of this outcome and formally expressed as:

Level-1 Model (Student)

$$\text{Prob}(CMA_{ijk}=1|\pi_{jk}) = \phi_{ijk},$$

$$\log[\phi_{ijk}/(1 - \phi_{ijk})] = \eta_{ijk},$$

$$\eta_{ijk} = \pi_{0jk} + \pi_{1jk}*(PS_{ijk}) + \pi_{2jk}*(SW_{ijk}),$$

Level-2 Model (Faculty)

$$\pi_{0jk} = \beta_{00k} + r_{0jk},$$

$$\pi_{1jk} = \beta_{10k},$$

$$\pi_{2jk} = \beta_{20k},$$

[3] We also ran a four-level model that broke out matched clusters as a separate level. In principle, we can think of the data as consisting of matched clusters for each Statway student associated with each faculty member and all of this in turn nested within colleges. The results from these four-level models closely mirrored the three-level analyses. For simplicity of presentation, we focus here on the three-level results.

Level-3 Model (College)

$$\beta_{00k} = \gamma_{000} + u_{00k},$$

$$\beta_{10k} = \gamma_{100},$$

$$\beta_{20k} = \gamma_{200},$$

where *CMA* represents college math achievement (1 for successfully completed and 0 for not successfully completed), and *SW* is a dummy variable indicating whether the student was enrolled in Statway (coded as 1) or one of the matched comparisons (coded as 0). As a further safeguard, we included individual students' propensity scores, *PS*, as an additional adjustment variable.

The results presented in Table 3 indicate that on average, Statway students demonstrated significantly higher odds of success, 5.31 and 7.40 (95% CIs [4.54, 6.21] and [6.20, 8.85]), in achieving college-level mathematics credit than the comparison students for Years 1 and 2, respectively. These translated into the estimated probabilities of success of 54.43% and 55.26% for the Statway groups and 18.36% and 14.30% for the comparison groups for Years 1 and 2.[4] Additionally, we found variation among colleges in student success (0.239 and 0.342 for the Years 1 and 2 variances). Figure 4 shows that for both cohorts, students in all but one college demonstrated greater success in Statway.

---

[4] We also conducted sensitivity analyses (Hong & Raudenbush, 2005, 2006) on Statway effects on college math achievement for Years 1 and 2 and college credit accumulation for Year 1. Results indicated that with adjustments for the largest potential hidden bias, none of the 95% confidence intervals for the new Statway effect estimates contained 0 or any negative values, thereby supporting the strong ignorability assumption. Thus, it is very unlikely that our general conclusion regarding the positive effects of Statway on the student outcomes has been influenced by the omission of unmeasured confounding factors.

[Insert Table 3 About Here]

[Insert Figure 4 About Here]

To estimate differences in college credits earned with a grade of C or higher in the subsequent year, we formulated a three-level Poisson model and estimated its model parameters using penalized quasi-likelihood estimation.[5] λ_{ijk} represents the event rate that student *i* within faculty member *j*'s class in college *k* successfully earns college credits in the following year. Thus, η_{ijk} is the corresponding log of this event and formally expressed as:

Level-1 Model (Student)

$$E(CCE_{ijk}|\pi_{jk}) = \lambda_{ijk},$$

$$\log[\lambda_{ijk}] = \eta_{ijk},$$

$$\eta_{ijk} = \pi_{0jk} + \pi_{1jk}*(PS_{ijk}) + \pi_{2jk}*(SW_{ijk}),$$

Level-2 Model (Faculty)

$$\pi_{0jk} = \beta_{00k} + r_{0jk},$$

$$\pi_{1jk} = \beta_{10k},$$

$$\pi_{2jk} = \beta_{20k},$$

Level-3 Model (College)

$$\beta_{00k} = \gamma_{000} + u_{00k},$$

$$\beta_{10k} = \gamma_{100},$$

$$\beta_{20k} = \gamma_{200},$$

[5] HLM 7 applies penalized quasi-likelihood estimation to a three or higher level Poisson model.

where CCE represents accumulated college-level units earned with a grade of C or higher in the subsequent year.[6]

The results presented in Table 4 indicate that on average, Statway students earned significantly more college credits than their matched comparison (with a higher event rate ratio, 1.37, 95% CI [1.11, 1.69]). The estimated accumulated credits for the Stataway and matched comparison groups were 5.57 vs. 4.08, respectively. We again found variation among colleges (Variance = 0.320). Figure 5 depicts that in the majority of colleges, Statway students earned more college credits in the subsequent year than their matched comparisons.

[Insert Table 4 About Here]

[Insert Figure 5 About Here]

Subgroup Analyses

To examine possible differential effects of Statway (a) by gender and race/ethnicity subgroups and (b) by math placement levels, we formulated a three-level HLM similar to those described above. In these subgroup analyses, however, we applied effect coding to the grouping variables in order to directly represent both main and interaction effects on the outcome. The reference categories were female, White, and a math placement three or more levels below college. Each of these was coded as -1. We excluded cases with the unknown gender status or the college math placement level.

Table 5 presents the model based results transformed back into their natural metrics of proportion of students successfully acquiring college math credit and accumulated college credits earned in the follow-up year. This metric transformation was made for the ease of interpretation.

[6] There were three quarter colleges, and accordingly, their college-level units were converted into semester units by dividing the units by 1.5.

The left and middle panels of Table 5 present the estimated proportions successful by gender and race/ethnicity and by math placement level. In general, large positive Statway effects appear consistently for all subgroups of students. The only exception was a somewhat smaller effect among Year 1 students who were placed one level below college level. Likewise, the right panel of Table 5 presents the estimated college credit accumulation by gender and race/ethnicity and math placement level. Positive effects of Statway were observed for each major race/ethnicity group: Black, Hispanic, and White. Also, regardless of math placement levels, overall, Statway students performed better than the comparison students. The effect appears the largest among students who were placed two levels below college level. The latter is not surprising as this was the subgroup of students whose outcomes Statway was specifically designed to improve.

[Insert Table 5 About Here]

Discussion

The current study sought to undertake a rigorous causal analysis of Statway's effectiveness for community college students. To assess this, we used a propensity score matching technique (Rosenbaum & Rubin, 1983) within a hierarchical linear modeling framework (Raudenbush & Bryk, 2002). Given the modest number of students participating in Statway in each college as compared to the college's population of developmental math students, we were able to secure a very high degree of propensity score matching across 44 different indicators. We also undertook a sensitivity analysis to examine the robustness of the estimated effects for possible unmeasured covariates. Given the large size of the estimated effects complemented with the results from the sensitivity analysis, we conclude that there is strong evidence that Statway substantially improves student success rates in acquiring college level math credit. Our results also suggest that significant Statway effects persist into the following school year. Statway students accumulated more college credits in the follow-up year than their matched comparisons. We speculate that the latter may be a consequence of Statway's emphasis on strengthening student growth mindset as a mathematical learner and doer and enhancing students' sense of belonging in a mathematical environment. These are the major foci for productive persistence, one of the six key drivers of Statway (see Figure 2).

These results are replicated across two different implementation cohorts. We also found significant improvements for all gender and race/ethnicity groups as well as for students with different math placement levels. The results suggest that Statway effectively advances more equitable outcomes for disadvantaged students than has been documented with traditional developmental math curricula and instruction (Bailey et al., 2010; Carnevale & Desrochers, 2003; National Research Council, 2002).

Our findings are consistent with the hypothesis formulated by Hodara (2013) that "structural changes to the developmental sequence may have limited effects, but the combination of structural, curricular, and pedagogical changes to a developmental math sequence as well as the provision of non-academic supports can impact the college success of students in long-lasting, meaningful ways" (p. 29). First, Statway accelerated the developmental math course-taking process and reduced the time required to earn college credit in one year by integrating developmental math skills and college-level statistics. This structural feature contributes to minimizing the confusion and ineffectiveness of the developmental system (Bailey et al., 2010). Second, the Statway instructional system is grounded in research-based learning principles: productive struggle, explicit connections to concepts, and deliberate practice (Boaler, 1998; Ericsson, 2008; Ericsson et al., 1993; Hiebert & Grouws, 2007; Pashler et al., 2007; Schmidt & Bjork, 1992). This instructional system helps to improve the content and pedagogy of developmental math and foster deeper student learning. Third, this system also integrates two types of research-based supports for students. One pertains to social-psychological factors of student learning (productive persistence) to sustain students' engagement and persistence (Dweck et al., 2011; Jamieson et al., 2010; Walton & Cohen, 2011; Yeager & Walton, 2011), and the other addresses language and literacy barriers to make leaning accessible to students (Gomez et al., 2013, Gomez et al., 2015). Fourth, Statway provides faculty with a rich set of opportunities for professional development to advance the quality of their teaching of Statway and quite possibly beyond as well (Edwards et al., 2015; Grubb, 1999; Grubb & Grabiner, 2013). Finally, Statway is organized as a NIC to accelerate collaboration and learning among college faculty and administrators, improvement specialists, and educational researchers (Bryk et al., 2011, Dolle et al., 2013). Although highly speculative, there is also the possibility of significant derivative effects associated with faculty participation in the Statway NIC. Although anecdotal in form, individual Statway faculty members have reported that their experiences here are changing the way they teach more generally as well as how they think about student learning. This is a tantalizing hypothesis that merits future empirical scrutiny.

It is important to acknowledge some limitations of the present study. First, although we identified large effects of Statway across two different implementation cohorts (complemented with the results from the sensitivity analyses), the matched comparisons in this study were not fully contemporaneous with Statway students. In order to allow comparison students to complete their program of study in two years and then permit us to follow up both groups for another year, comparison students began essentially a year earlier than Statway students. This raises the possibility of cohort effects that might confound our results. However, we found no evidence of a significant overall improvement in outcomes over that two-year period in the data provided by college institutional researchers on all of their developmental math students. If cohort effects were operating, we would expect to see this improvement. Regardless, it is important to consider this alternative hypothesis in future analyses as more data become available over time. To do this, we would need to match comparison students from the same enrollment cohorts and then follow them forward for staggered periods (i.e. two years for Statway students and three years for the matched comparisons).

Second, no information on programs of study or majors declared was available for use in propensity score matching in this study. Although the reliability of such data has been questioned (for example, see Jenkins & Cho, 2012), information on students' declared programs of study or majors might further improve the matching of non-Statway students with Statway students especially in cases where prior course-taking data are limited (e.g., the first semester students).

Third, we also want to examine performance for the Year 2 cohort in the follow-up year to see if the results on increased college course credit accumulation also replicate. Even longer term as more extended longitudinal data become available, we want to examine more distal outcomes such as transfer rates and academic success of Statway students in four-year institutions. These analyses would further illuminate the dimensions and possible limitations of Statway's effectiveness.

Finally, we note that we found no evidence in our analyses of significant variability in student outcomes among faculty within colleges. However, up to this point in time, the number of faculty members teaching Statway per college has been small and therefore the power to detect such variation in performance has been limited. As the Statway initiative now starts to scale to many more sections within colleges, future studies should explicitly focus here. It is important to investigate whether the effects reported in this study generalize as a larger and presumably more diverse sample of faculty subsequently take up this work.

In conclusion, these overall results suggest that Statway is a very promising alternative to the traditional developmental algebra pathway. It is also an effective solution that advances equity. Statway takes a holistic, systemic approach as a multifaceted change initiative to tackle complex problems in developmental math education. Addressing all those issues simultaneously makes Statway distinct from traditional developmental math programs and seems the key to student success. Through a NIC structure, college faculty and administrators, improvement specialists, and educational researchers collaborate with each other and accelerate learning to improve Statway in order to change developmental math education from crisis to hope for students to sustain their academic and career aspirations.

References

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics, 10,* 150-161.Bailey, T., Jeong, D. W., & Cho, S.W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review, 29*, 255–270.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*, 246–263.

Boaler, J. (1998). Open and closed mathematics: Student experiences and understandings. *Journal for Research in Mathematics Education, 29,* 41-62. Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. T. Hallinan (Ed.), *Frontiers in sociology of education* (pp. 127-162). New York, NY: Springer.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). Learning to improve: How America's schools can get better at getting better. Cambridge, MA: Harvard Education Press.

Carnevale, A. P. & Desrochers, D. M. (2003). *Standards for what?: The economic roots of K-16 reform*. Princeton, NJ: Educational Testing Service.

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, *324,* 400-403.

Cullinane, J., & Treisman, P. U. (2010). *Improving developmental mathematics education in community colleges: A prospectus and early progress report on the Statway initiative.* An NCPR Working Paper. New York, NY: National Center for Postsecondary Research.

Dolle, J. R., Gomez, L. M., Russell, J. L., & Bryk, A. S. (2013). "More than a network: building communities for educational improvement." In B. J. Fishman, W. R. Penuel, A. R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research: Theories, methods, and exemplars.* National Society for the Study of Education Yearbook. New York, NY: Teachers College Record.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* New York, NY: Random House.

Dweck, C. S., Walton, G. M., & Cohen, G. L. (2011). *Academic tenacity: Mindsets and skills that promote long-term learning*. Seattle, WA: White paper prepared for the Gates Foundation.

Edwards, A. R., Sandoval, C., & McNamara, H. (2015). Designing for improvement in professional development for community college developmental mathematics faculty. *Journal of Teacher Education, 66,*466-481. doi: 10.1177/0022487115602313

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine, 15*, 988-994.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

Gomez, K., Rodela, K., Lozano, M., & Mancevice, N. (2013). Designing embedded language and literacy supports for developmental mathematics teaching and learning. *MathAMATYC Educator, 5,* 43-56.

Gomez, K., Gomez, L. M., Rodela, K. C., Horton, E. S., Cunningham, J., & Ambrocio, R. (2015). Embedding language support in developmental mathematics lessons: Exploring the value of design as professional development for community college mathematics instructors. *Journal of Teacher Education, 66,*450-465. doi: 10.1177/ 0022487115602127

Grubb, W. N. (1999). *Honored but invisible: An inside look at teaching in community colleges*. New York, NY: Routledge.

Grubb, W. N., & Gabriner, R. (2013). *Basic skills education in community colleges: Inside and outside of classrooms.* New York, NY: Routledge.

Haynes, T. L., Perry, R. P., Stupnisky, R. H., & Daniels, L. M. (2009). A review of attributional retraining treatments: Fostering engagement and persistence in vulnerable college students. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 227–272). New York, NY: Springer.

Hiebert, J., & Grouws, D. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age.

Hodara, M. (2013). *Improving students' college math readiness: A review of the evidence on postsecondary interventions and reforms.* A CAPSEE Working Paper. Center for Analysis of Postsecondary Education and Employment.

Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis, 27,* 205-224.

Hong, G., & Raudenbush, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association, 101,* 901-910.

Jamieson, J. P., Mendes, W. B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology*, *46*, 208-212.

Jenkins, D., & Cho, S. W. (2012). *Get with the program: Accelerating community college students' entry into and completion of programs of study* (CCRC Working Paper No. 32). New York, NY: Columbia University, Teachers College, Community College Research Center.

Ming, K., & Rosenbaum, P. (2000). Substantial gains in bias reduction from matching with a variable number of controls. *Biometrics, 56*, 118-124.

National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 14*, 187-193.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70,* 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39,* 33-38.

Schmidt, R.A., & Bjork, R.A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207-217.

Strother, S., Van Campen, J., & Grunow. (2013) *Community College Pathways: 2011-2012 Descriptive Report.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Van Campen, J., Sowers, N., & Strother, S. (2013) *Community College Pathways: 2012-2013 Descriptive Report.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes among minority students. *Science, 331,* 1447–1451.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267-301.

Appendix

List of Participating Community Colleges

- American River College
- Austin Community College
- Capital Community College
- Gateway Community College
- El Paso Community College
- Foothill College
- Housatonic Community College
- Houston Community College
- Los Angeles Pierce College
- Miami Dade College
- Mt. San Antonio College
- Naugatuck Valley Community College
- Northwest Vista College
- Richland College
- San Diego City College
- Seattle Central Community College
- Tacoma Community College
- Tallahassee Community College
- Valencia College

Table 1.
Descriptive Statistics of Covariates in the Two-Level Propensity Models

	Year 1			Year 2		
	Non-Statway		Statway	Non-Statway		Statway
Covariate	Before matching	After matching		Before matching	After matching	
	%	%	%	%	%	%
Cohort						
First year*	57	43	40	51	33	33
Second year or older	43	57	60	49	67	67
Gender						
Female*	57	57	58	58	61	60
Male	43	43	42	42	39	40
Unknown	0	0	0	0	0	0
Race/Ethnicity						
Black	21	24	25	23	25	25
Hispanic	37	28	29	38	30	29
White*	29	32	29	25	31	32
Other	8	10	11	8	9	9
Unknown	5	6	6	6	5	5
Type of first-time student						
First-time college*	82	75	74			
First-time transfer	18	25	26			
Dual enrollment in a previous term						
Yes	4	4	4	4	2	2
No*	87	78	78	85	79	76
Unknown	9	18	18	11	19	22
Math placement level						
College Level	6	1	2	4	2	2

1 level below college level	18		17		15		15		8		6	
2 levels below college level	35		43		43		31		31		29	
3+ levels below college level*	26		21		21		31		21		24	
Unknown	15		18		19		19		38		39	
English placement level												
College level*	29		33		32		27		19		19	
Developmental level	47		41		41		45		30		31	
Unknown	24		26		27		28		51		50	
Reading placement level												
College level*	31		33		32		29		24		25	
Developmental level	39		33		33		37		25		26	
Unknown	30		34		35		34		51		49	
Part time vs. Full time												
Full time*	54		56		54		48		52		54	
Part time	46		44		46		52		48		46	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (in years)	23.2	7.3	25.6	9.8	26.1	10.3	23.5	7.3	27.0	9.6	27.3	9.8
Prior course enrollment and performance												
College math units attempted	0.1	0.6	0.1	0.6	0.1	0.6	0.1	0.5	0.1	0.7	0.1	0.9
College math units completed	0.0	0.3	0.0	0.3	0.0	0.3	0.0	0.3	0.0	0.3	0.0	0.3
College math courses attempted (Year 1)	0.0	0.2	0.0	0.2	0.0	0.2						
							%		%		%	
College math courses attempted (Year 2)												
0*							99		97		96	
1							1		2		3	
2 or more							0		1		1	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
College math courses completed	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1
Developmental math units attempted	2.5	4.1	3.5	5.4	3.6	5.5	2.7	4.0	3.2	4.4	3.2	4.5

Developmental math units completed	1.4	2.8	1.9	3.5	2.0	3.7	1.5	2.6	1.5	2.6	1.5	2.4
Developmental math courses attempted	0.7	1.2	1.0	1.5	1.0	1.5	0.8	1.2	1.0	1.3	1.0	1.3
Developmental math courses completed	0.4	0.8	0.6	0.9	0.6	1.0	0.5	0.8	0.5	0.8	0.4	0.7
College non-math units attempted	7.6	13.3	12.9	19.2	13.9	20.2	8.6	13.4	13.5	16.5	14.4	17.5
College non-math units completed	5.6	10.8	9.7	15.8	10.7	16.6	6.3	10.7	10.5	13.7	11.4	14.6
College non-math courses attempted	2.5	4.3	4.2	6.0	4.5	6.1	2.9	4.4	4.5	5.4	4.7	5.7
College non-math courses completed	1.9	3.5	3.2	4.9	3.5	5.1	2.1	3.5	3.4	4.4	3.7	4.6
Developmental non-math units attempted	2.0	5.1	2.4	6.3	2.5	6.8	2.1	4.1	1.7	3.8	1.7	3.8
Developmental non-math units completed	1.6	4.3	2.0	5.5	2.0	5.8	1.8	3.7	1.5	3.5	1.5	3.5
Developmental non-math courses attempted	0.6	1.5	0.7	1.8	0.8	2.0	0.7	1.3	0.5	1.2	0.5	1.2
Developmental non-math courses completed	0.5	1.3	0.6	1.6	0.6	1.7	0.6	1.2	0.5	1.1	0.5	1.1
College STEM courses attempted	0.3	1.0	0.5	1.2	0.4	1.2	0.4	0.8	0.4	0.9	0.4	0.9
College STEM courses completed	0.2	0.8	0.3	0.9	0.3	0.9	0.2	0.6	0.3	0.7	0.3	0.7
College non-STEM courses attempted	2.2	3.8	3.8	5.5	4.1	5.7	2.5	3.8	3.9	4.8	4.2	5.0
College non-STEM courses completed	1.7	3.1	2.9	4.5	3.2	4.9	1.8	3.1	3.1	4.0	3.3	4.2
GPA of college STEM courses	2.0	1.4	2.0	1.4	2.1	1.3	2.1	1.4	2.1	1.4	2.0	1.4
GPA of college non-STEM courses	2.4	1.1	2.5	1.1	2.5	1.0	2.3	1.1	2.5	1.0	2.6	1.0
Missing on college STEM GPA	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4	0.8	0.4
Missing on college non-STEM GPA	0.6	0.5	0.5	0.5	0.4	0.5	0.5	0.5	0.4	0.5	0.4	0.5

Note. Terms with "*" were used as reference categories (coded as 0, otherwise 1) when formulating dummy variables. First year under Cohort was defined as Summer/Fall enrollment in a given college for the first time in 2010/2011 for Year 1 non-Statway/Statway and 2011/2012 for Year 2. Type of first-time student was available only for Year 1. Part time vs. Full time status was based on Fall 2010/2011 enrollment for Year 1 non-Statway/Statway and Fall 2011/2012 enrollment for Year 2, with 12 or more units considered as full time. Age was computed by subtracting a birth year from 2010/2011 for Year 1 non-Statway/Statway and 2011/2012 for Year 2; in the current analyses, we centered Age around age 18. "Completed" was defined as course credit attained with a grade of C or higher (C- or higher if a college employs a +/– grading system) or Pass for developmental courses.

Table 2.

Balance in Logit of the Propensity Score for non-Statway and Statway Students

	Year 1										Year 2									
	Non-Statway						Statway				Non-Statway						Statway			
	Sample before matching			Sample after matching							Sample before matching			Sample after matching						
College	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*t*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*t*
1	3463	-3.98	0.81	477	-3.18	0.92	97	-3.13	0.96	-0.43	3517	-3.83	0.97	626	-2.94	0.93	126	-2.92	0.94	-0.20
2	637	-3.25	0.72	171	-2.83	0.75	36	-2.72	0.90	-0.71	580	-3.78	1.02	115	-3.30	1.02	23	-3.31	1.02	0.04
3	3857	-4.15	0.61	385	-3.65	0.62	77	-3.65	0.62	0.00	3969	-4.33	0.96	418	-3.17	1.10	91	-3.01	1.20	-1.19
4	2270	-3.74	0.51	320	-3.53	0.54	65	-3.50	0.61	-0.33										
5	2610	-3.93	0.48	286	-3.46	0.67	60	-3.40	0.74	-0.54	2905	-4.34	0.86	280	-3.59	0.68	56	-3.59	0.68	0.00
6	1214	-3.16	0.79	341	-2.80	0.58	70	-2.76	0.64	-0.52	987	-3.36	0.79	241	-2.64	0.64	50	-2.63	0.62	-0.06
7	2408	-4.12	0.82	254	-3.83	0.62	51	-3.82	0.64	-0.07	2129	-4.46	0.72	170	-3.50	0.31	34	-3.50	0.32	0.01
8	1451	-3.71	0.79	228	-3.09	0.64	48	-3.00	0.72	-0.74	1618	-3.81	0.71	238	-2.98	0.82	51	-2.86	0.94	-0.87
9	2243	-3.70	0.65	341	-3.34	0.64	70	-3.28	0.73	-0.58	1976	-3.82	0.73	310	-2.80	0.60	62	-2.80	0.59	0.06
10	3975	-4.61	0.54	240	-4.16	0.50	48	-4.17	0.49	0.08	3902	-5.07	0.62	145	-4.58	0.69	30	-4.51	0.80	-0.48
11	8623	-5.67	1.07	255	-4.85	0.69	51	-4.85	0.69	-0.01										
12	6779	-4.76	0.58	340	-4.22	0.57	69	-4.16	0.62	-0.65	6999	-5.02	0.90	365	-4.38	1.09	73	-4.38	1.10	-0.02
13	4763	-5.40	0.50	110	-5.10	0.49	22	-5.10	0.50	-0.01	4613	-5.73	0.69	90	-4.93	0.69	18	-4.93	0.71	0.01
14	8955	-5.21	0.54	280	-4.69	0.71	56	-4.69	0.72	0.01	9994	-5.95	0.78	175	-5.48	1.05	35	-5.48	1.06	-0.01
15	2970	-4.76	0.78	171	-4.00	0.72	35	-3.96	0.76	-0.29	3317	-5.25	0.87	130	-4.33	0.78	26	-4.33	0.79	-0.01
16	714	-3.26	0.59	171	-2.91	0.55	36	-2.83	0.68	-0.69	789	-2.99	0.79	104	-2.09	0.80	60	-1.89	0.91	-1.43
17	1102	-3.84	0.88	179	-3.18	1.01	37	-3.12	1.13	-0.27	1088	-4.03	1.13	176	-3.62	1.20	36	-3.56	1.28	-0.28

Table 3.
Model-Based Estimation of Statway Effect on College Math Achievement

Year 1	Fixed effect	Coef.	*SE*	*t*	*p*-value	Odds ratio
	Intercept	-1.67	0.13	-13.06	<0.001	0.19
	Propensity score	0.17	0.05	3.37	<0.001	1.19
	Statway effect	1.67	0.08	20.84	<0.001	5.31
	Random effect	Variance	*df*	χ^2	*p*-value	
	Level 3 (college)	0.239	16	153.32	< 0.001	
	Level 2 (faculty)	0.014	24	31.43	0.142	
Year 2	Fixed effect	Coef.	*SE*	*t*	*p*-value	Odds ratio
	Intercept	-1.81	0.16	-11.27	<0.001	0.16
	Propensity score	0.02	0.05	0.35	0.726	1.02
	Statway effect	2.00	0.09	22.02	<0.001	7.40
	Random effect	Variance	*df*	χ^2	*p*-value	
	Level 3 (college)	0.342	14	179.22	<0.001	
	Level 2 (faculty)	0.000	21	17.53	>0.500	

Note. The *df*'s, χ^2 statistics, and *p*-values are derived from penalized quasi-likelihood estimation but reported here to indicate approximate significance levels for the random effects.

Table 4.
Model-Based Estimation of Statway Effect on Accumulated College Credits Earned in the Subsequent Year - Year 1

Fixed effect	Coef.	*SE*	*T*	*p*-value	Event rate ratio
Intercept	1.51	0.14	10.51	<0.001	4.52
Propensity score	-0.10	0.03	-3.09	0.002	0.90
Statway effect	0.31	0.11	2.89	0.004	1.37
Random effect	Variance	*df*	$\chi 2$	*p*-value	
Level 3 (college)	0.320	16	373.82	<0.001	
Level 2 (faculty)	0.032	24	367.71	<0.001	

Table 5.
Model-Based Success Rates and College-Level Credits Accumulated in the Subsequent Year by Gender and Race/Ethnicity and by Math Placement Level

		Year 1				Year 2				Year 1			
		Non-Statway		Statway		Non-Statway		Statway		Non-Statway		Statway	
		%	*n*	%	*n*	%	*n*	%	*n*	Credits	*n*	Credits	*n*
Female	Black	13	685	48	146	8	572	53	121	3.20	685	5.04	146
	Hispanic	22	743	58	154	13	655	55	134	4.55	743	5.91	154
	White	20	764	67	150	16	682	67	144	4.40	764	5.93	150
	Other	25	246	55	59	22	175	58	41	6.00	246	6.11	59
	Unknown	26	155	61	28	17	107	54	20	4.81	155	3.88	28
Male	Black	11	412	43	83	10	308	41	68	2.64	412	5.27	83
	Hispanic	15	537	49	113	18	405	50	90	4.23	537	5.17	113
	White	22	654	57	122	20	463	60	105	4.00	654	6.54	122
	Other	26	230	51	42	20	135	55	28	5.33	230	4.15	42
	Unknown	21	110	64	28	13	73	61	17	4.31	110	7.71	28
1 level below		29	778	48	136	27	279	62	50	5.68	778	6.90	136
2 levels below		17	1960	59	397	16	1098	49	222	4.08	1960	6.49	397
3+ levels below		14	931	54	205	10	765	60	183	3.50	931	4.91	205
Unknown		19	814	56	172	14	1370	59	299	3.52	814	4.30	172

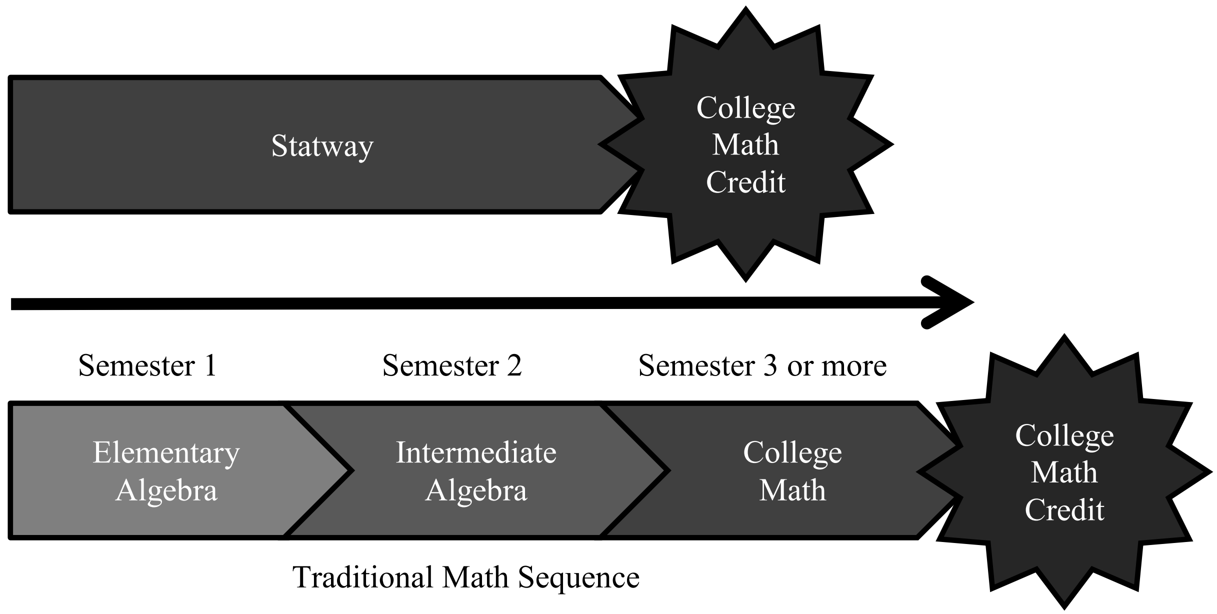

Figure 1. Statway vs. Traditional math sequence

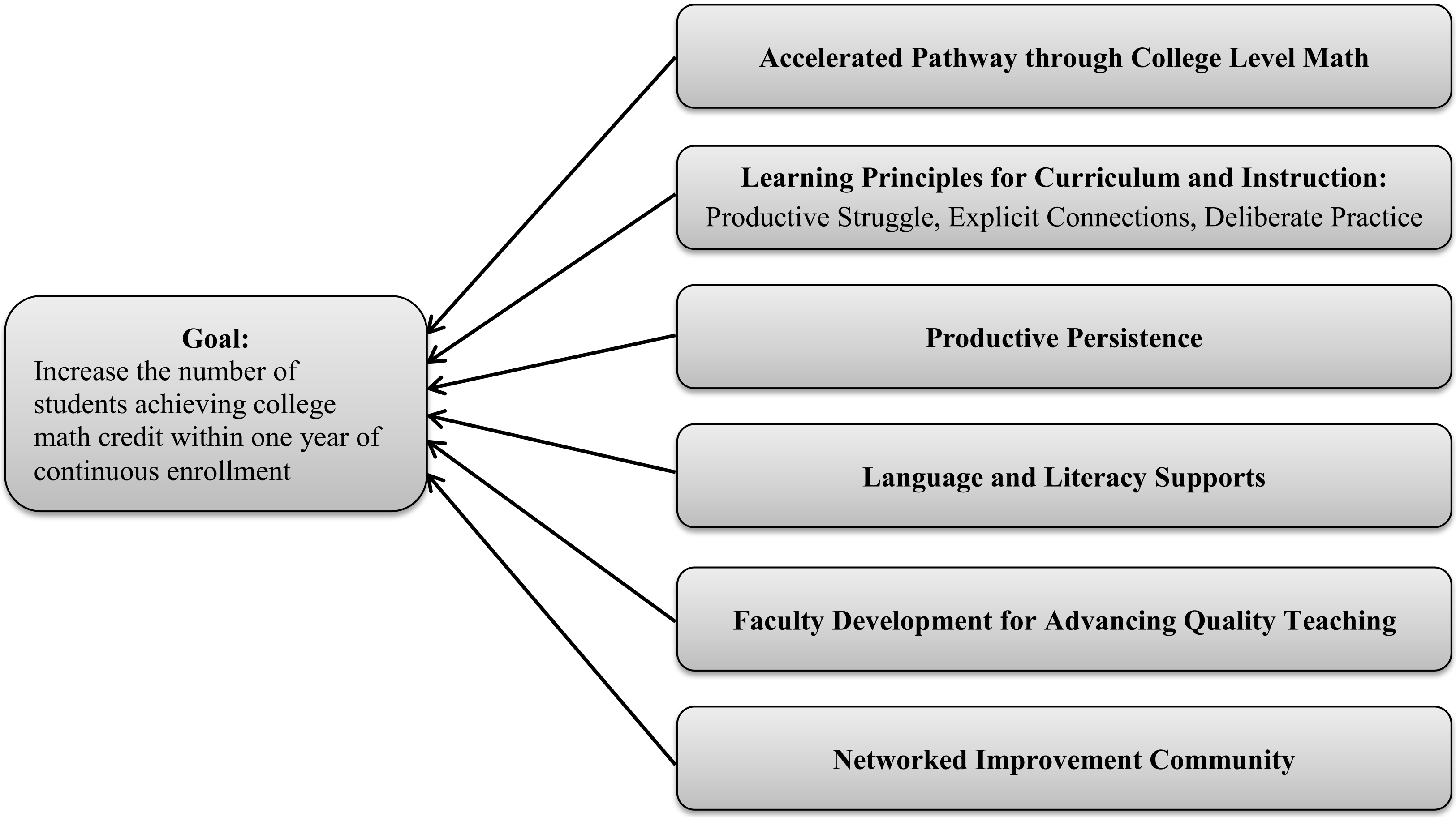

Figure 2. Six key drivers of Statway.

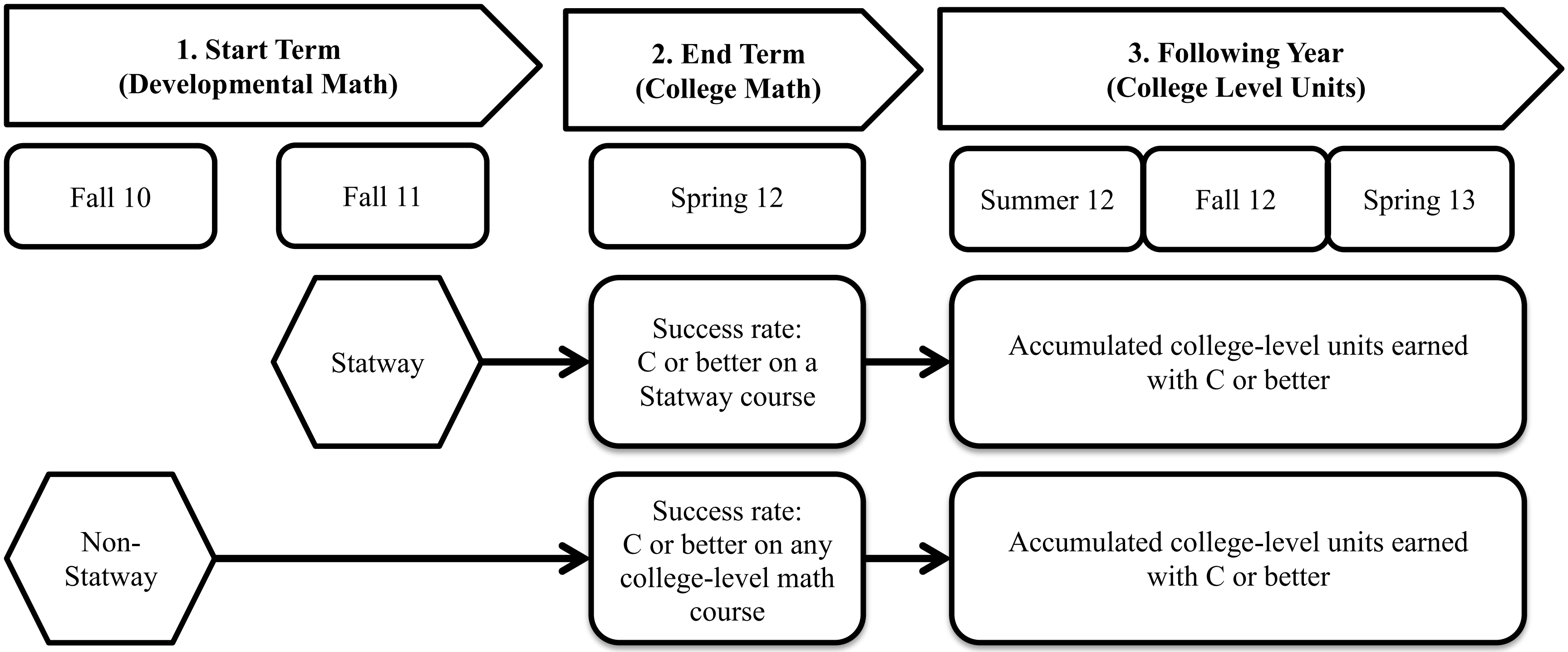

Figure 3. Study design with Year 1 as an exemplar. For 2. End Term, data for Statway students were based on registrar reports from Statway classrooms. For non-Statway students, results were based on whichever college level math course, if any, was successfully completed by Spring 2012. If more than one course was completed during the two years, data from the course (minimum 3 units) with a higher grade were included. For 3. In the following year, if non-Statway students successfully completed a college math course before Spring 2012, their college level units were tracked three terms immediately after the completed term. A grade of C- or higher was employed for six colleges that use a +/– grading system to define college math success and college-level units earned.

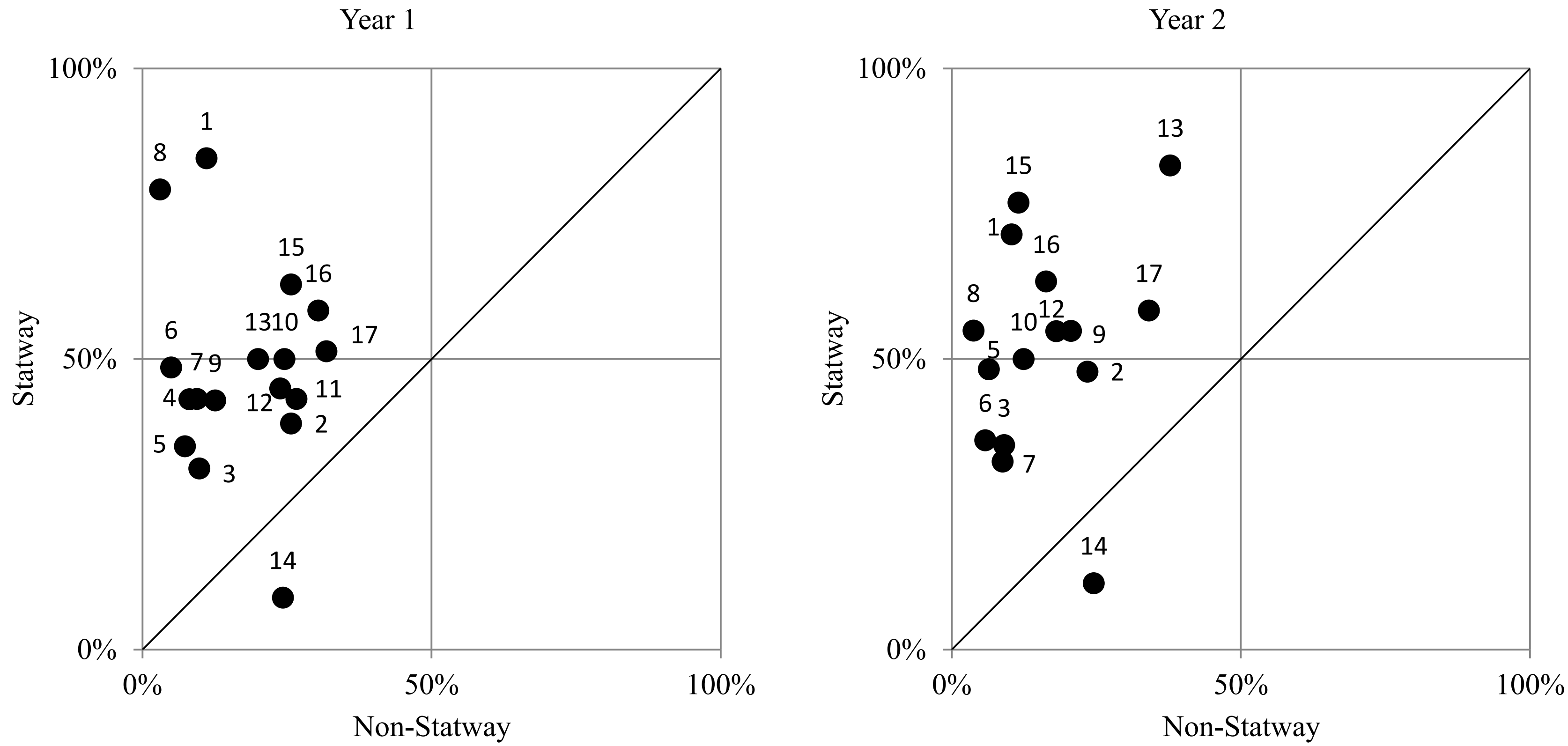

Figure 4. Comparative success rates by college - Year 1 (left) and Year 2 (right). The success rates of the matched comparison students are represented along the x-axis, and those of the Statway student are represented along the y-axis. For ease of interpretation, a 45 degree line is provided as a reference, indicating no difference in outcome (Statway vs. Non-Statway). The numeric values represent pseudo-college IDs.

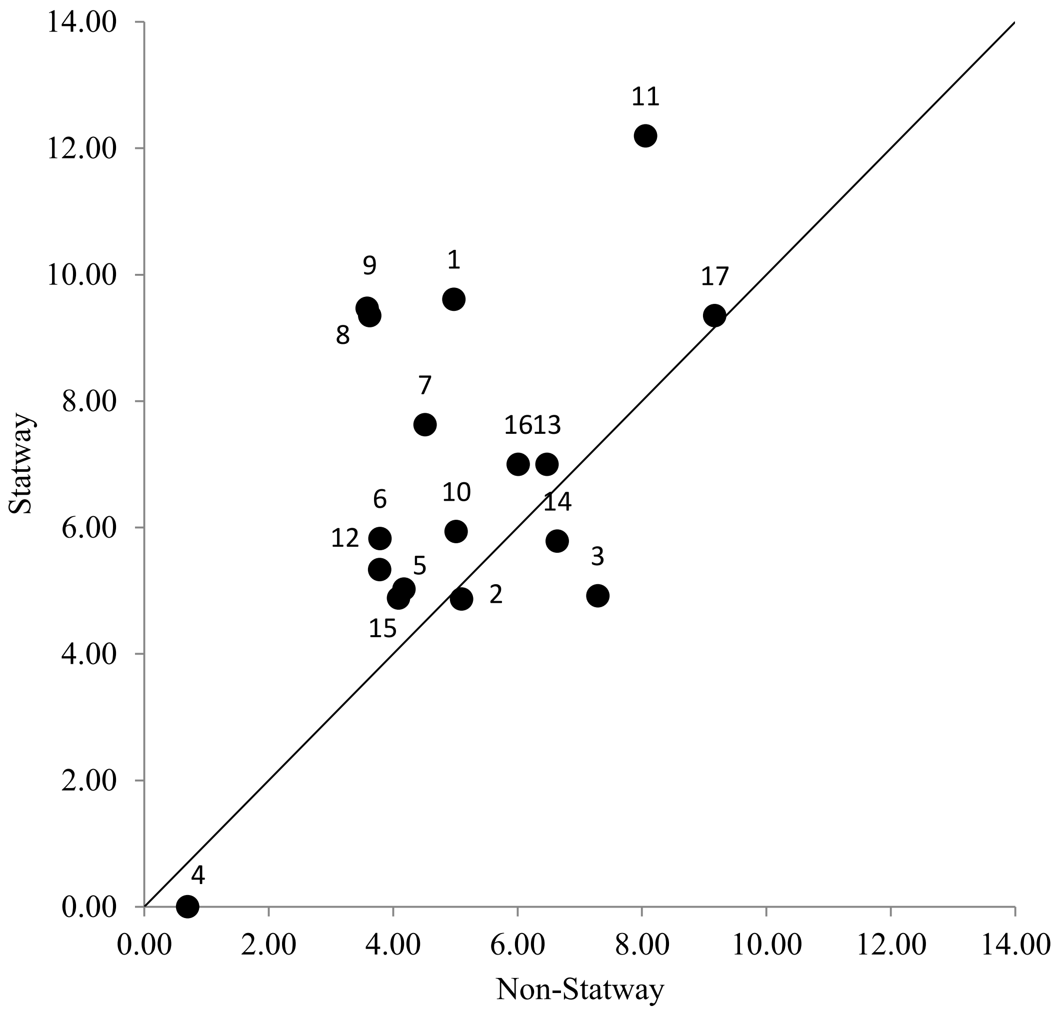

Figure 5. Comparative college-level credits accumulated in the subsequent year by college - Year 1. The college units of the matched comparison students are represented along the x-axis, and those of the Statway student are represented along the y-axis. For ease of interpretation, a 45 degree line is provided as a reference, indicating no difference in outcome (Statway vs. Non-Statway). The numeric values represent pseudo-college IDs.