Carnegie Foundation
for the Advancement of Teaching

Maintaining Success Rates: Does Statway® sustain its impact as it scales to new classrooms and institutions?

February 2017

Melrose Huang and Hiroyuki Yamada
Carnegie Foundation for the Advancement of Teaching
Stanford, CA

CARNEGIE MATH PATHWAYS
TECHNICAL REPORT

Abstract

Statway® is an accelerated developmental mathematics intervention program for college students who are not yet prepared to succeed in a college-level math course. The Carnegie Foundation for the Advancement of Teaching created the initiative to increase student success rates in developmental mathematics and subsequent credit-bearing, college-level mathematics. Statway differs from traditional developmental mathematics sequences in that it allows students to complete both their developmental and college-level mathematics requirements in a single pathway over the course of one academic year. A previous study (Yamada & Bryk, 2016) found that students—across all sex and race/ethnic groups—who enrolled in Statway, had higher rates of completion for both developmental and college-level math compared to carefully matched students in traditional developmental math programs. Statway provides a multifaceted approach to addressing challenges that impede student success through psychosocial and academic supports, as well as lessons based on examples relevant to students' lives. The course is structured so the same cohort of students is enrolled in a consecutive two-term sequence with the same faculty instructor, students work collaboratively, and instructors learn and apply new and more interactive teaching methods. Faculty also receive professional development and support through a networked improvement community (NIC) that connects them with other practitioners, administrators, and researchers from diverse institutions across the nation.

We assessed the effectiveness of the program in its third and fourth years by utilizing a multilevel modeling approach with propensity score matching to compare Statway student outcomes with their counterparts in traditional developmental mathematics programs. We found that in comparison to traditional developmental mathematics students, Statway students had a significantly higher likelihood of successfully attaining college-level mathematics credit—even when comparison group students were allowed twice the amount of time to achieve the same outcome. These results are consistent across all sex and race/ethnic subgroups, as well as in almost all participating classrooms and institutions. Additionally, in replicating earlier findings from the first two years of the program, the results suggest that Statway continues to advance equitable outcomes even as it scales to greater numbers of students in an increasing number of institutions. This report also suggests potential areas for future investigation and research.

This report provides an important sequel to the Carnegie Foundation's previous assessment of the effectiveness of the first two years of Statway®, a mathematics program designed to expedite the process by which students may fulfill their developmental requirements and earn college mathematics credit. Employing a multilevel modeling technique with propensity score matching, the previous peer-reviewed study found that Statway considerably improved college mathematics success rates among students from the 2011 and 2012 cohorts (Yamada & Bryk, 2016). Since then, Statway has continued to scale within current institutions and spread to new ones, nearly doubling its enrollment and expanding its geographic reach to three new states, and expanding in total four-fold since it began five years ago (Huang, Hoang, Yesilyurt, & Thorn, 2016). Scaling innovations in education is notoriously difficult (Cannata & Nguyen, 2015). Thus, the current study, which focuses on the third and fourth years of Statway's implementation, seeks to further substantiate Statway as an evidence-based initiative by replicating results from the previous analysis. To better inform our efforts to address challenges in scaling Statway, the study also examines variation in performance across instructors and institutions.

The Developmental Mathematics Problem

Level of attainment in higher education predicts income and is a primary means of attaining upward social mobility (Deming & Dynarski, 2009). As of 2014, 12.3 million individuals were enrolled in community colleges in order to develop relevant knowledge and skills to participate meaningfully in the workforce (American Association of Community Colleges, 2016). For some, successfully completing community college serves as the final step toward earning an associate degree or professional certification. For others, community college represents a stepping stone for transitioning to a four-year institution and working toward a higher degree. Unfortunately, without first achieving college-level mathematics credit, many of these students can never realize their dreams.

Approximately 60% of community college students are assigned to take one or more developmental mathematics courses. However, three years after enrollment, 80% of these students do not successfully complete the course, thereby precluding them from going on to earn college-level mathematics credit (Bailey, Jeong, & Cho, 2010). Minority students are disproportionately required to take more developmental mathematics courses, but are also less likely to achieve college-level mathematics credit than white students are (Bailey et al., 2010). Developmental mathematics sequences thus act as gatekeepers in granting or denying students, particularly those from traditionally disadvantaged groups, the means of continuing to progress toward their academic and professional goals. Consequently, developmental mathematics reform is one of the most salient education and equity issues of our time.

Prior research suggests that a number of factors in how developmental mathematics courses are traditionally designed and implemented may adversely affect student outcomes. At present, developmental mathematics programs are generally offered as a sequence of foundational courses that increase in difficulty, starting with basic arithmetic and culminating in intermediate algebra (Stigler, Givvin, & Thompson, 2010). Prior to even attempting a college-level mathematics course, students must first successfully make it through this sometimes lengthy sequence (Hodara, 2013). Depending on their placement test scores, students may be required to take three or more of these prerequisite courses; most students, however, are placed at least two levels below college math, meaning a minimum of a year in non-credit-bearing math (Cullinane & Treisman, 2010). Among those who do attempt a developmental mathematics course, even those who successfully complete one or multiple courses, may find themselves struggling to persevere through the entirety of the sequence (Bailey et al., 2010).

Structural barriers to student success and enrollment in developmental mathematics courses may be further exacerbated by a myriad of other factors. For instance, researchers have found little use of evidence-based classroom activities and practices that may encourage student success by facilitating deeper learning (Mesa, 2011; National Research Council, 2002). In addition, much of the material presented in conventional developmental mathematics courses lacks accessibility and relevance by failing to demonstrate real world applications of mathematics or incorporate student interests (Carnevale & Desrochers, 2003; Edwards, Sandoval, & McNamara, 2015). Lamentably, students may believe that successfully learning mathematics is a skill some people are born with, while others are inherently incapable of becoming successful mathematics students. Such beliefs may derive from previous negative experiences with mathematics, which may also induce excessive anxiety among students when they are confronted with challenging concepts (Blackwell, Trzesniewski, & Dweck, 2007; Haynes, Perry, Stupinsky, & Daniels, 2009). Students who identify with traditionally marginalized groups may be especially vulnerable; in particular, African-Americans and females may internalize stereotypes of themselves as poor mathematics students (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). Additionally, students with limited language and literacy skills, such as students from immigrant backgrounds, may find it difficult to effectively communicate mathematical reasoning or understand mathematics problems (Gomez, Rodela, Lozano, & Mancevice, 2013).

Statway's Theory of Improvement

Statway was initiated by the Carnegie Foundation for the Advancement of Teaching as an innovative means to improve upon the generally dismal success rate of students in developmental mathematics. Statway offers students an accelerated pathway by which they can meet their developmental mathematics requirements and achieve college-level mathematics credit in statistics within a single academic year. The theory behind the program's design is underpinned by six key drivers (see Figure 2): 1) an accelerated pathway through college-level mathematics, 2) curriculum and instruction principles that emphasize productive struggle, explicit connections, and deliberate practice, 3) explicit integration of socioemotional supports in the form of productive persistence, 4) language and literacy supports, 5) professional development for faculty, and 6) support of a networked improvement community comprised of researchers, practitioners, college administrators, and designers (Bryk, Gomez, Grunow, & LeMahieu, 2015).

Designing an accelerated pathway through college-level mathematics. Statway offers an accelerated and simplified means by which students can meet their developmental and college-level mathematics requirements. As shown in Figure 1, under traditional developmental mathematics programs, students placed into elementary algebra within the sequence would require one and a half years, at minimum, to achieve college-level mathematics credit. Others, who may repeatedly fail developmental mathematics courses and/or face scheduling conflicts or do not have full course loads due to family and work duties, may require even more time to achieve the same outcome (Bailey et al., 2010; Cullinane & Treisman, 2010). Accordingly, Statway is offered as an intensive two-course sequence that is intended to be taken consecutively by the same cohort of students and with the same faculty member. Its curriculum integrates elementary and intermediate algebra with college-level statistics, and generally meets liberal arts or general education requirements for introductory college-level mathematics at both participating community colleges and four-year universities. In particular, its content is designed to prepare students working toward non-STEM majors (Cullinane & Treisman, 2010; Hodara, 2013).

Instructional emphasis on productive struggle, explicit connections, and deliberate practice. Statway's instructional system is rooted in three evidence-based practices that promote student learning: productive struggle, explicit connections, and deliberate practice. Through productive struggle, students work through rich mathematics assignments that challenge them to make sense of mathematical concepts on their own, while still offering enough scaffolding such that students can feasibly complete the task (Schmidt & Bjork, 1992; Hiebert & Grouws, 2007). The idea of explicit connections suggests that understanding mathematical concepts and procedures in relation to one another can foster deeper learning, whereas traditional mathematics instruction generally places too much emphasis on developing procedural knowledge as discrete process, while offering little to bolster students' conceptual knowledge (Boaler, 1998; Hiebert & Grouws, 2007). Similarly, rather than emphasizing repetitive tasks and procedural knowledge through rote memorization, deliberate practice entails assigning tasks specifically designed to bridge gaps in student knowledge, while enriching student understanding of core concepts (Ericsson, 2008; Ericsson, Krampe, & Tescher-Römer, 1993; Pashler, Rohrer, Cepeda, & Carpenter, 2007).

Integration of socioemotional supports via Productive Persistence. Statway hosts an intervention package made up of established change ideas that foster a sense of productive persistence among students. These activities strive to empower students' belief in themselves as mathematics learners (Dweck, 2006), mitigate their mathematics-induced anxiety (Jamieson, Mendes, Blackstock, & Schmaeder, 2010), and heighten their sense of belonging in the classroom (Walton & Cohen, 2011). Other activities focus on helping students develop good learning strategies (Dweck, Walton, & Cohen, 2011; Yeager & Walton, 2011). The productive persistence change package includes activities such as formalizing team members' roles during group work, committing to contributing to a healthy learning environment and personal engagement via a contract, and completing a reading about the malleability of the brain and its potential to acquire new knowledge and skills.

Inclusion of language and literacy supports. Language and literacy supports are embedded in Statway's curriculum and related materials to assist student acquisition of key vocabulary, identification of core concepts in question prompts and written materials, as well as development of critical thinking and reasoning skills. These structures work in conjunction to promote student understanding and learning, while eliminating language as a barrier to that learning (Gomez et al., 2013; Gomez, Gomez, Rodela, Horton, Cunningham, & Ambrocio, 2015).

Addressing faculty professional development needs. Instructors accustomed to teaching traditional mathematics courses may find it difficult to adapt to Statway's holistic and multifaceted approach. Accordingly, onboarding and continuous professional development opportunities are offered to assist faculty members in transitioning from a teacher-centered perspective to a student-centered perspective that is less dependent upon lecture-heavy instruction, focusing instead on group work and discussion of mathematics problems (Edwards et al. 2015; Grubb, 1999; Grubb & Grabiner, 2013; Mesa, 2011). Avenues for professional development include attending in-person trainings, participating in regular online sync meetings, and mentoring by faculty members with more experience in successfully implementing Statway.

Support of a networked improvement community. As discussed earlier in this report, Statway is structured as a networked improvement community. Within the NIC, instructors, administrators, and researchers work together on a measurable and specific aim. That work is supported by research of the underlying problem and the system through which it exists, as well as a working theory on how to address the problem (Bryk, Gomez, & Grunow, 2011; Dolle, Gomez, Russell, & Bryk, 2013). Members of the NIC do this by leveraging the key tenets of improvement science, including iteratively developing, testing, and refining ideas for change. Learning is communicated throughout and accelerated by the network, which works to incorporate validated change ideas into the extant system. In the case of Statway, the NIC provides an opportunity for postsecondary practitioners, improvement coaches, and educational researchers to collaborate both within and across institutions. As a learning community, they implement common measures, procedures, and concepts to achieve quantifiable improvement in Statway at both the outcome and process levels (Bryk et al., 2015).

In light of the previous study's promising results and the evolving nature of the Statway initiative, the aim of the present study was to determine whether Statway's initial positive results were sustainable beyond the first two years of its establishment. In particular, our focus centered on whether Statway continued to be effective in the face of scaling and operating across an increasingly diverse range of contexts. To investigate these areas of interest, we used a hierarchical linear modeling (HLM) approach (Raudenbush & Bryk, 2002) with propensity score matching (Rosenbaum & Rubin, 1983) to increase both the validity of causal inference and the quality of the matched sample by accounting for potential selection bias in Statway group membership (e.g., students with certain characteristics may have been more likely to participate in Statway, which may have artificially inflated its success rates). The HLM analyses also address the nested structure of our sample (i.e., students enrolled in classrooms within colleges), which may be relatively homogeneous compared to a random sample, and thus manifest problems when using conventional linear modeling techniques (Osborne, 2000). We then compared college mathematics performance between Statway and the non-Statway matched comparison groups, again applying an HLM approach. Although matching was conducted separately for Year 3 and Year 4, data from the two cohorts were combined prior to examining college mathematics credit achievement.

Beyond assessing the average effect of Statway, we also explored variation in performance to ensure that Statway benefits a wide spectrum of students, instructors, and institutions within the NIC. This is extremely important from an improvement science perspective as it seeks to ensure effective implementation for all participants and across diverse contexts, in contrast to typical attributional analyses that seek only to examine mean effects across populations. Additionally, we analyzed student performance across a variety of sex and race/ethnic categories to determine Statway's capacity to achieve equitable outcomes for all. These analytic results will be examined to identify systemic shortcomings and prioritize strategic action within the NIC to further improve student success and reliably achieve quality at scale (Bryk et al., 2015).

Method

Participants

Students were drawn from the Year 3 and Year 4 Statway cohorts that began the sequence in the fall of 2013 and 2014, respectively. The Year 3 cohort consisted of 2,283 students enrolled in 158 sections taught by 70 faculty members (Sowers & Yamada, 2015). These students attended 19 community colleges and three four-year universities across seven states: California, Connecticut, Florida, Minnesota, New York, Texas, and Washington. (For more information, please see Appendix). The Year 4 cohort comprised 2,862 students enrolled in 166 sections taught by 81 faculty members across 22 community colleges and four four-year universities in 10 states, including the aforementioned states with the addition of Hawaii, Massachusetts, and Wyoming (Huang et al., 2016). Across these two cohorts (N = 5,145), more than 60% of the students were female and nearly 70% were minorities.

Data Collection and Study Design

The institutional researchers of each participating college supplied data on student background characteristics, prior course enrollment, and performance. The Year 3 analytic sample consisted of 1,422 students enrolled across 18 community colleges and two four-year universities. The Year 4 analytic sample comprised 1,284 students across 12 community colleges and one four-year university.[1]

Figure 3 illustrates the standard design used in this study. It followed the previous study design (Yamada & Bryk, 2016) to take into account the amount of time students who enrolled in traditional developmental mathematics sequences typically require to attain college-level mathematics credit (see Figure 1). Specifically, we selected students who enrolled in developmental mathematics courses one year prior to Statway students. This method provided

[1] Due to a lack of or insufficient institutional data, two colleges were excluded from the Year 3 analyses. In Year 4, three participating colleges only offered the first course within Statway's two-course sequence; accordingly, they were excluded from the analytic sample. An additional ten colleges were excluded from the Year 4 analyses due to either a lack of or insufficient institutional data. The data used in our analysis were submitted prior to November 2016.

comparison group students with twice the amount of time to achieve college-level mathematics credit as their Statway counterparts, and allowed us to measure the final outcome for each group at the same time point. Thus, we compared students who began traditional developmental mathematics in Fall 2012 with students who began Statway in Fall 2013, and compared students who began traditional developmental mathematics in Fall 2013 with students who began Statway in Fall 2014. College-level mathematics achievement was then measured for each cohort in Spring 2014 and Spring 2015, respectively. We followed this approach for students in all participating institutions with the exception of those who attended two community colleges, one of which only implemented an accelerated one-term form of Statway, and another which implemented both the standard one-year and accelerated one-term forms of Statway. That is, Statway students in these institutions were to complete the full Statway sequence and acquire college-level mathematics credit in a single academic term. Thus, for the purpose of these analyses, students in these institutions' comparison groups were provided a full academic year to achieve the same outcomes.

We utilized a two-level hierarchical linear model (HLM) to calculate propensity scores for each student. Propensity score matching is a statistical technique applied to observational data to 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). There are two main steps involved in this procedure: (a) obtain a propensity score per student, which represents the likelihood of a student enrolling in Statway, and (b) identify students whose propensity scores are similar to each other as matches. Typically, a logistic regression approach is used with a set of factors or covariates hypothesized to influence student enrollment in Statway.

We selected a total of 32 student-level covariates including student background characteristics, as well as prior course taking and success patterns during the two years prior to the Statway term, to generate propensity scores from HLM (Hong & Raudenbush, 2005, 2006; Raudenbush & Bryk, 2002; Yamada & Bryk, 2016). These covariates were selected based on prior research findings and advice from institutional researchers in the participating colleges. Specifically, student background characteristics including sex and race/ethnicity were included because they may influence student performance in developmental mathematics (Bailey et al., 2010). Prior course records were also taken into account because they may serve as more accurate proxies of students' professional and educational goals than declared academic majors (Jenkins & Cho, 2012). Table 1 displays a comprehensive list of covariates as well as their descriptive statistics. Some student background data were not available in institutional records, and were therefore treated as "unknown" when matching students. Prior to matching in both Years 3 and 4, Statway students were more likely to be female and/or Hispanic than non-Statway students. Additionally, Statway students were generally more likely to have previously attempted and successfully completed developmental English and other college-level courses outside of STEM than comparison group students.

For step two, we employed a nearest neighbor matching algorithm to conduct propensity score matching separately for each participating institution (Rosenbaum & Rubin, 1985). We selected this approach in order to preserve as many Statway students as possible while drawing their closest matches from a large batch of potential comparison group students. For each Statway student, we found up to five matches (5:1 ratio matching) in order to select the most proximal matches from the non-Statway population without sacrificing precision (Ming & Rosenbaum, 2000). To achieve this, we set a caliper distance of up to 0.2 as the maximum acceptable difference between matched students to mitigate the possibility of bad nearest neighbor matches based on recommendations in the literature (Austin, 2011; Rosenbaum & Rubin, 1985). For propensity score matching, we used the package MatchIt (Ho, Imai, King, & Stuart, 2011) in the statistics analytics system R (R Core Team, 2015). For more details regarding propensity score matching, see Yamada et al. (2016). The sample of comparison students was 82.165 prior to matching; 2,607 out of 2,706 Statway students were matched with 10,089 comparison group students across both years and all schools.

As in the earlier analysis of Quantway (Yamada, Bohannon, & Grunow, 2016), we applied a four-level HLM model to estimate the effectiveness of Statway against that of traditional developmental mathematics programs. Our dichotomous outcome was defined as follows: successful students either completed Statway with a grade of C or higher within a single academic year (or as aforementioned, within a single academic term for students enrolled in accelerated Statway courses), or college-level mathematics with a grade of C or higher within two academic years (or as previously described, within a single academic year for non-Statway students matched with students enrolled in accelerated Statway courses). Matched clusters (level 1) were nested within Statway students (level 2), who were nested within their instructor classrooms (level 3), who were in turn were nested within their respective institutions (level 4). We used HLM 7 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) for all of the HLM analyses.

Analytic Approach

Propensity Score Matching

We utilized a two-level Bernoulli model and estimated its model parameters using maximum likelihood via adaptive Gaussian quadrature. ϕ_{ij} represents the probability of student *i* enrolling in Statway in college *j* while η_{ij} indicates the log-odds of this event, which may be formally expressed as:

Level-1 Model (Student)

$$\mathrm{Prob}(SW_{il}=1|\beta_l) = \phi_{il}$$

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

$$\eta_{il} = \beta_{0l} + \beta_{1l}(COV1_{il}) + \ldots + \beta_{37l}(COV32i_l),$$

Level-2 Model (College)

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

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

.

.

.

$$B_{32l} = \gamma_{320},$$

where *SW* is a dummy variable indicating whether students were in enrolled in Statway (coded as 1) or in a traditional developmental mathematics sequence (coded as 0), *COV1…COV32* represent all covariates, *i* denotes a student, and *j* is a college. We matched a total of 5,853 comparison students to 1,353 Statway students for Year 3 and a total of 4,240 comparison students matched to 1,254 Statway students for Year 4.[2] Thus, there were a total of 12,700 cases included in the analysis reported here. 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 3 and 4. For both cohorts, there were no significant differences in mean propensity scores between the Statway and matched students in any of the colleges (see *t*-values in Table 2). Taken together, this provides strong evidence that comparability of the groups was achieved on the measured covariates.

Estimating Statway Effects

To estimate Statway's effect on student success, we ran a four-level Bernoulli model and estimated model parameters using maximum likelihood via penalized quasi-likelihood estimation. ϕ_{ijkl} represents the probability that student *i* in Statway student *j*'s matched group enrolled in instructor *k's* course within college *l* successfully achieved college mathematics credit. Correspondingly, η_{ijkl} is the log-odds of this outcome and formally expressed as:

Level-1 Model

$$\mathrm{Prob}(SUCC_{ijkl}=1|\pi_{jkl}) = \phi_{ijkl}$$
$$\log[\phi_{ijkl}/(1 - \phi_{ijkl})] = \eta_{ijkl}$$
$$\eta_{ijkl} = \pi_{0jkl} + \pi_{1jkl}*(SW_{ijkl}) + \pi_{2jkl}*(PS_{ijkl}),$$

Level-2 Model

$$\pi_{0jkl} = \beta_{00kl} + r_{0jkl}$$
$$\pi_{1jkl} = \beta_{10kl} + r_{1jkl}$$
$$\pi_{2jkl} = \beta_{20kl},$$

Level-3 Model

$$\beta_{00kl} = \gamma_{000l} + u_{00kl}$$
$$\beta_{10kl} = \gamma_{100l} + u_{10kl}$$
$$\beta_{20kl} = \gamma_{200l},$$

Level-4 Model

$$\gamma_{000l} = \delta_{0000} + v_{000l}$$
$$\gamma_{100l} = \delta_{1000} + v_{100l}$$
$$\gamma_{200l} = \delta_{2000},$$

where *SUCC* represents college mathematics credit achievement (1 for success and 0 for non-success), 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

[2] While we attempted to match five non-Statway students to each Statway student, we lowered our matching ratio at ten institutions in Cohort 3 and eight institutions in Cohort 4. These reduced matching ratios ranged from 1:1 to 4:1, as illustrated in Table 2. We also dropped select Statway students for whom we were unable to find appropriate matches from the original analytic sample.

included individual students' propensity scores, *PS*, as an additional adjustment variable.[3]

Results

Overall, Statway students were significantly more likely to succeed at earning college level math credits compared to their peers. The findings, presented in Table 3, indicate that on average, Statway students demonstrated significantly higher odds of success, 4.96 (95% CI [3.42, 7.19][4]), in achieving college-level mathematics credit than their matched counterparts. This translated into the estimated probabilities of success of 58% for the Statway group and 22% for the comparison group.[5] In Table 3, -0.51 and -0.85 represent the estimated correlation coefficients between the intercept and slope at the college and faculty levels, indicating that the Statway effect is greater when the outcome of matched non-Statway students is lower. While this relationship was observed at both the college and faculty levels, it was stronger in the latter.

Variation in Performance

Statway had a positive effect across all institutions, as illustrated by Figure 4, in which the center line represents the average Statway effect across all students. The lines above and below the main line represent the upper and lower bounds of the average Statway effects and are positioned two standard errors away from the center line. Figure 4 also shows that Statway effects varied at the institutional level (with a variance component of 0.51). For instance, Colleges 11 and 12 were positioned below the lower bound of the center line, showing smaller Statway effects relative to other colleges. On the other hand, College 1 was above the upper bound and could be seen as a positive deviant.

All faculty, with the exception of instructor 11_01, saw an overall positive effect of Statway on student outcomes. This finding is illustrated by Figure 5, in which the center line again represents the average Statway effect across all students. Similarly, the lines above and below the main line represent the upper and lower bounds of the average Statway effects and are positioned two standard errors away from the center line. Figure 5 also depicts faculty-level variation in Statway effect (with a variance component of 0.21). For example, the Statway effect tended to be homogeneous at College 1, where instructors consistently achieved student success within their respective classes. In contrast, College 16 exhibited the greatest faculty-level variation: some of its instructors demonstrated effects slightly below the NIC average whereas others saw effects well above average. Our findings suggest that it may be of interest to investigate ways to reduce the substantial variation in Statway effects both between and

---

[3] We initially included binary covariates indicating institution type (2 year versus 4 year institutions), mode of adaptation (Statway offered as a standard two-course sequence versus an accelerated one-course sequence), and cohort (Year 3 versus Year 4) at the intercept and slope for levels 4 and 2. However, our initial model did not indicate that these factors had any significant effects on student success. To maintain a parsimonious model, we therefore excluded these non-significant covariates in our final analysis. However, these findings do not necessarily preclude the possibility of differences in student outcomes between different modes of adaptation or institutional types.

[4] HLM 7 produces 95% confidence intervals of odds ratios.

[5] We also conducted a sensitivity analysis (Hong & Raudenbush, 2005) on Statway effects on college math achievement. 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.

within colleges.

Subgroup Analyses

We constructed a four-level HLM to account for potential differences in the Statway effect between sex and race/ethnicity subgroups. In order to represent main and interaction effects on our outcome of interest, we utilized effect coding for each subgroup variable. The respective reference categories for sex and race/ethnicity groups were female and White, which were each coded as -1. We excluded students of unknown sex from the analytic sample.

To enhance interpretability, we converted the model-based results to proportions of students successfully earning college mathematics credit, as displayed in Figure 6. We found large positive Statway effects for both female and male students, as well as for all race/ethnicity subgroups. Black females demonstrated the largest gain in college mathematics achievement relative to the baseline performance. Our results suggest that although Statway advances equitable outcomes for diverse groups of students, those from certain groups may benefit more from the program than others. Accordingly, future improvement efforts may focus on reducing the gap in Statway effects between various race/ethnicity subgroups.

Discussion

Our study provides an assessment of Statway's effectiveness for post-secondary students in Years 3 and 4 of the initiative by implementing a propensity score matching technique (Rosenbaum & Rubin, 1983) within a hierarchical linear modeling framework (Raudenbush & Bryk, 2002). This rigorous analytic approach allowed us to examine the relationship between the Statway program and student success in achieving college-level mathematics credit. We maintained a high matching ratio (approaching 1:5) between Statway students and traditional developmental mathematics students while matching students across 32 covariates. We found that even when providing non-Statway students with twice the amount of time to achieve the same outcome as their Statway counterparts, Statway students exhibited significantly higher probability of success (often approaching odds that were twice as great). Additionally, we conducted a sensitivity analysis and found no evidence that the estimated effects were affected by confounding factors not previously addressed by our analyses. While Statway has both scaled within institutions and spread to new ones, our results replicated those from an earlier analysis on the first two years of Statway (Yamada & Bryk, 2016). Our findings held across all sex and race/ethnicity subgroupings, suggesting that Statway promotes equitable outcomes for all students, including females and students of color, who are historically at higher risk of failing developmental mathematics (Bailey et al., 2010; Carnevale & Desrochers, 2003; Chen, 2016; National Research Council, 2002). Strikingly, the results from our subgroup analysis indicate that Statway may be most beneficial to minority students.

Statway's multifaceted approach to developmental mathematics reform may be key to its evidenced ability to improve student success across diverse populations and conditions (Hodara, 2013). Notably, Statway's accelerated format, which may be easier to navigate than traditional developmental mathematics sequences (Bailey et al., 2010), is bolstered by additional foci on productive persistence, a quality system of instruction, relevant content, support for faculty learning, and networked improvement. Students, especially those of color, may suffer from negative academic stereotypes and may question whether they truly belong or will be respected in a college setting (Center for Community College Student Engagement, 2010; Gardenhire-Crooks, Collado, Martin, & Castro, 2010; Walton & Cohen, 2011). Accordingly, Statway promotes productive persistence, which addresses a number of such psychosocial factors. These include developing a growth mindset toward learning, academic tenacity, good learning strategies, and a sense of classroom belonging—all of which have been demonstrated to empower student motivation, engagement, and therefore success (Beilock, 2011; Dweck, 2006; Stigler & Heibert, 1999; Yeager & Dweck, 2012). Furthermore, the program's instructional system fosters quality learning by drawing upon the research-based learning principles of productive struggle, explicit connections among concepts, and deliberate practice (Boaler, 1998; Ericsson, 2008; Ericsson et al., 1993; Hiebert & Grouws, 2007; Pashler et al., 2007; Schmidt & Bjork, 1992). The initiative also addresses the needs of many traditionally disadvantaged students by providing language and literacy support as well as contextualized lessons and culturally relevant pedagogy (Gomez et al., 2013; Gomez et al., 2015; Morest, V.S., 2014; Muro, A., 2015). Not only does Statway advance learning at the student level, but it also equips instructors with ample training and professional development resources to ensure faculty success (Edwards et al., 2015; Grubb, 1999; Grubb & Grabiner, 2013). Moreover, Statway bridges administrators, practitioners, researchers, and improvement experts in a NIC to encourage collaboration, learning, and sharing of best practices across institutions (Bryk et al., 2011, Dolle et al., 2013). For example, several instructors across participating institutions have formed teams for the purposes of revising Statway curricula, developing a more user-friendly onboarding system for new faculty, iteratively testing change ideas for building productive persistence, and improving student assessments.

One of the foundational principles of improvement science posits that achieving quality improvement requires reducing variation in performance. Accordingly, we investigated variation in Statway's effect among participating faculty members and institutions. As mentioned above, we found positive Statway effects for all institutions and all but one faculty member. Significant variation in performance also existed both across and within colleges. Some instructors and colleges represented positive outliers that outperformed others, while other faculty members and institutions underperformed (relative to the rest). By leveraging the primary function of a NIC to collaboratively learn how to improve (Bryk et al., 2015), Carnegie Math Pathways team members plan on conducting qualitative interviews with faculty members at colleges that showed above average performance in this analysis to illuminate promising instructional and supportive practices that promote student success. For instance, positive deviant instructors within College 16 could potentially lead training sessions on how to implement additional homework interventions for underperforming colleagues, thereby improving instructional practices within their respective institutions. On the other hand, College 1, whose instructors were all positive deviants, may well have initiated practices at the institutional level to promote such success across the board. College 1 might have designed, prototyped, and refined additional faculty development opportunities that it could share with other Statway institutions. By testing and scaling innovative practices among its members, the NIC has a unique opportunity to continuously improve Statway's design, implementation, and student outcomes.

Although we did not find any single mode of adaptation or institution type to significantly affect Statway outcomes, there may still be differences between accelerated and standard Statway formats, as well as two-year and four-year institutions, that merit further investigation. Specifically, institution type may not have been significant due to a relatively small number of participating four-year institutions, (which composed only two of the twenty colleges in our sample). In addition, other analyses of relevant data suggest that four-year universities tend to have higher success rates than community colleges (Huang et al., 2016). As we continue to scale to four-year institutions, including large universities, we will have the opportunity to examine whether they owe high success rates to student body characteristics or distinct teaching and/or administrative practices. Descriptive data also indicate that students taking an accelerated one-course form of Statway generally exhibit success rates at least equal to their standard two-term Statway counterparts. However, one institution only offered accelerated Statway, rendering it difficult to disentangle college-level effects from those induced by mode of adaptation. In this case, it may be of interest to compare post-participation outcomes between students from the aforementioned institution and students who took Statway as a standard two-course sequence. Should accelerated Statway students go on to be just as successful as their standard counterparts in subsequent years, it may be worthwhile to look into whether a more intensive design could be established as an option for implementing Statway.

Overall, we find that Statway students in the third and fourth years of the program were not only more likely to achieve college-level mathematics credit than their comparison group counterparts, but were also able to do so in half the time. These findings replicate those of a previous analysis conducted on the first two years of Statway. By leveraging a rigorous multi-level modeling technique with propensity score matching, our findings offer further evidence of the causal relationship between Statway and student success, suggesting that the program provides a viable alternative means of navigating the developmental mathematics sequence. Notably, Statway continues to benefit students even as the program spreads to a growing number of students and colleges. Our findings indicate that Statway’s unique and multifaceted approach to developmental mathematics reform enables students, instructors, and institutions to overcome impediments to student success across a diverse range of contexts.

References

American Association of Community Colleges. Community college fact sheet. (2016). Retrieved November 9, 2016, from www.aacc.nche.edu.

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.

Beilock, S. (2011). Back to school: Dealing with academic stress. *Psychological Science Agenda*, *25*(9).

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.

Bonham, B. S., & Boylan, H. R. (2011). Developmental mathematics: Challenges, promising practices, and recent initiatives. *Journal of Developmental Education*, *34*(3), 2.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In *Frontiers in sociology of education* (pp. 127-162). Springer Netherlands.

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.

Cannata, M., & Nguyen, T. (2015). *Consensus versus Clarity: Tensions in Designing for Scale.* Paper presented at the annual meeting of the American Educational Research Association in Chicago, IL, April 16-21, 2015.

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

Center for Community College Student Engagement. (2010). *The Heart of Student Success: Teaching, Learning, and College Completion (2010 CCCSE Findings)*. Austin, TX: University of Texas at Austin, Community College Leadership Program.

Chen, X. (2016). *Remedial coursetaking at U.S. public 2- and 4-year institutions: Scope, experiences, and outcomes* (NCES 2016-405). U.S. Department of Education, Washington, DC: National Center for Education Statistics.

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.
Deming, D., & Dynarski, S. (2009). *Into college, out of poverty? Policies to increase the postsecondary attainment of the poor* (No. w15387). National Bureau of Economic 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.
Gardenhire-Crooks, A., Collado, H., Martin, K., & Castro, A. (2010). Terms of engagement: Men of color discuss their experiences in community college Educational Resources Information Center (ERIC) Document No. *ED508982. New York: MDRC*.
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.
Ho, D.E., Imai, K., King, G., & Stuart, E.A. (2011). MatchIt: Nonparametric Preprocessing for

Parametric Causal Inference. *Journal of Statistical Software, 42 (8)*, 1-28.

Hodara, M. (2013). *Improving students' college mathematics 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.

Huang, M., Hoang, H., Yesilyurt, S., & Thorn, C. (2016). *Community College Pathways: 2014-2015 Impact Report.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

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.

Mesa, V. (2011). Similarities and differences in classroom interaction between remedial and college mathematics courses in a community college. *Journal on Excellence in College Teaching, 22*(4), 21-55.

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

Morest, V. S. (2014). The Community's College: Contextualizing Service-Learning in the Mission and Development of Community Colleges. In Service-Learning at the American Community College (pp. 37-49). Palgrave Macmillan US.

Muro, A. (2015). Retention and Completion of Underrepresented Students: A Look at the Seattle Colleges (Doctoral dissertation, University of Washington).

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

Osborne, J. W. (2000). The Advantages of Hierarchical Linear Modeling. ERIC/AE Digest.

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.

R Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage

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.

Sowers, N., & Yamada, H. (2015). Community College Pathways: 2013-2014 Descriptive Report. *Carnegie Foundation for the Advancement of Teaching*.

Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *MathAMATYC Educator*, *1*(3), 4-16.

Stigler, J. & Hiebert, J. (1999). The teaching gap. New York: The Free Press.

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

Yamada, H., Bohannon, A., & Grunow, A. (2016). *Assessing the Effectiveness of Quantway®: A Multilevel Model with Propensity Score Matching.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Yamada, H., & Bryk, A. S. (2016). Assessing the first two years' effectiveness of Statway®: A multilevel model with propensity score matching. Community College Review, 44, 179-204.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302-314.

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 Colleges

- American River College
- Austin Community College
- Bunker Hill Community College
- California State University, Los Angeles
- California State University, Sacramento
- California State University, San Francisco
- California State University, San Jose
- Capital Community College
- De Anza College
- Diablo Valley College
- Foothill College
- LaGuardia Community College
- Laramie County Community College
- Los Angeles Pierce College
- Maui College
- Minneapolis Community and Technical College
- Mountain View College
- Mt. San Antonio College
- Naugatuck Valley Community College
- Normandale Community College
- North Hennepin Community College
- Richland College
- San Diego City College
- Seattle Central Community College
- South Seattle Community College
- Tacoma Community College
- Tallahassee Community College
- University of Washington, Bothell
- Valencia College

Table 1.

Descriptive Statistics of Covariates in the Two-Level Propensity Model

Year 3	Sample before matching		Sample after matching	
	Non-Statway	Statway	Non-Statway	Statway
	%	%	%	%
Sex				
Female*	54	63	61	61
Male	42	37	38	39
Unknown	4	>1	1	>1
Race/Ethnicity				
Asian	7	6	6	6
Black	18	22	19	18
Hispanic	31	35	38	39
White*	30	29	28	28
Multiracial	1	3	3	3
Other	1	1	1	2
Unknown	11	4	5	5
Any course records in past two years				
No*	42	41	38	41
Yes	58	59	62	59

	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (in years)	24.01	8.06	25.35	9.73	25.21	9.30	25.51	8.18
Terms since first developmental math course	1.52	2.30	1.69	2.54	1.75	2.47	1.75	2.46
Course load	3.25	1.21	3.23	1.17	3.22	1.25	3.19	1.14
Developmental math								
One level below college level								
Number of courses attempted	0.18	0.50	0.20	0.58	0.21	0.56	0.21	0.59
Success rate	0.04	0.20	0.02	0.15	0.03	0.17	0.03	0.16
Two levels below college level								
Number of courses attempted	0.31	0.62	0.32	0.69	0.34	0.67	0.33	0.70
Success rate	0.15	0.34	0.13	0.33	0.14	0.33	0.14	0.33
Three or more levels below college level								
Number of courses attempted	0.21	0.53	0.21	0.60	0.22	0.55	0.22	0.61
Success rate	0.12	0.32	0.13	0.33	0.13	0.33	0.13	0.33
Developmental English								
Number of courses attempted	0.17	0.69	0.20	0.57	0.19	0.56	0.21	0.59
Success rate	0.07	0.25	0.12	0.32	0.11	0.31	0.12	0.32
Developmental reading								
Number of courses attempted	0.05	0.27	0.04	0.21	0.04	0.22	0.04	021
Success rate	0.03	0.17	0.03	0.17	0.04	0.18	0.03	0.18
Developmental writing								
Number of courses attempted	0.03	0.22	0.05	0.28	0.06	0.29	0.05	0.28
Success rate	0.02	0.15	0.03	0.18	0.04	0.20	0.04	0.18
College math								
Number of courses attempted	0.02	0.22	0.06	0.31	0.05	0.30	0.06	0.30
Success rate	0.01	0.08	0.01	0.10	0.01	0.10	0.01	0.10
College non-math								

Number of courses attempted	2.33	3.85	2.89	4.71	3.02	4.71	2.95	4.72
Success rate	0.33	0.42	0.37	0.44	0.44	0.42	0.39	0.44
College STEM								
Number of courses attempted	0.27	0.93	0.40	1.24	0.41	1.22	0.41	1.24
Success rate	0.08	0.27	0.11	0.29	0.12	0.31	0.11	0.30
College non-STEM								
Number of courses attempted	2.07	3.50	2.55	4.33	2.66	4.29	2.60	4.34
Success rate	0.33	0.43	0.36	0.44	0.38	0.44	0.37	0.44

Year 4	Sample before matching		Sample after matching	
	Non-Statway	Statway	Non-Statway	Statway
	%	%	%	%
Sex				
Female*	56	64	60	64
Male	43	35	39	35
Unknown	>1	>1	1	>1
Race/Ethnicity				
Asian	7	7	3	4
Black	20	21	32	30
Hispanic	31	40	21	26
White*	34	26	36	33
Multiracial	2	2	1	1
Other	1	1	0	0
Unknown	5	4	6	5

Any course records in past two years								
No*	43		38		46		41	
Yes	57		62		54		59	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (in years)	24.79	8.59	25.07	9.49	24.95	9.14	25.07	9.51
Terms since first developmental math course	1.64	2.43	1.60	2.61	1.70	2.56	1.58	2.59
Course load	3.10	1.21	3.32	1.54	3.15	1.20	3.32	1.55
Developmental math								
One level below college level								
Number of courses attempted	0.15	0.47	0.23	0.58	0.24	0.59	0.22	0.57
Success rate	0.02	0.15	0.03	0.16	0.03	0.17	0.02	0.15
Two levels below college level								
Number of courses attempted	0.30	0.64	0.24	0.57	0.26	0.56	0.24	0.58
Success rate	0.13	0.33	0.12	0.31	0.11	0.31	0.12	0.32
Three or more levels below college level								
Number of courses attempted	0.27	0.59	0.15	0.50	0.18	0.49	0.16	0.51
Success rate	0.16	0.36	0.09	0.29	0.12	0.32	0.10	0.29
Developmental English								
Number of courses attempted	0.19	0.80	0.23	0.65	0.15	0.58	0.23	0.66
Success rate	0.08	0.26	0.13	0.33	0.08	0.28	0.13	0.33
Developmental reading								
Number of courses attempted	0.04	0.23	0.03	0.18	0.03	0.20	0.02	0.17
Success rate	0.02	0.15	0.02	0.14	0.03	0.17	0.02	0.14
Developmental writing								
Number of courses attempted	0.05	0.27	0.03	0.19	0.04	0.23	0.03	0.20
Success rate	0.03	0.17	0.03	0.16	0.03	0.17	0.03	0.16
College math								

Number of courses attempted	0.02	0.19	0.05	0.29	0.04	0.25	0.04	0.24
Success rate	0.01	0.08	0.01	0.09	0.01	0.09	0.01	0.09
College non-math								
Number of courses attempted	2.06	3.71	2.80	4.85	3.00	4.82	2.77	4.82
Success rate	0.27	0.40	0.37	0.44	0.35	0.43	0.36	0.44
College STEM								
Number of courses attempted	0.28	0.95	0.29	0.82	0.34	1.04	0.28	0.82
Success rate	0.08	0.26	0.08	0.26	0.10	0.28	0.08	0.26
College non-STEM								
Number of courses attempted	1.80	3.38	2.57	4.64	2.69	4.49	2.54	4.59
Success rate	0.26	0.40	0.37	0.44	0.34	0.43	0.36	0.44

Note. "*" denotes covariates that served as reference categories (coded as 0) for creating dummy variables. Birthdates were subtracted from Sept. 1, 2013 and Sept. 1, 2014 to calculate the ages of Statway students in Years 3 and 4, respectively. For non-Statway students, Sept. 1, 2012 and Sept. 1, 2013 were used to compute age. These ages were then centered around age 18 in the propensity score model. In terms since first developmental math course, 0 indicates that a student took her/his first developmental math course (including Statway) in the term concurrent with the Start term,, and 1 indicates one term before the Start term, etc. Course load represents the total number of courses a student was enrolled in while taking the first Statway course within the sequence; for comparison group students, it refers to the total number of courses a student was enrolled in during the Start term. Success rates were calculated by dividing the total number of courses successfully completed over the total number of attempted courses. In this case, success was defined as receiving a grade of C or higher (C- or higher in schools that used +/- grading systems) or pass within a pass/fail grading scheme.

Table 2.
Balance in Logit of the Propensity Score for non-Statway and Statway Students

		Sample before matching						Sample after matching							
		Non-Statway			Statway			Non-Statway			Statway				
College	Cohort	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*t*	Matched ratio
1	Fall 2013	3321	-4.13	1.59	98	-3.24	0.68	490	-3.29	0.66	98	-3.24	0.68	-0.63	5.00
1	Fall 2014	4961	-4.02	0.56	103	-3.88	0.49	515	-3.88	0.49	103	-3.88	0.49	-0.11	5.00
2	Fall 2013	2977	-4.31	0.59	46	-3.99	0.59	230	-4.00	0.58	46	-3.99	0.59	-0.14	5.00
3	Fall 2013	2704	-4.03	0.73	60	-3.73	1.03	286	-3.63	0.82	59	-3.69	0.99	0.41	4.85
3	Fall 2014	2384	-3.55	0.70	87	-2.99	0.65	432	-3.01	0.63	87	-2.99	0.65	-0.27	4.97
4	Fall 2013	842	-5.29	2.48	30	-2.47	0.70	147	-2.57	0.53	30	-2.47	0.70	-0.70	4.90
5	Fall 2013	2083	-2.11	0.44	280	-1.90	0.51	807	-1.96	0.45	277	-1.92	0.49	-1.18	2.91
5	Fall 2014	1685	-1.48	0.65	494	-0.90	0.61	474	-0.97	0.56	474	-0.94	0.59	-0.77	1.00
6	Fall 2013	3254	-4.21	0.62	58	-3.93	0.60	290	-3.94	0.59	58	-3.93	0.60	-0.09	5.00
6	Fall 2014	2489	-3.85	0.63	63	-3.49	0.57	311	-3.52	0.54	63	-3.49	0.57	-0.29	4.94
7	Fall 2013	419	-3.20	1.57	80	-0.30	0.24	70	-0.47	0.11	25	-0.46	0.11	-0.46	2.80
8	Fall 2013	146	-1.37	0.27	39	-1.38	0.29	121	-1.37	0.26	39	-1.38	0.29	0.34	3.10
8	Fall 2014	145	-1.11	0.23	50	-1.09	0.24	141	-1.11	0.22	48	-1.08	0.23	-0.86	2.94
9	Fall 2013	1236	-4.00	1.72	45	-3.33	0.59	225	-3.27	0.54	45	-3.33	0.59	0.63	5.00
10	Fall 2013	1452	-5.09	1.22	12	-4.06	0.66	60	-4.12	0.62	12	-4.06	0.66	-0.29	5.00
11	Fall 2013	2243	-3.49	0.63	83	-2.87	0.93	364	-3.09	0.71	82	-2.90	0.90	-1.02	4.44
11	Fall 2014	2304	-3.72	0.46	62	-3.50	0.72	299	-3.58	0.63	61	-3.54	0.67	-0.38	4.90
12	Fall 2013	10362	-4.65	0.53	113	-4.53	0.52	565	-4.53	0.51	113	-4.53	0.52	-0.08	5.00
13	Fall 2013	2024	-3.48	0.59	73	-3.37	0.59	365	-3.37	0.57	73	-3.37	0.59	-0.02	5.00
13	Fall 2014	1785	-3.48	0.69	68	-3.27	0.56	335	-3.30	0.47	67	-3.30	0.48	-0.02	5.00

14	Fall 2013	1647	-3.39	0.56	65	-3.18	0.59	325	-3.18	0.58	65	-3.18	0.59	0.05	5.00
14	Fall 2014	1631	-3.64	0.59	51	-3.40	0.60	251	-3.44	0.56	51	-3.40	0.60	-0.44	4.92
15	Fall 2013	1904	-3.58	0.59	63	-3.23	0.55	315	-3.24	0.54	63	-3.23	0.55	-0.06	5.00
15	Fall 2014	1900	-3.66	0.62	59	-3.29	0.48	295	-3.30	0.47	59	-3.29	0.48	-0.15	5.00
16	Fall 2013	2966	-3.53	0.59	104	-3.03	0.66	498	-3.10	0.59	103	-3.04	0.64	-0.76	4.83
16	Fall 2014	3074	-3.23	0.57	148	-2.73	0.97	702	-2.89	0.65	144	-2.85	0.71	-0.74	4.88
17	Fall 2013	6771	-5.30	0.59	39	-5.04	0.49	195	-5.06	0.49	39	-5.04	0.49	-0.16	5.00
17	Fall 2014	6258	-5.92	0.51	18	-5.62	0.56	90	-5.63	0.54	18	-5.62	0.56	-0.02	5.00
18	Fall 2013	1120	-3.16	0.27	51	-2.66	0.60	184	-2.93	0.29	43	-2.84	0.36	-1.65	4.28
18	Fall 2014	2878	-4.35	0.52	42	-4.44	0.34	210	-4.44	0.33	42	-4.44	0.34	-0.03	5.00
19	Fall 2013	355	-2.79	1.61	45	-1.62	0.79	128	-1.79	0.67	45	-1.62	0.79	-1.29	2.84
20	Fall 2013	1598	-4.80	1.95	38	-3.17	0.87	188	-3.25	0.79	38	-3.17	0.87	-0.49	4.95
20	Fall 2014	1247	-3.72	0.64	39	-3.35	1.03	181	-3.56	0.68	37	-3.50	0.81	-0.44	4.90

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

Fixed effect	Coefficient	*SE*	*t*	*p*	Odds ratio
Intercept	-1.44	0.15	-9.60	<0.001	0.24
Propensity score	0.16	0.03	4.69	<0.001	1.17
Statway effect	1.60	0.18	8.73	<0.001	4.96
Random effect at level 4 (college)	Variance	*df*	χ^2	*p*	Correlation
Intercept	0.43	19	507.38	<0.001	-0.51
Statway effect	0.51	19	106.25	<0.001	
Random effect at level 3 (faculty)	Variance	*df*	χ^2	*p*	Correlation
Intercept	0.00	52	56.72	0.300	-0.85
Statway effect	0.21	52	117.71	<0.001	

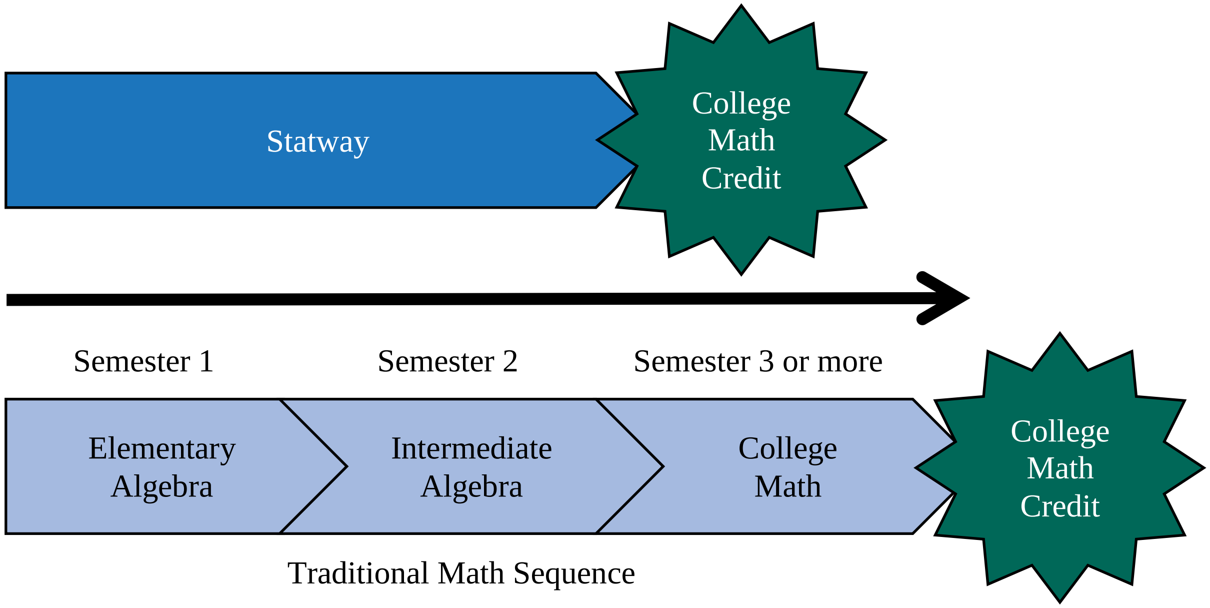

Figure 1. Statway vs. Traditional mathematics sequence

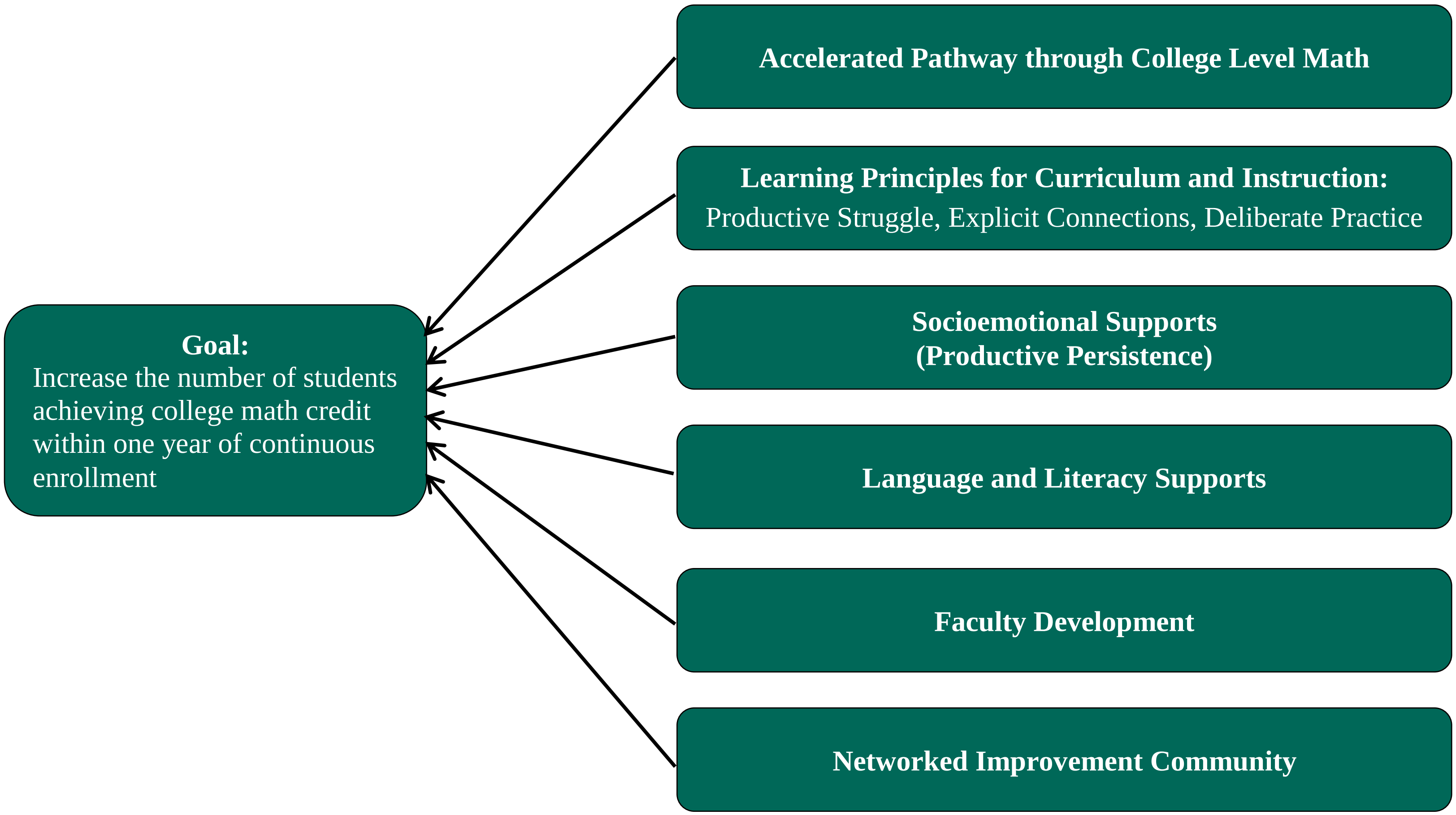

Figure 2. Six key drivers of Statway.

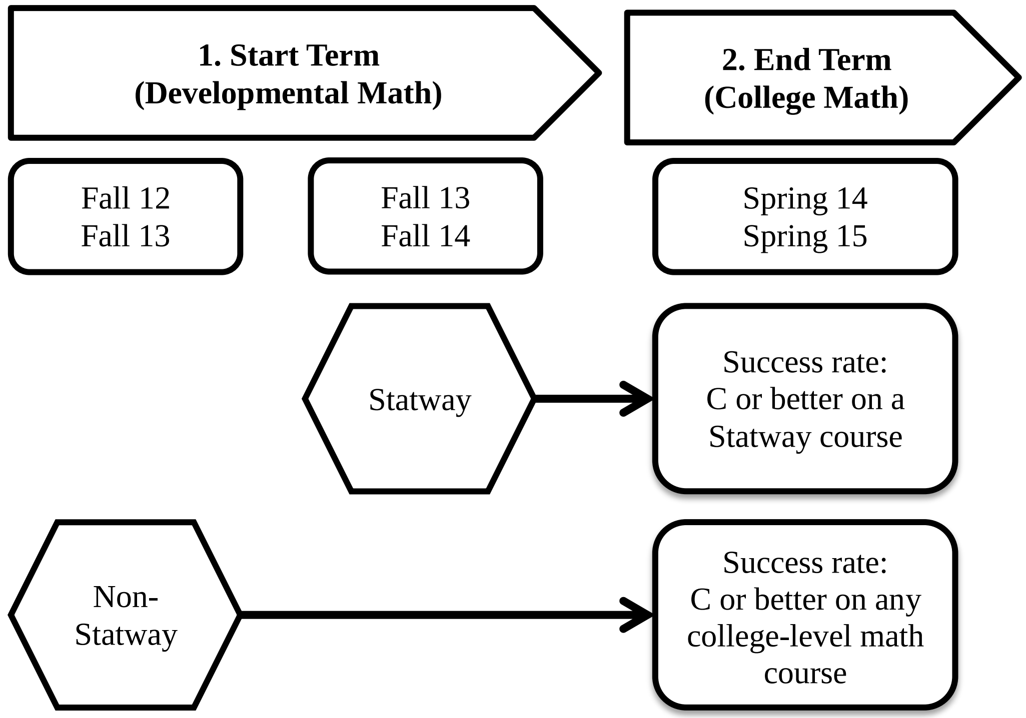

Figure 3. Standard study design.
Outcome data for Statway students were based on grade records submitted by institutional researchers. For non-Statway students, results were based on whichever college level mathematics course, if any, was successfully completed by Spring 2014 and Spring 2015 for Year 3 and Year 4, respectively. If more than one course was completed during the two years, data from the course (minimum 3 units) with a higher grade were included immediately after the completed term. A grade of C or higher (or C- or higher for colleges that follow a +/– grading system) was used to define college mathematics success.

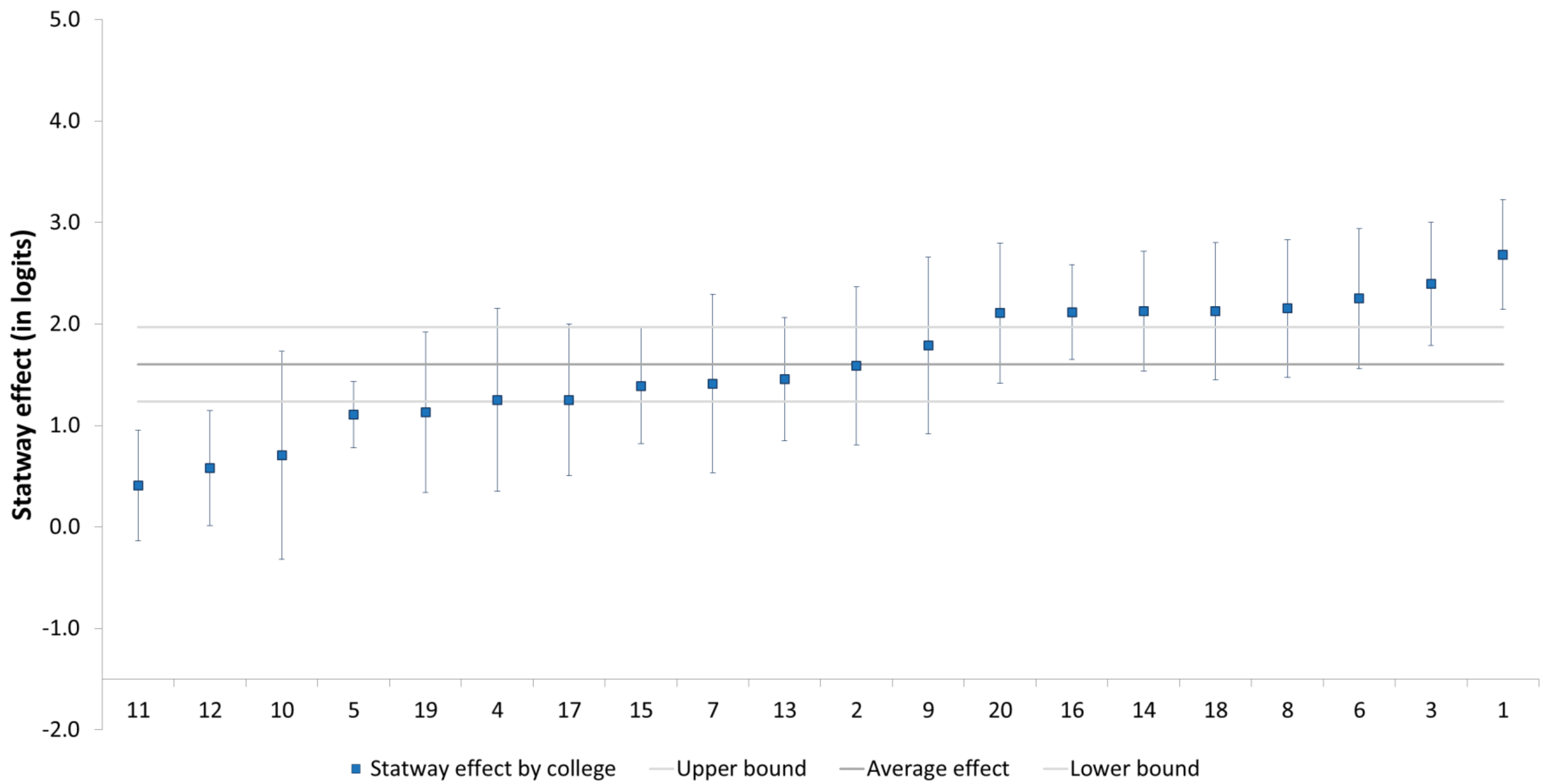

Figure 4. Variation in Statway effect among colleges

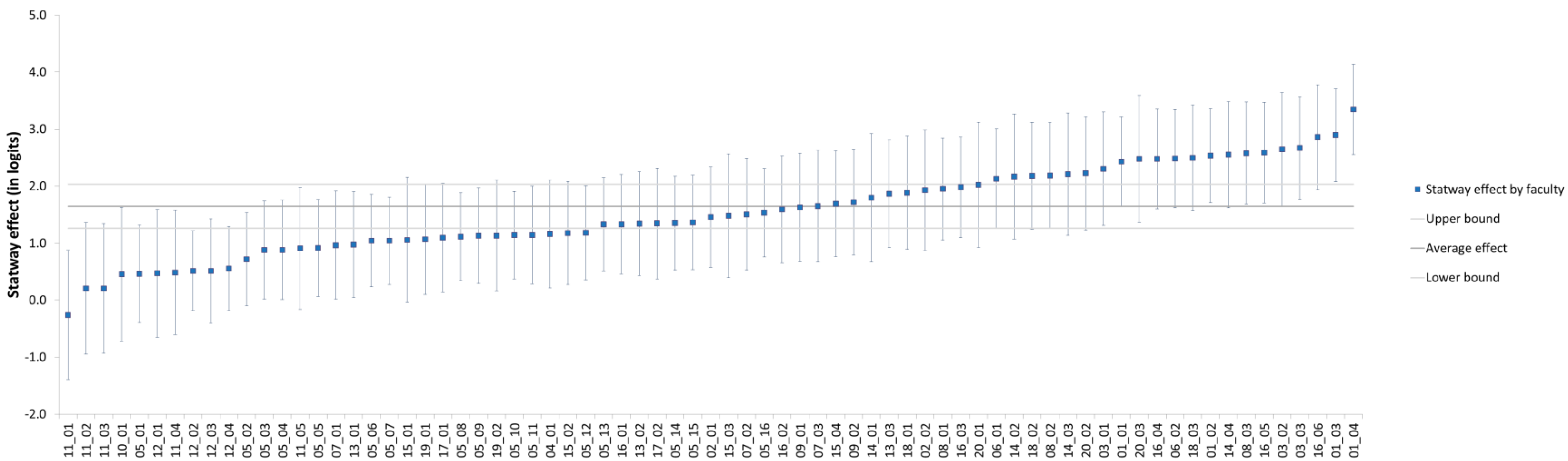

Figure 5. Variation in Statway effect among faculty members

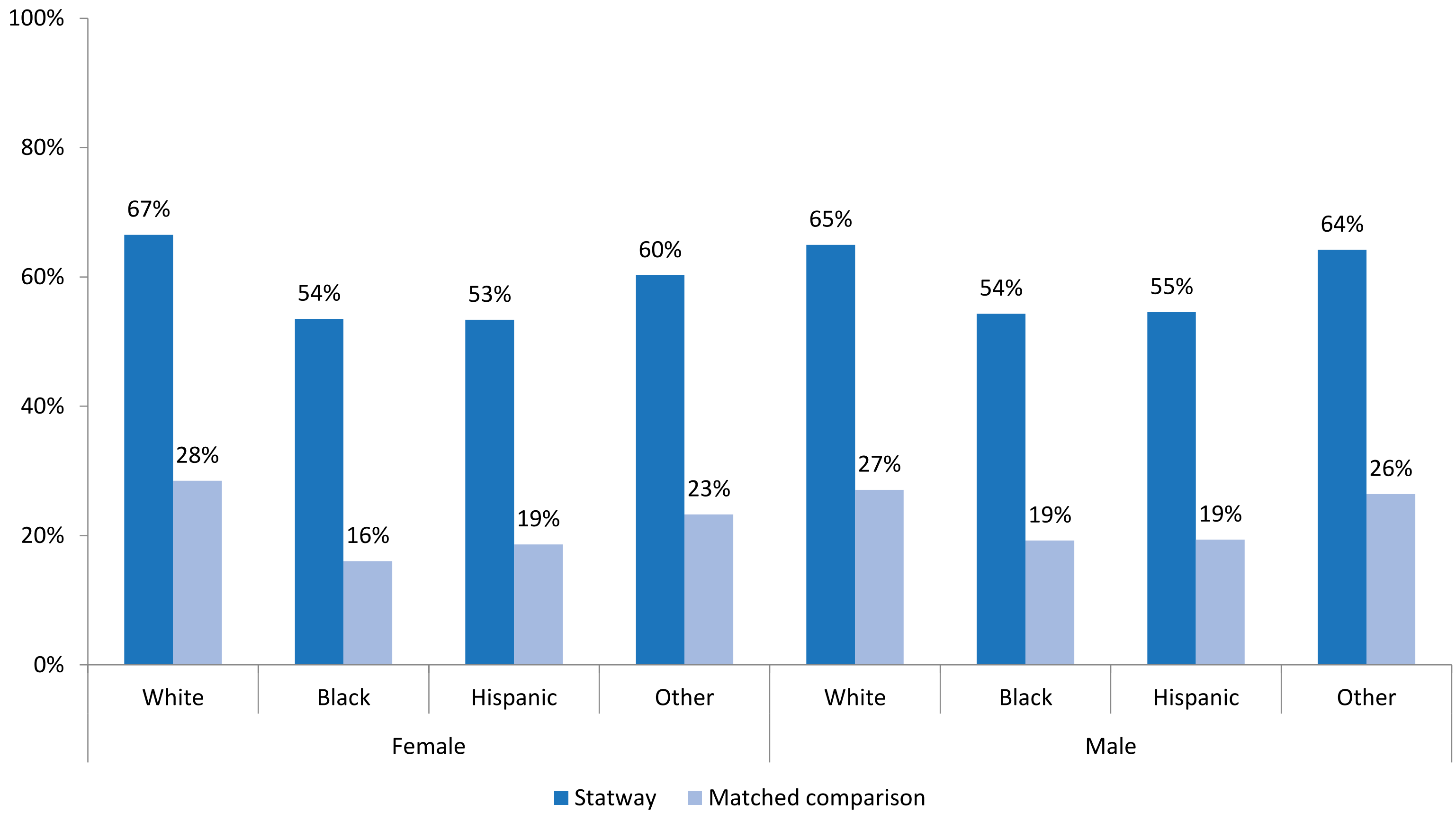

Figure 6. Model-based success rates by sex and race/ethnicity

Carnegie Foundation for the Advancement of Teaching
51 Vista Lane
Stanford, California 94305
650-566-5100

Carnegie Foundation for the Advancement of Teaching seeks to vitalize more productive research and development in education. We bring scholars, practitioners, innovators, designers, and developers together to solve practical problems of schooling that diminish our nation's ability to educate all students well. We are committed to developing networks of ideas, expertise, and action aimed at improving teaching and learning and strengthening the institutions in which this occurs. Our core belief is that much more can be accomplished together than even the best of us can accomplish alone.

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 in cooperation with the Carnegie Foundation for the Advancement of Teaching.

www.carnegiefoundation.org

We invite you to explore our website, where you will find resources relevant to our programs and publications as well as current information about our Board of Directors, funders, and staff.