Article

Community College Review
2018, Vol. 46(3) 257–287

DOI: 10.1177/0091552118771754
journals.sagepub.com/home/crw
SAGE

Assessing the Effectiveness of Quantway®: A Multilevel Model With Propensity Score Matching

Hiroyuki Yamada[1], Angel X. Bohannon[1], Alicia Grunow[1], and Christopher A. Thorn[1]

Abstract

Objective: Quantway is a Carnegie Math Pathways initiative, which redesigns the content, pedagogy, and structure of traditional developmental mathematics courses to simultaneously tackle traditional barriers to student success and support a broad range of developmental students in achieving their mathematics potential. Specifically, Quantway is a quantitative reasoning sequence that is comprised of a single term accelerated developmental mathematics course called Quantway 1 and a college-level mathematics course called Quantway 2. This study assesses the effectiveness of the developmental mathematics course, Quantway 1, during its first six semesters of implementation. **Method:** We used a hierarchical linear modeling technique to conduct propensity score matching across 37 student characteristics to compare the course performance of Quantway 1 students with matched comparison students from traditional developmental mathematics courses. **Results:** Quantway 1 students demonstrated significantly higher odds of success in fulfilling developmental mathematics course requirements and enrolling in college mathematics courses in the following year than matched comparison students. In addition, Quantway 1 effects were positive across all sex and race/ethnicity subgroups as well as in nearly all classrooms and colleges. **Contributions:** This study provides robust evidence that Quantway 1 increases student success in fulfilling developmental mathematics requirements and advances equity in student outcomes. Implications of and future directions for the Pathways are discussed.

Keywords

causal inference, multilevel modeling, networked improvement community, propensity score matching, remedial mathematics reform

[1]Carnegie Foundation for the Advancement of Teaching, Stanford, CA, USA

Corresponding Author:
Hiroyuki Yamada, Carnegie Foundation for the Advancement of Teaching, 51 Vista Lane, Stanford, CA 94305, USA.
Email: yamada@carnegiefoundation.org

Traditional developmental or remedial mathematics sequences serve as an impediment for community college students, often preventing them from obtaining technical credentials and associate degrees, as well as blocking their transfer to 4-year institutions. Nearly 60% of community college students nationwide are required to take at least one developmental mathematics course, and 80% of these students do not complete a college mathematics course within 3 years (Bailey, Jeong, & Cho, 2010), which could be translated to more than 500,000 students who fail to earn college mathematics credit (Hoang, Huang, Sulcer, & Yesilyurt, 2017). Students spend long periods of time repeating courses and accruing student loan debt, ultimately leaving college without a degree and the personal (Johnstone, 2013), societal (Carnevale, Smith, & Strohl, 2013), and economic benefits that come with such credentials.

Long remedial course sequences, which typically involve arithmetic, pre-algebra, elementary algebra, and intermediate algebra, are not working as expected (Merseth, 2011). Research indicates that whereas remedial mathematics programs are highly effective at resolving skill deficits for students who remediate successfully, the majority of those who need remediation do not make it through the sequences (Bahr, 2008; Waycaster, 2001). This finding suggests a need to improve the remedial process for those who struggle. Accordingly, one reform effort focuses on accelerating student progress through developmental course sequences by reorganizing instruction and curricula (Edgecombe, 2011). There are two main acceleration models: One offers only the content students need to succeed in college mathematics and the other involves mainstreaming students into college-level mathematics courses with an additional support course. Another reform effort addresses mathematics instruction and pedagogy to improve students' course completion and learning outcomes (Hodara, 2011; Stigler, Givvin, & Thompson, 2010). As promising instructional practices, instructors are encouraged to focus more on conceptual reasoning than procedural knowledge, utilize structured forms of student collaboration and group work, and represent the same mathematical problem in varied ways (i.e., numerically, graphically, and algebraically). In addition to these structural and pedagogical changes to remedial mathematics education, other reform efforts have addressed language and socioemotional supports for students (Merseth, 2011).

There are important equity implications associated with the low success rates we observe in traditional developmental mathematics course sequences. Traditionally underserved students are disproportionately likely to encounter developmental mathematics as a stumbling block on the road to community college completion. The community college student population is more racially diverse, older, and lower income than 4-year university students (Bueschel, 2004). Minority students are placed in more developmental mathematics courses and less likely to complete these courses to achieve college-level mathematics credit than White students (Bailey et al., 2010; Chen, 2016). Improving the success rates of students in developmental mathematics sequences is thus a key lever for improving access for all students. The goal is not merely to improve success rates but to eliminate achievement gaps traditionally associated with income and race/ethnicity.

Several aspects of traditional developmental mathematics sequences have been proposed as contributors to negative student outcomes. Students must take long sequences of courses with increasing levels of difficulty to fulfill developmental mathematics requirements (Hodara, 2013). This structure drastically hinders student completion, and even when students complete one course in a sequence, many fail to enroll in subsequent courses (Bailey et al., 2010; Cullinane & Treisman, 2010). The utility of the algebra-heavy content of traditional mathematics curricula has also been called into question. A study on the Survey of Workplace Skills, Technology, and Management Practices found that only 19% of employees use any algebra in their work (Handel, 2007). The instruction in many mathematics classrooms does not incorporate research-based curriculum design and pedagogical practices that foster deeper student learning and engagement (Mesa, 2011). Traditional mathematics courses emphasize transmission of content over a more participatory approach (Edwards, Sandoval, & McNamara, 2015), factual and procedural knowledge over conceptual knowledge (Mesa, 2011), and do not demonstrate the relevance of mathematical concepts (Carnevale & Desrochers, 2003). Furthermore, these courses do not address either language and literacy or noncognitive barriers (e.g., belief about mathematics ability and anxiety) that impede many students' ability to learn mathematics (Blackwell, Trzesniewski, & Dweck, 2007; Gomez, Rodela, Lozano, & Mancevice, 2013; Haynes, Perry, Stupnisky, & Daniels, 2009). On her extensive review on interventions in college mathematics readiness, Hodara (2013) proposed that we might need a more holistic approach that simultaneously addresses these structural, curricular, pedagogical, linguistic, and noncognitive hindrances to student success. More recent work reinforces those recommendations (Bailey et al., 2016).

Carnegie Math Pathways' Theory of Improvement

To spur progress on this problem, the Carnegie Foundation for the Advancement of Teaching convened a networked improvement community (NIC)—a national collective of mathematical scholarly society representatives, community college administrators and faculty, educational researchers, and improvement specialists (Bryk, Gomez, Grunow, & LeMahieu, 2015). Through an improvement science approach, the NIC redesigned the content, pedagogy, and structure of traditional mathematics sequences to increase the number of students completing their mathematics requirements. The outcome of this work was a specific set of design principles[1] used to create two accelerated alternatives to traditional developmental mathematics sequences, Statway and Quantway, which are intended for non-STEM (science, technology, engineering, and mathematics) students who were placed into two levels below college-level mathematics. Figure 1 summarizes the Carnegie Math Pathways' theory of improvement, aiming to increase student success through working on six key drivers: (a) acceleration of developmental mathematics requirements, (b) implementation of a research-based instructional system, (c) socioemotional supports (productive persistence), (d) language and literacy supports, (e) faculty development, and (f) participation in a NIC.

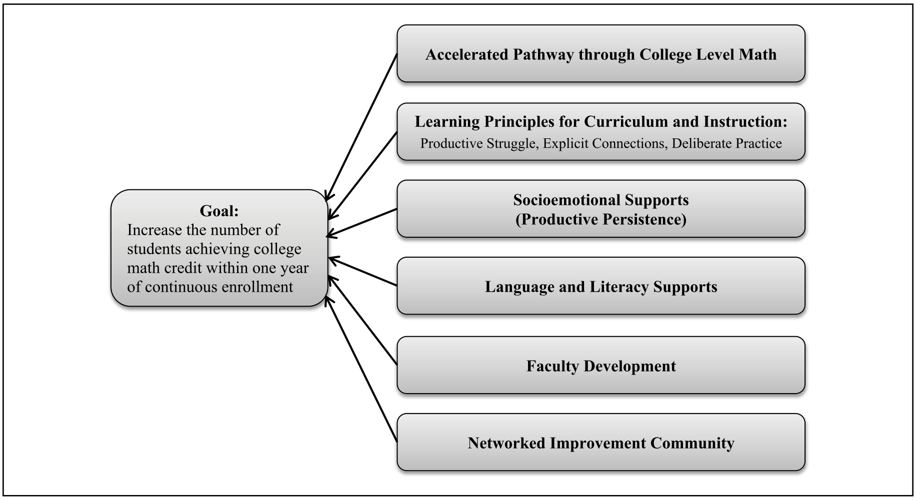

Figure 1. Six key drivers of Carnegie Math Pathways.

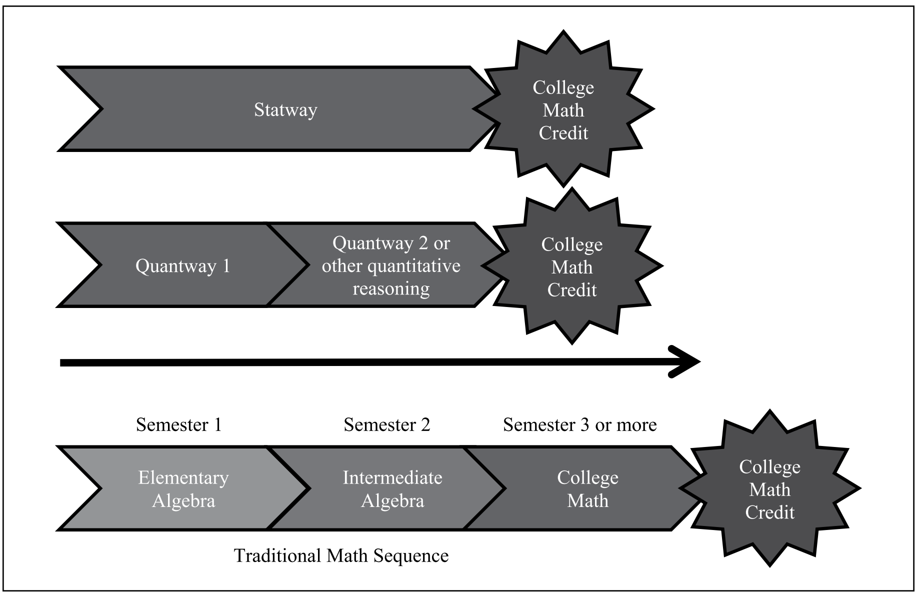

Figure 2. Carnegie Math Pathways Versus traditional math sequence.

Figure 2 visualizes the first driver of accelerated pathways (Edgecombe, 2011). Statway is an accelerated yearlong introductory college-level statistics course that integrates developmental mathematics content. A previous study demonstrated

Statway's efficacy and impact on student success over traditional developmental mathematics programs (Yamada & Bryk, 2016). In contrast, Quantway focuses on quantitative literacy, which is described as "the ability to adequately use elementary mathematical tools to interpret and manipulate quantitative data and ideas that arise in an individual's private, civic, and work life" (Gillman, 2004, p. 5). These quantitative literacy concepts are codified in a set of rigorous learning outcomes that were collaboratively established and vetted by a committee that included representatives from several mathematical professional societies (Blair & Getz, 2011).[2] Because Quantway's learning outcomes provide students with a strong foundation in numerical and quantitative reasoning concepts, it has served as a preparatory quantitative reasoning course for many non-STEM pathways and majors and as the culminating mathematics course for technical certificate programs. Quantway 1 enables students to complete their developmental mathematics requirements in a single term. Those who successfully complete Quantway 1 are then prepared for college-level mathematics and eligible to enroll in Quantway 2 or another college-level quantitative reasoning course. Some colleges offer Quantway as a pathway through college-level mathematics by combining Quantway 1 with Quantway 2 or another college-level quantitative reasoning course. Because Quantway has such a flexible design, and changes to developmental course offerings generally have limited implications for articulation or transfer that can delay the adoption of credit-bearing courses, community colleges can more easily integrate Quantway into their current institutional structures. Accordingly, since its launch in 2012, Quantway 1's implementation has grown more rapidly than Statway (Huang, Hoang, Yesilyurt, & Thorn, 2016). This study thus focused on the effectiveness of Quantway 1.[3]

To address the second driver in Figure 2 (Hodara, 2011; Stigler et al., 2010), Quantway 1's instructional system is designed to foster robust and sustained mathematical learning, emphasizing the teaching of concepts to improve both procedural and conceptual understanding (Hiebert & Grouws, 2007). Its instructional model is organized around three research-based learning opportunities: productive struggle, explicit connections, and deliberate practice. In productive struggle, faculty engage students in substantive mathematical tasks that encourage students to wrestle with key mathematical concepts and solve problems that are challenging but still within reach (Hiebert & Grouws, 2007). By productively struggling, students can make meaning of the mathematical content for themselves and develop strategies for engaging with the content. Explicit connections refers to instruction that creates opportunities for students to forge associations between mathematical procedures and underlying conceptual knowledge. Deliberate practice aims to improve student performance through a series of highly structured, increasingly sophisticated, and challenging tasks that deepen facility with key concepts (Edwards & Beattie, 2016). These learning opportunities are supported by instructional practices that facilitate student discussion and support collaborative learning around rich mathematical problems (Edwards & Beattie, 2016; Edwards et al., 2015). Quantway 1's instructional system is also designed to ground unfamiliar mathematics concepts in familiar situations through contextualization. Quantway 1's lessons use authentic, relevant contexts and real data to increase student motivation to learn. The course is organized around three intentional themes (citizenship, health care, and financial literacy) that reflect everyday concepts and are critically important in engaging in society. By illustrating the real world applications of mathematics concepts, Quantway 1 can empower previously unsuccessful students to have meaningful and positive interactions with quantitative reasoning content.

The third and fourth drivers address student supports (Merseth, 2011). Quantway 1 integrates two types of research-based student supports designed to meet the needs of diverse student learners: productive persistence and language and literacy supports. One set of supports is designed to promote students' ability to productively persist through rigorous math coursework. The socioemotional intervention, called productive persistence, consists of a collection of student activities and faculty actions that address the high-leverage noncognitive factors that promote student tenacity and effective learning strategies (Edwards & Beattie, 2016). NIC members worked together with social psychologists to iteratively develop this *package* of productive persistence routines and practices that work to promote growth mindset, reduce math anxiety, and increase students' sense of belonging. A second set of interventions is designed to support students in successfully grappling with the complex language and literacy demands of mathematics, which often incorporates different forms of representation and elaborate grammatical forms. Quantway 1 lessons embed language and literacy tools to facilitate the comprehension and organization of information in quantitative situations. These lessons are written to avoid literacy barriers that developmental math students commonly face (Gomez, Gomez, et al., 2015, Gomez, Rodela et al., 2013).

Finally, the fifth and sixth drivers address faculty development and participation in a NIC. Because the Quantway 1 curriculum and pedagogy significantly differ from traditional methods of teaching, Quantway 1 faculty are invited to participate in a comprehensive professional development program (Edwards et al., 2015). This faculty support program prepares instructors to teach Quantway 1 and supports them in their first year of teaching, and provides ongoing opportunities for instructional improvement and professional learning. Through online resources, faculty mentorship, and workshops, this program equips faculty to effectively implement the Pathways' collaborative instructional approach, learning opportunities, and productive persistence and language and literacy supports. Quantway 1 faculty and administrators participate in a NIC that provides a collaborative learning community to support them in teaching and implementing Quantway 1. The NIC social structure supports community college faculty and administrators in collectively generating and disseminating practical learning about what works, for whom, and under what conditions to reliably deliver efficacy at scale (Bryk et al., 2015).

Study Objective

The objective of the current study was to assess the effectiveness of Quantway 1. We used a propensity score matching technique (Rosenbaum & Rubin, 1983) to statistically reduce possible selection bias by comparing Quantway 1 student success to a reasonable counterfactual that represents how similar students would have performed if they had not taken Quantway 1. Given the nested structure of the data, with students enrolled within institutions in the network, we employed a hierarchical linear modeling (HLM) framework (Raudenbush & Bryk, 2002) to obtain propensity scores. We then compared completion rates of developmental mathematics sequences between Quantway 1 and the non-Quantway 1 matched comparison groups. Furthermore, we tracked their college mathematics achievement throughout the following year to determine if Quantway 1 students perform comparably or better than their matched students in college mathematics courses. As Quantway 1 was designed not only to get students through their developmental mathematics sequences but also to prepare them to meet their college mathematics requirements, this analysis was particularly important in determining Quantway 1's effectiveness. We applied an HLM approach to all of these outcome analyses because potential outcomes most likely depended on the institutions and classrooms to which students belong (Hong & Raudenbush, 2005, 2006).

In addition to the average impact of Quantway 1, we looked at variation in performance across classrooms and institutions in the network. These investigations into variation at different levels enabled us to assess Quantway 1's ability to scale with efficacy and inform where improvement efforts should be targeted to further increase success rates (Bryk et al., 2015). We examined possible differential effects of Quantway 1 across sex, race/ethnicity subgroups to determine its potential to promote an equity agenda by improving outcomes across all these subgroups.

Another objective of the current study was to discuss the potential of the Carnegie Math Pathways as a developmental mathematics reform initiative and any policy implications. Two pathways, Statway and Quantway, represent two different programs with different focuses, and an earlier study demonstrated the success of the former (Yamada & Bryk, 2016). By providing empirical evidence of Quantway 1's efficacy, this study offered a unique opportunity to assess the Pathways approach in general. We aimed to share our knowledge of developmental mathematics reform efforts, particularly those with an application of improvement science and NICs, with the field.

Method

Study Setting

Quantway 1 was first implemented during the spring of 2012.[4] The initial cohort of students spanned eight community colleges across three states (Georgia, New York, and Ohio). In the past 4 academic years, Quantway 1 served a total of 5,561 students from 14 colleges (see Appendix) across eight states (Georgia, Minnesota, New Jersey, New York, Ohio, Washington, West Virginia, and Wisconsin; Huang et al., 2016).

Data and Study Design

Institutional researchers from participating colleges provided background data on student characteristics, course enrollment, and course performance. The analytic sample consisted of 4,243 Quantway 1 students from 10 colleges (see Appendix) who enrolled in a Quantway 1 course between the spring of 2012 and the fall of 2014, as well as 83,887 potential comparison group students from the same respective institutions and corresponding semesters. The sample size of Quantway 1 students was somewhat smaller than the total number of students described above due to insufficient data for conducting adequate propensity score matching and outcome analyses.

Table 1 presents all of the covariates used in propensity score matching and their descriptive statistics before and after propensity score matching was conducted. We selected those covariates based on prior research findings (Yamada & Bryk, 2016) and advice from institutional researchers in the participating colleges. The list includes standard student background data such as sex and race/ethnicity. It has been shown that these characteristics tend to differentiate students' progress in developmental mathematics sequences (Bailey et al., 2010). We also matched on students' prior course-taking history and performance in the past 2 years. Previous research demonstrated that students' prior course-taking history and success patterns are a more reliable indicator of students' educational and career goals than their declared program of study (Jenkins & Cho, 2012).

The descriptive data on the left panel of Table 1 shows that, overall, the Quantway 1 group comprised higher proportions of female and Hispanic students than the non-Quantway 1 group. Quantway 1 students had more course records in the 2 years before taking a Quantway 1 course, suggesting that the term in which they took a Quantway 1 course was less likely to be their first semester or year of enrollment at a given institution. Quantway 1 students also started their developmental course(s) earlier, and attempted more developmental mathematics courses and college-level courses than non-Quantway 1 students.

Propensity score matching. First, we conducted propensity score matching to identify a group of students who followed traditional developmental mathematics sequences with similar characteristics to Quantway 1 students. To obtain propensity scores, we took an HLM approach (Hong & Raudenbush, 2005, 2006; Raudenbush & Bryk, 2002; Yamada & Bryk, 2016) and constructed a two-level model with a total of 37 student-level covariates including student background characteristics and prior course taking and success patterns during the 2 years prior to the Quantway 1 term (see Table 1). We estimated model parameters by leveraging maximum likelihood via penalized quasi-likelihood estimation using HLM 7 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011).[5] In this model, η_{il} is the log-odds of student i enrolling in Quantway 1 in college l and formally expressed as follows:

Level 1 Model (student)

$$\eta_{il} = \beta_{0l} + \beta_{1l}\left(\mathrm{COV1}_{il}\right) + \cdots + \beta_{37l}\left(\mathrm{COV37}_{il}\right),$$

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

	Sample before matching				Sample after matching			
	Non-Quantway 1		Quantway 1		Non-Quantway 1		Quantway 1	
	%		%		%		%	
Sex								
Female[a]	56		62		60		61	
Male	44		37		39		38	
Unknown	0		1		1		1	
Race/Ethnicity								
Asian	3		4		3		4	
Black	31		30		32		30	
Hispanic	18		26		21		26	
White[a]	36		34		36		33	
Multiracial	1		1		1		1	
Other	1		0		0		0	
Unknown	9		5		6		5	
Any course records in past 2 years								
No[a]	45		38		46		41	
Yes	55		62		54		59	
Cohort group								
Winter 2012	6		2		3		2	
Spring 2012	16		11		16		11	
Fall 2012	14		13		20		13	
Spring 2013	16		20		14		20	
Fall 2013	16		21		20		21	
Spring 2014[a]	17		15		14		15	
Fall 2014	15		19		14		19	
Age missing	21		14		23		15	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (in years)	24.98	8.47	24.09	8.18	24.09	7.55	23.97	8.11
Semesters since first developmental math course	0.92	1.70	1.62	2.46	1.27	2.12	1.37	2.25
Course load	3.69	1.20	3.95	1.14	3.95	1.19	3.93	1.14
Developmental math								
One level below college level								
Number of courses attempted	0.09	0.31	0.21	0.55	0.15	0.41	0.17	0.47
Success rate	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.03
Two levels below college level								
Number of courses attempted	0.21	0.48	0.35	0.65	0.26	0.56	0.30	0.58
Success rate	0.11	0.31	0.19	0.38	0.13	0.33	0.17	0.36

(continued)

Table 1. (continued)

	M	SD	M	SD	M	SD	M	SD
Three or more levels below college level								
Number of courses attempted	0.18	0.50	0.16	0.55	0.12	0.45	0.13	0.49
Success rate	0.10	0.29	0.08	0.27	0.07	0.25	0.07	0.25
Developmental English								
Number of courses attempted	0.10	0.37	0.08	0.33	0.08	0.35	0.08	0.32
Success rate	0.06	0.24	0.05	0.21	0.05	0.21	0.05	0.21
Developmental reading								
Number of courses attempted	0.07	0.28	0.11	0.38	0.09	0.32	0.11	0.37
Success rate	0.05	0.21	0.06	0.24	0.05	0.22	0.06	0.23
Developmental writing								
Number of courses attempted	0.10	0.32	0.11	0.34	0.09	0.32	0.10	0.34
Success rate	0.07	0.25	0.07	0.26	0.06	0.24	0.07	0.26
College math								
Number of courses attempted	0.03	0.18	0.08	0.29	0.05	0.27	0.07	0.28
Success rate	0.01	0.07	0.01	0.08	0.00	0.06	0.01	0.07
College nonmath								
Number of courses attempted	2.10	3.57	4.32	6.01	3.29	5.17	3.76	5.42
Success rate	0.33	0.42	0.42	0.42	0.35	0.42	0.40	0.42
College STEM								
Number of courses attempted	0.30	1.11	0.49	1.30	0.44	1.48	0.45	1.25
Success rate	0.07	0.24	0.09	0.26	0.08	0.25	0.08	0.25

Note. Age was computed in years using a date of birth and 9/1 for the fall cohorts and 3/1 for the spring cohorts (1/1 for one winter cohort group). In the current propensity model, we centered age around age 18. Semesters since first developmental math course takes an integer, such as 0, 1, 2, and so on, where 0 means a student took a developmental math course for the first time in the same term as the Quantway 1 term, 1 means one semester before, 2 means two semesters before, and so on. Course load refers to the number of courses a student took during the Quantway 1 term. Success rate was computed by dividing the number of courses completed with a pass in a pass/fail grading scheme, or a C or higher (C– if a ± grading scheme is used) by the number of courses attempted. STEM = science, technology, engineering, and mathematics.
[a]Reference categories (coded as 0) when formulating dummy variables.

Level 2 Model (college)

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

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

$$\ldots,$$

$$\mathrm{B}_{37l} = \gamma_{370.}.$$

In the equations, COV1 to COV37 are the set of propensity score covariates.[6] We found a substantial number of unknown records for students' date of birth when computing students' age in years. To factor these cases into the propensity model, we constructed a dummy variable in which missing age was coded as 1, and nonmissing age was coded as 0 (Stuart, 2010). Also, we accounted for six cohort groups by formulating a set of dummy variables with Spring 2014 as a reference category.

We conducted propensity score matching separately for each cohort and college by applying a nearest neighbor matching algorithm (Rosenbaum & Rubin, 1985). This algorithm was appropriate for our study because we wanted to retain as many Quantway 1 students as possible and had a large pool of non-Quantway 1 students for creating matches. We attempted to find up to five matches per Quantway 1 student (5:1 ratio matching) to maximize the best matches from the non-Quantway 1 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). For propensity score matching, we used the package MatchIt (Ho, Imai, King, & Stuart, 2011) in R (R Core Team, 2015).

HLM outcome analysis. We next estimated Quantway 1's effectiveness by comparing: (a) success rates in developmental mathematics sequences, (b) enrollment rates in college-level mathematics courses in the subsequent calendar year including a summer term where applicable (e.g., tracking student mathematics course enrollments over the spring, summer, and fall terms for the fall cohorts), and (c) the corresponding college mathematics grade point average (GPA) of Quantway 1 students with their matched counterparts. Success was defined as a passing grade or a grade of C or higher[7] (coded as 1, otherwise 0) on a Quantway 1 course for Quantway 1 students and a developmental mathematics course one level below college level (or another course deemed equivalent to a Quantway 1 course by faculty) for the matched comparison students. For the latter group, we tracked course outcomes over the entire academic year (i.e., tracking course outcomes over the fall and spring semesters for the fall cohorts and the spring, summer, and fall semesters for the spring cohorts). As described earlier, Quantway 1 was designed to accelerate traditional developmental mathematics sequences in one semester for students placed two levels below college mathematics. Similar students following the traditional developmental mathematics route would need 1 to 1½ years to complete the developmental sequence (Bailey et al., 2010; Cullinane & Treisman, 2010). Accordingly, if comparison students in the fall cohorts had failed a developmental mathematics course one level below college level in the fall semester but passed it the following spring semester, we counted it as success. Therefore, the analysis was conservative, providing comparison students twice as much time to reach the same success benchmark as Quantway 1 students. As student follow-up measures, we tracked student enrollment in college-level mathematics courses and their corresponding GPA[8] right after the Quantway 1 enrollment period for Quantway 1 students. For matched comparison students, we tracked the same outcomes over the same time period right after: (a) they had successfully completed their developmental mathematics sequences (e.g., completing the requirements in one semester) or (b) the entire academic year had passed regardless of their success in developmental mathematics sequences. College mathematics enrollment was defined as having attempted at least one course (coded as 1, otherwise 0).

To estimate differences in success rates and college mathematics enrollment rates, we constructed a four-level model and estimated its model parameters using maximum likelihood via penalized quasi-likelihood estimation. This four-level model was an extended application of the standard HLM model that took advantage of our matching procedure in which each Quantway 1 student was matched with up to five other students based on similar propensity scores. One of the study's objectives was to examine variation in performance across classrooms and institutions to see if and where improvement efforts should be targeted. With the HLM model, we could capture not only the average effect of Quantway 1 but also variability in its effect for each level (Hong & Raudenbush, 2005, 2006; Raudenbush & Bryk, 2002). In the Level 1 model below, η_{ijkl} is the log-odds of success or enrollment, and QW_{ijkl} is a dummy variable indicating whether the outcome was from the Quantway 1 group (coded as 1) or the matched comparison group (coded as 0). We included individual propensity scores, PS_{ijkl}, as a further safeguard to remove any potential remaining bias (Hong & Raudenbush, 2005, 2006) The Level 1 Model (measurement) is formally expressed as: $\eta_{ijkl} = \pi_{0jkl} + \pi_{1jkl}(QW_{ijkl}) + \pi_{2jkl}(PS_{ijkl})$. The Level 1 model is in essence a measurement model for the individual effect of Quantway 1 on each separate student (Raudenbush & Bryk, 2002). Both parameters at Level 1, π_{0jkl} and π_{1jkl}, are substantively interpretable. Of primary interest for this research, π_{1jkl} captures the effect of Quantway 1 on student j in faculty member k's classroom nested within college l. In comparison, π_{0jkl} is the outcome for their individually matched comparisons. In the context of modern causal theory, this is the potential outcome we might have observed had this particular Quantway 1 student followed a traditional program of study instead (Holland, 1986; Neyman, 1990; Rubin, 1974, 1978). We note that the correlation between π_{0jkl} and π_{1jkl} as it provides information about how program effects are distributed among different students, classrooms, and colleges (Raudenbush & Bloom, 2015). Specifically, a negative correlation indicates that positive effects are more likely to accrue for students whose prognosis for success is otherwise very low. In contrast, a positive correlation would indicate that those who are already more likely to succeed also tend to gain more benefit from Quantway 1. To complete the four-level model, we included a set of covariates at Level 2 as additional adjustment variables for the outcome.

Level 2 Model (student):

$$\pi_{0jkl} = \beta_{00kl} + \beta_{01kl}\left(\mathrm{COH1}_{jkl}\right) + \cdots + \beta_{0gkl}\left(\mathrm{COH}g_{jkl}\right) + r_{0jkl},$$
$$\pi_{1jkl} = \beta_{10kl} + \beta 1_{1kl}\left(\mathrm{COH1}_{jkl}\right) + \cdots + \beta_{1gkl}\left(\mathrm{COH}g_{jkl}\right) + r_{1jkl},$$
$$\pi_{2jkl} = \beta_{20kl}.$$

$COH1_{jkl}$ to $COHg_{jkl}$ are dummy variables for the cohort groups described earlier, where $g = 6$ for the success rates, and $g = 4$ for the enrollment rates.[9] No covariates were included at Levels 3 and 4.[10]

To estimate differences in college mathematics GPAs, we ran a four-level model similar to those described above using maximum likelihood via full maximum likelihood estimation.[11] $QWEC_{ijkl}$ indicates whether the outcome was from the Quantway 1 group (coded as 1) or the matched comparison group (coded as –1), $ENROLLEC_{ijkl}$ represents college mathematics enrollment (1 for having attempted at least one course and –1 for not having attempted any courses), and INT_{ijkl} is an interaction term of these

Level 3 Model (faculty): Level 4 Model (college):

$$\begin{aligned}
&\beta_{00kl} = \gamma_{000l} + u_{00kl}, && \gamma_{000l} = \delta_{0000} + v_{000l},\\
&\beta_{01kl} = \gamma_{010l}, && \gamma_{010l} = \delta_{0100},\\
&\ldots, && \ldots,\\
&\beta_{0gjkl} = \gamma_{0g0l}, && \gamma_{0g0l} = \delta_{0g00},\\
&\beta_{10kl} = \gamma_{100l} + u_{10kl}, && \gamma_{100l} = \delta_{1000} + v_{100l},\\
&\beta_{11kl} = \gamma_{110l}, && \gamma_{110l} = \delta_{1100},\\
&\ldots, && \ldots,\\
&\beta_{1gkl} = \gamma_{1g0l}, && \gamma_{1g0l} = \delta_{1g00},\\
&\beta_{20kl} = \gamma_{200l}. && \gamma_{200l} = \delta_{2000}.
\end{aligned}$$

two variables. We applied effect coding to these grouping variables. Effect coding was more appropriate than dummy coding to directly estimate both main and interaction effects on the outcome for ease of interpretation.

Level 1 Model (measurement):

$$\begin{aligned}
\text{GPA}_{ijkl} = \pi_{0jkl} + \pi_{1jkl}\left(\text{QWEC}_{ijkl}\right) +\\
\pi_{2jkl}\left(\text{ENROLLEC}_{ijkl}\right) + \pi_{3jkl}\left(\text{INT}_{ijkl}\right) + \pi_{4jkl}\left(\text{PS}_{ijkl}\right),
\end{aligned}$$

Level 2 Model (student):

$$\begin{aligned}
\pi_{0jkl} &= \beta_{00kl} + \beta_{01kl}\left(\text{COH1}_{jkl}\right) + \cdots + \beta_{0gkl}\left(\text{COH}g_{jkl}\right) + r_{0jkl},\\
\pi_{1jkl} &= \beta_{10kl} + \beta_{11kl}\left(\text{COH1}_{jkl}\right) + \cdots + \beta_{1gkl}\left(\text{COH}g_{jkl}\right),\\
\pi_{2jkl} &= \beta_{20kl} + \beta_{21kl}\left(\text{COH1}_{jkl}\right) + \cdots + \beta_{2gkl}\left(\text{COH}g_{jkl}\right),\\
\pi_{3jkl} &= \beta_{30kl} + \beta_{31kl}\left(\text{COH1}_{jkl}\right) + \cdots + \beta_{3gkl}\left(\text{COH}g_{jkl}\right),\\
\pi_{4jkl} &= \beta_{40kl}.
\end{aligned}$$

Level 3 Model (faculty): Level 4 Model (college):

$$\begin{aligned}
&\beta_{00kl} = \gamma_{000l} + u_{00kl}, && \gamma_{000l} = \delta_{0000} + v_{000l},\\
&\beta_{01kl} = \gamma_{010l}, && \gamma_{010l} = \delta_{0100},\\
&\ldots, && \ldots,\\
&\beta_{0gjkl} = \gamma_{0g0l}, && \gamma_{0g0l} = \delta_{0g00},\\
&\beta_{11kl} = \gamma_{110l}, && \gamma_{110l} = \delta_{1100},\\
&\ldots, && \ldots,\\
&\beta_{1gjkl} = \gamma_{1g0l}, && \gamma_{1g0l} = \delta_{1g00},\\
&\beta_{21kl} = \gamma_{210l}, && \gamma_{210l} = \delta_{2100},\\
&\ldots, && \ldots,\\
&\beta_{2gjkl} = \gamma_{2g0l}, && \gamma_{2g0l} = \delta_{2g00},\\
&\beta_{31kl} = \gamma_{310l}, && \gamma_{310l} = \delta_{3100},\\
&\ldots, && \ldots,\\
&\beta_{3gjkl} = \gamma_{3g0l}, && \gamma_{3g0l} = \delta_{3g00},\\
&\beta_{40kl} = \gamma_{400l}. && \gamma_{400l} = \delta_{4000}.
\end{aligned}$$

Subgroup analysis. To examine possible differential effects of Quantway 1 by sex and race/ethnicity subgroups on the three outcomes of interest, we ran four-level HLM models similar to those described above. In this set of subgroup analyses, we applied effect coding to the grouping variables to directly estimate both main and interaction effects on the outcomes, as we did in the analysis on college mathematics GPAs. The reference categories were female and White for the sex and race/ethnicity variables, respectively, each of which was coded as –1. We excluded cases where sex had not been specified.

Sensitivity analysis. We conducted sensitivity analyses when Quantway 1 effects were significant. The validity of the effects was based on an assumption of a strongly ignorable treatment assignment. In other words, all relevant covariates were included in the propensity score analysis, so that the bias due to unmeasured covariates could be ignored. Thus, we examined the sensitivity of the estimated Quantway 1 effects to possible unmeasured confounders (Hong & Raudenbush, 2005, 2006; Lin, Psaty, & Kronmal, 1998). Given some unmeasured covariates (U), the Quantway 1 effect (δ) can be reestimated by adjusting for hypothesized hidden bias $\{\gamma[E(U_1) - E(U_0)]\}$ as $\delta^* = \delta - \gamma[E(U_1) - E(U_0)]$, where γ is the unmeasured covariates' association with the outcome and $[E(U_1) - E(U_0)]$ is their association with treatment assignment (i.e., enrolled in Quantway 1 or one of the matched comparisons). Adapting the approach of Hong and Raudenbush (2005, 2006), we operationally defined a proxy for γ as a coefficient derived from a four-level model designed to predict the outcome with the same set of covariates used in the propensity score analysis and $[E(U_1) - E(U_0)]$ as the observed mean difference between the Quantway 1 and matched comparison groups on the corresponding covariate. We then selected the largest positive value of the product of these two values as the largest possible bias[12] and obtained an adjusted Quantway 1 estimate (δ^*). Accordingly, we reestimated the effect of Quantway 1 on the outcome and constructed a 95% confidence interval (CI) for the adjusted estimate.

Results

Propensity Score Matching

We matched a total of 12,448 comparison students to 3,992 Quantway 1 students and ran an HLM analysis on their outcome data with regard to developmental mathematics completion. Table 1 compares the descriptive statistics on each covariate before and after matching to the Quantway 1 group. Table 2 documents the balance in propensity score cohort by cohort for each college. There were no significant differences in mean propensity scores between the Quantway 1 and matched comparison students in any of the cohorts for each college (see t values). These results provide strong evidence that comparability of the groups was achieved on the measured covariates.

It may be worthwhile here to mention the matched ratios we accomplished. As described earlier in the Method section, we attempted to find up to five matches per Quantway 1 student. The matched ratios in the far right column suggest that in general,

Table 2. Balance in Logit of the Propensity Score for Non-Quantway and Quantway Students.

		Sample before matching						Sample after matching								
		Non-Quantway 1			Quantway 1			Non-Quantway 1			Quantway 1					
College	Cohort	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*t*	Difference in *SD*	Matched ratio
1	2012 Spring	585	−2.78	0.48	43	−2.72	0.37	212	−2.74	0.32	43	−2.72	0.37	−0.42	−0.05	4.93
1	2012 Fall	470	−2.41	0.31	34	−2.20	0.54	149	−2.36	0.35	31	−2.31	0.42	−0.59	−0.06	4.81
1	2013 Fall	305	−1.88	0.19	59	−1.79	0.38	249	−1.89	0.15	54	−1.88	0.16	−0.12	−0.01	4.61
1	2014 Spring	273	−2.32	0.26	29	−2.01	0.52	112	−2.23	0.26	25	−2.16	0.34	−0.94	−0.08	4.48
2	2013 Fall	690	−2.34	0.96	69	−2.29	0.80	337	−2.31	0.75	69	−2.29	0.80	−0.16	−0.05	4.88
2	2014 Spring	270	−2.93	0.62	17	−3.02	0.27	85	−3.02	0.27	17	−3.02	0.27	−0.01	−0.01	5.00
2	2014 Fall	402	−2.86	0.41	47	−2.99	0.36	217	−2.95	0.26	46	−2.96	0.32	0.18	−0.06	4.72
3	2012 Spring	4,138	−3.38	0.39	72	−3.37	0.27	360	−3.37	0.27	72	−3.37	0.27	−0.05	0.00	5.00
3	2012 Fall	3,234	−2.86	0.40	177	−2.74	0.42	875	−2.76	0.38	175	−2.76	0.38	−0.07	0.00	5.00
3	2013 Spring	3,745	−2.24	0.43	584	−1.93	0.75	1,057	−2.11	0.53	544	−2.06	0.59	−1.66	−0.06	1.94
3	2013 Fall	2,358	−2.11	0.51	408	−1.66	0.93	739	−1.89	0.64	378	−1.84	0.72	−1.28	−0.08	1.96
3	2014 Spring	3,242	−2.58	0.54	290	−2.12	0.81	559	−2.20	0.72	287	−2.14	0.78	−1.08	−0.06	1.95
3	2014 Fall	4,696	−2.46	0.36	402	−1.90	0.94	368	−2.08	0.69	368	−2.07	0.71	−0.16	−0.01	1.00
4	2012 Spring	594	−3.20	0.66	38	−2.78	1.27	175	−3.03	0.88	36	−2.97	0.97	−0.33	−0.09	4.86
4	2012 Fall	570	−2.88	0.47	45	−2.40	0.83	193	−2.67	0.48	42	−2.54	0.65	−1.22	−0.17	4.60
4	2013 Spring	599	−2.20	0.48	42	−1.91	0.73	188	−2.09	0.50	41	−1.96	0.68	−1.18	−0.18	4.59
4	2013 Fall	681	−2.33	0.46	69	−2.01	0.89	278	−2.28	0.48	63	−2.20	0.66	−0.84	−0.17	4.41

(continued)

Table 2. (continued)

		Sample before matching						Sample after matching								
		Non-Quantway 1			Quantway 1			Non-Quantway 1			Quantway 1					
College	Cohort	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*t*	Difference in *SD*	Matched ratio
4	2014 Spring	617	−2.78	0.50	49	−2.43	0.81	213	−2.65	0.57	47	−2.51	0.72	−1.16	−0.15	4.53
4	2014 Fall	827	−2.66	0.38	87	−2.48	0.72	408	−2.63	0.41	85	−2.56	0.53	−1.17	−0.13	4.80
5	2013 Fall	712	−3.88	0.58	6	−3.60	0.76	30	−3.61	0.68	6	−3.60	0.76	−0.02	−0.07	5.00
5	2014 Spring	648	−4.58	0.58	19	−3.65	0.93	82	−3.88	0.76	19	−3.65	0.93	−0.99	−0.17	4.32
6	2012 Spring	1,601	−4.72	0.45	21	−4.74	0.35	105	−4.74	0.34	21	−4.74	0.35	−0.05	−0.01	5.00
6	2013 Spring	1,481	−3.75	0.55	35	−3.44	0.87	167	−3.57	0.63	34	−3.52	0.73	−0.34	−0.10	4.91
6	2013 Fall	1,534	−3.82	0.49	49	−3.49	0.78	232	−3.60	0.65	48	−3.54	0.73	−0.53	−0.08	4.83
6	2014 Spring	1,436	−4.20	0.45	32	−3.76	0.87	140	−3.88	0.55	29	−3.83	0.58	−0.41	−0.03	4.83
6	2014 Fall	1,954	−4.16	0.35	17	−3.44	0.85	71	−3.70	0.63	16	−3.55	0.74	−0.72	−0.11	4.44
7	2012 Spring	3,675	−4.79	0.91	63	−4.49	1.02	305	−4.53	0.99	62	−4.49	1.03	−0.08	−0.04	4.92
7	2012 Fall	2,929	−4.68	0.58	65	−4.47	0.67	320	−4.52	0.58	65	−4.47	0.67	−0.35	−0.09	4.92
7	2013 Spring	3,831	−4.05	0.49	42	−3.92	0.52	204	−3.97	0.45	42	−3.92	0.52	−0.08	−0.07	4.86
7	2013 Fall	3,566	−4.21	0.44	68	−4.17	0.36	338	−4.18	0.34	68	−4.17	0.36	−0.02	−0.02	4.97
7	2014 Spring	4,097	−4.65	0.45	29	−4.29	0.80	127	−4.54	0.47	27	−4.41	0.70	−1.11	−0.23	4.70
8	2012 Winter	4,693	−5.45	1.30	70	−3.90	1.55	348	−3.93	1.52	70	−3.90	1.55	−0.16	−0.03	4.97
8	2012 Spring	1,955	−4.39	0.91	108	−2.32	1.83	275	−2.91	1.22	97	−2.72	1.36	−1.23	−0.14	2.84

(continued)

Table 2. (continued)

		Sample before matching						Sample after matching								
		Non-Quantway 1			Quantway 1			Non-Quantway 1			Quantway 1					
College	Cohort	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*t*	Difference in *SD*	Matched ratio
8	2012 Fall	3,420	−3.94	0.88	124	−2.14	1.69	443	−2.51	1.30	120	−2.29	1.47	−1.48	−0.17	3.69
8	2013 Spring	3,615	−3.36	0.60	129	−1.93	1.36	119	−2.13	1.17	119	−2.10	1.20	−0.19	−0.03	1.00
8	2013 Fall	3,230	−3.33	0.62	119	−1.87	1.46	105	−2.23	1.19	105	−2.20	1.20	−0.22	−0.02	1.00
8	2014 Spring	3,088	−3.79	0.52	118	−2.49	1.23	165	−3.16	0.74	88	−3.04	0.83	−1.10	−0.08	1.88
8	2014 Fall	3,796	−3.61	0.49	158	−2.55	1.37	253	−3.07	0.97	138	−2.89	1.09	−1.60	−0.12	1.83
9	2012 Spring	683	−2.86	0.56	41	−2.88	0.70	195	−2.97	0.56	40	−2.93	0.62	−0.37	−0.07	4.88
9	2012 Fall	776	−2.75	0.34	55	−2.33	1.00	226	−2.71	0.37	50	−2.58	0.54	−1.58	−0.16	4.52
9	2014 Spring	543	−2.59	0.38	30	−2.63	0.38	142	−2.59	0.33	29	−2.62	0.38	0.38	−0.04	4.90
9	2014 Fall	574	−2.49	0.29	71	−2.36	0.45	303	−2.43	0.27	68	−2.39	0.41	−0.91	−0.13	4.46
10	2012 Spring	520	−2.36	0.52	72	−2.41	0.58	343	−2.44	0.52	72	−2.41	0.58	−0.36	−0.07	4.76
10	2012 Fall	376	−2.33	0.38	48	−2.29	0.41	229	−2.32	0.37	48	−2.29	0.41	−0.43	−0.04	4.77
10	2013 Fall	243	−1.94	0.15	39	−1.87	0.16	164	−1.91	0.14	39	−1.87	0.16	−1.33	−0.02	4.21
10	2014 Spring	291	−2.03	0.40	27	−2.13	0.32	134	−2.14	0.31	27	−2.13	0.32	−0.13	0.00	4.96
10	2014 Fall	324	−2.12	0.31	27	−1.76	0.65	109	−2.05	0.25	22	−2.04	0.26	−0.22	−0.01	4.95

Table 3. Model-Based Estimation of Quantway 1 Effect on Developmental Math Success Rate.

	Coefficient	*SE*	*t*	*p*	Odds ratio
Fixed effect					
Matched comparison group mean outcome (intercept)	−0.49	0.15	−3.20	.005	0.61
Quantway 1 effect (slope)	0.72	0.21	3.47	.003	2.05
	Variance	*df*	χ^2	*p*	Correlation
Random effect at Level 4 (college)					
Matched comparison group mean outcome (intercept)	0.22	9	311.84	<.001	−.70
Quantway 1 effect (slope)	0.35	9	97.93	<.001	
Random effect at Level 3 (faculty)					
Matched comparison group mean outcome (intercept)	0.02	70	112.41	<.001	−.41
Quantway 1 effect (slope)	0.20	70	182.71	<.001	

Note. For brevity, we omitted in this table the coefficient estimates for the covariates including the cohort groups and the term tracked for the outcomes of the matched comparison students.

we identified four to five matches per Quantway 1 student. For some cohorts from Colleges 3 and 8, however, we identified fewer matches and needed to exclude some Quantway 1 students to maintain the comparability of the groups. It appears that both colleges have a relatively large population of students who were placed into developmental mathematics courses and accordingly more students at varying levels of developmental mathematics. Therefore, it may be possible that certain kinds of students (e.g., those who failed developmental courses multiple times) were advised to take Quantway 1 so as to limit the number of appropriate students for matching.

Developmental Mathematics Completion

The results presented in Table 3 indicate that on average, Quantway 1 students demonstrated significantly higher odds of success, odds ratio (OR) = 2.05, 95% CI = [1.33, 3.18],[13] in successfully completing the developmental mathematics course than the matched comparison students. The corresponding estimated probabilities of success were 56.50% for the Quantway 1 group and 38.74% for the matched comparison group. The estimated coefficients between the intercept and the slope at both college and faculty levels were negative (–.70 and –.41, respectively), suggesting that colleges and their associated faculty's classrooms with the lower mean outcomes of the matched comparison group produced larger Quantway 1 effects than those with the higher mean outcomes of the matched comparison group. Hence, Quantway 1 tends to reduce inequality in student outcomes across these different levels of settings. This tendency was stronger at the college level.

In addition, we found variation in Quantway 1 effects among colleges and faculty members (0.35 and 0.20 for the college and faculty variances). Figures 3 and 4 display

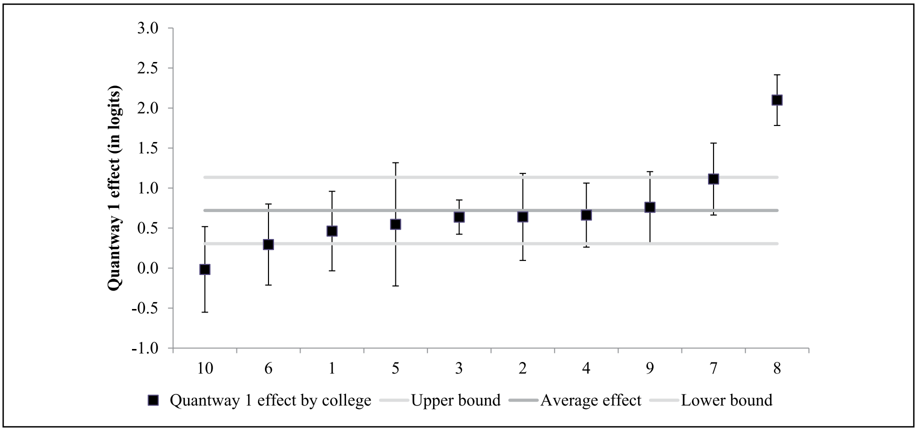

Figure 3. Variation among colleges in Quantway 1 effect on developmental math completion.

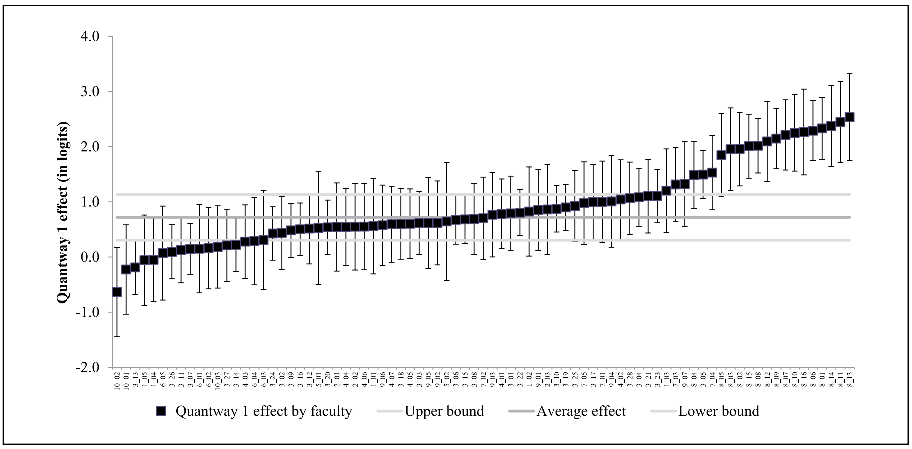

Figure 4. Variation among faculty members in Quantway 1 effect on developmental math completion.

the variation in Quantway 1 effect size at the college and faculty levels, respectively. In both charts, we added three lines as references. The center line represents the average effect of Quantway 1, and the upper and lower lines represent the upper and lower bounds of the average effect (which are deviated in two *SE*s from the center line). A value of 0 in logits means no Quantway 1 effects. Figure 3 demonstrates that there were positive Quantway 1 effects on student outcomes in all but College 10 (which showed no effect of Quantway 1). College 8 stands out as a positive deviant with a Quantway 1 effect outside the upper bound of the average effect. Figure 4 shows the variation in Quantway 1 effectiveness across the classrooms in the network. The vast

Table 4. Model-Based Estimation of Quantway 1 Effect on College Math Enrollment Rate.

	Coefficient	*SE*	*t*	*p*	Odds ratio
Fixed effect					
Matched comparison group mean outcome (intercept)	−1.11	0.17	−6.68	<.001	0.33
Quantway 1 effect (slope)	0.85	0.21	3.99	.001	2.33
	Variance	*df*	χ^2	*p*	Correlation
Random effect at Level 4 (college)					
Matched comparison group mean outcome (intercept)	0.22	8	154.40	<.001	−.55
Quantway 1 effect (slope)	0.30	8	66.07	<.001	
Random effect at Level 3 (faculty)					
Matched comparison group mean outcome (intercept)	0.01	44	60.36	.051	−.81
Quantway 1 effect (slope)	0.09	44	76.14	.002	

Note. For brevity, we omitted in this table the coefficient estimates for the covariates including the cohort groups and the individual propensity scores.

majority of Quantway 1 faculty at College 8 drastically outperformed the average Quantway 1 faculty, suggesting internal coherence at this institution. In contrast, a wide range of variation was observed among faculty members at College 3.

College Mathematics Achievement

Based on the results of the propensity score matching, the analysis of developmental mathematics completion was based on a total of 16,440 students (3,992 Quantway 1 students and 12,448 matched comparison students) across 10 colleges, leading to an average of approximately three matched students for each Quantway 1 participant. For the analysis of college mathematics achievement, however, we had to reduce the analytic sample to 10,184 students (2,406 Quantway 1 students and 7,778 matched comparison students) across nine colleges due to the limited availability of course-taking data following the enrollment of developmental mathematics courses (Quantway 1 courses for Quantway 1 students). The ratio of Quantway 1 students to the matched comparison students remained the same at 1:3.

The results presented in Table 4 indicate that on average, Quantway 1 students demonstrated significantly higher odds of attempt, OR = 2.33, 95% CI = [1.49, 3.66], at taking at least one college mathematics course than the matched comparison students. The corresponding estimated probabilities of enrollment were 50.35% for the Quantway 1 group and 30.31% for the matched comparison group, suggesting that about half of the Quantway 1 students enrolled in at least one college mathematics course in the subsequent year, and that less than a third of the matched comparison students did so. The estimated correlations between the intercept and the slope at both college and faculty levels were negative (–.55 and –.81, respectively), suggesting that

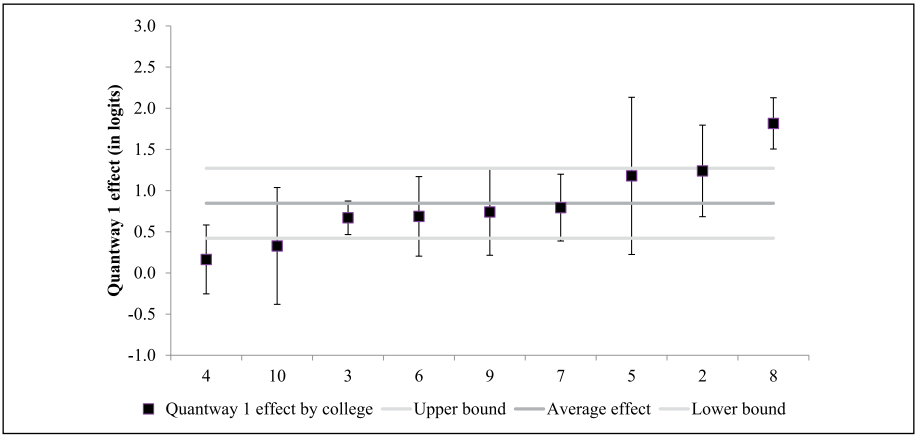

Figure 5. Variation among colleges in Quantway 1 effect on college math enrollment rates.

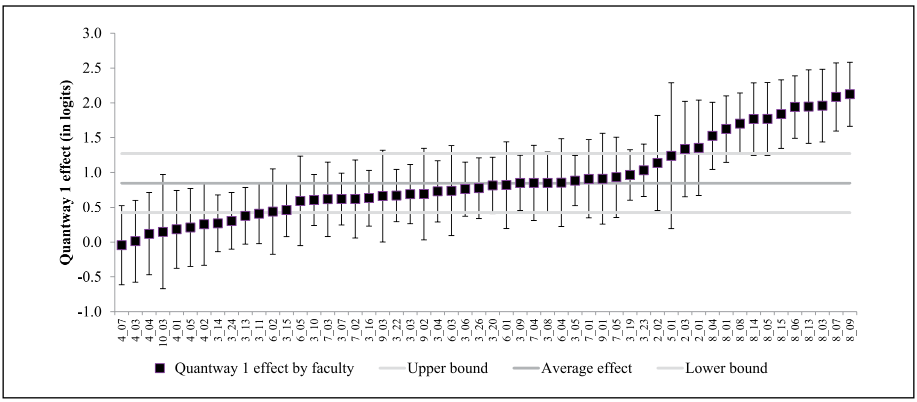

Figure 6. Variation among faculty members in Quantway 1 effect on college math enrollment rates.

as observed in the analysis of developmental mathematics completion, Quantway 1 tends to reduce inequality in student outcomes, and that this tendency was stronger at the faculty level.

In addition, we found variation in Quantway 1 effect among colleges and faculty members (0.30 and 0.09 for the college and faculty variances). Figures 5 and 6 display the variation in Quantway 1 effect size at the college and faculty levels, respectively. Figure 5 demonstrates that there were positive Quantway 1 effects on student outcomes in all colleges. College 8 stands out as a positive deviant with a Quantway 1 effect outside the upper bound of the average effect. Figure 6 shows the variation in Quantway 1 effectiveness across the classrooms in the network. The vast majority of Quantway 1 faculty at College 8 drastically outperformed the average Quantway 1 faculty,

Table 5. Model-Based Estimation of Quantway 1 Effect on College Math GPA.

	Coefficient	*SE*	*t*	*p*
Fixed effect				
Grand mean outcome (intercept)	0.57	0.04	14.50	<.001
Quantway 1 effect (slope)	0.04	0.01	3.78	<.001
College math enrollment (slope)	1.03	0.01	87.80	<.001
Interaction (slope)	0.04	0.01	3.87	<.001
	Variance	*df*	χ^2	*p*
Random effect at Level 4 (college)				
Grand mean outcome (intercept)	0.01	8	77.95	<.001
Random effect at Level 3 (faculty)				
Grand mean outcome (intercept)	<0.01	44	60.36	.051

Note. For brevity, we omitted in this table the coefficient estimates for the covariates including the cohort groups and the individual propensity scores. GPA = grade point average.

suggesting internal coherence at this institution. In contrast, a wide range of variation was observed among faculty members at College 3. These patterns in Colleges 8 and 3 were consistent with those found in the analysis of developmental mathematics completion.

The results presented in Table 5 indicate a significant interaction effect of Quantway 1 and college mathematics enrollment as well as significant main effects of these two variables. Our main focus was on this interaction effect because we were interested to see, among those who enrolled in college mathematics courses, whether Quantway 1 students earned a college mathematics GPA higher or lower than, or comparable with, the matched comparison students. For ease of interpretation, we transformed the model-based results into group mean GPAs and found that the GPAs, 2.22 for the Quantway 1 group and 2.06 for the matched comparison group, were comparable with each other. Although significant variation among colleges was also observed (0.01), this variation pertained to the intercept (grand mean of GPA), but not the slope (effect size of Quantway 1 on GPA among those who enrolled in college mathematics courses).[14] Overall, the results derived from college mathematics enrollment and GPA suggested that Quantway 1 students were more likely than comparable students who enrolled in traditional developmental mathematics sequences to attempt a college mathematics course in the following year and demonstrate comparable performance on college mathematics.

Subgroup Analysis

Figures 7, 8, and 9 present the model-based results transformed back into their natural metrics of proportions of students successfully completing a developmental mathematics sequence, enrolling in college mathematics courses in the subsequent year, and earning an associated GPA, respectively. This metric transformation was made for ease of interpretation. Positive effects of Quantway 1 were observed for all subgroups. More specifically, Black and Hispanic male students exhibited the largest increase in

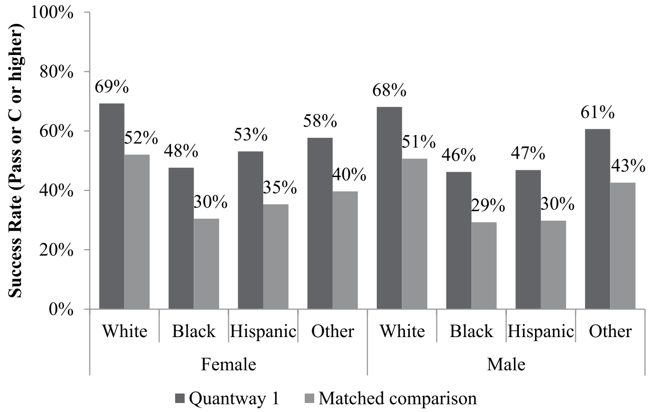

Figure 7. Model-based success rates by sex and race/ethnicity.

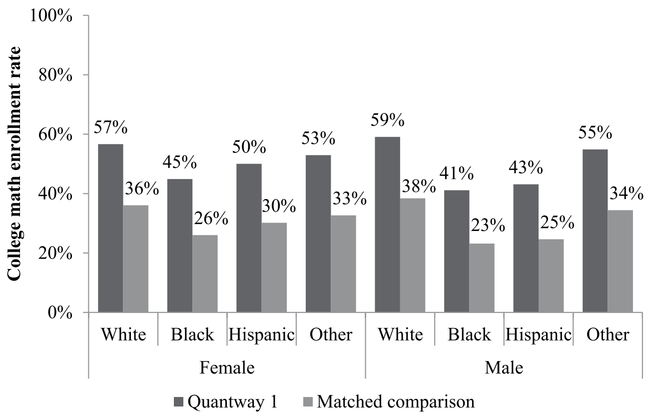

Figure 8. Model-based college math enrollment rates by sex and race/ethnicity.

developmental mathematics success rates and college mathematics enrollment rates relative to the corresponding subgroups of the matched comparison students, suggesting that they benefited most from Quantway 1. Each subgroup of students also showed a college mathematics GPA comparable with its matched comparison students. The results were consistent with the negative correlation estimates reported earlier in that Quantway 1 tends to reduce inequality in student outcomes.

Sensitivity Analysis

We also conducted sensitivity analyses when we obtained significant Quantway 1 effects.[15] The adjusted estimates for the Quantway 1 effect on developmental

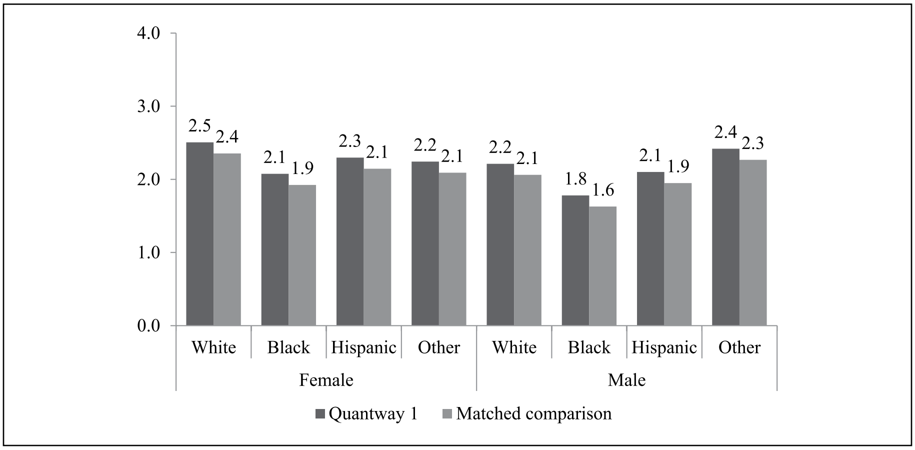

Figure 9. Model-based college math GPA by sex and race/ethnicity.
Note. GPA = grade point average.

mathematics success rates and college mathematics enrollment rates were .67 in logits (95% CI = [0.26, 1.08]) and .81 in logits (95% CI = [0.39, 1.22]), respectively. The corresponding CIs did not contain 0 or any negative value, thereby supporting the strong ignorability assumption. Our sensitivity analysis concluded that it was very unlikely that our general conclusion regarding Quantway 1's positive effects was influenced by the omission of unmeasured confounding factors.

Discussion

This study assessed Quantway 1's effectiveness for community college students across six semesters of implementation by means of a propensity score matching technique (Rosenbaum & Rubin, 1983) within an HLM framework (Raudenbush & Bryk, 2002). This approach allowed us to control for possible selection bias and increase the validity of causal inference. Throughout the outcome analyses, Quantway 1 students demonstrated significantly higher odds of success in completing developmental mathematics sequences and enrolling in college mathematics courses in the following year than their matched comparison students, and when enrolling in college mathematics courses, their outcomes were comparable with their counterparts. Our sensitivity analyses also indicated that these results were not due to unmeasured differences between the two groups. We conclude that Quantway 1 substantially improves student success in fulfilling developmental mathematics course requirements and student opportunity for acquiring college mathematics credit.

Although typical evaluations may stop at estimating the average impact of Quantway 1, this study also sought to understand its variation across different colleges, faculty members' classrooms, and student subgroups. To achieve efficacy at scale, Quantway 1 must not only produce a positive effect on average, but must also be effective for diverse student populations across a range of different classroom and institutional contexts. We found that Quantway 1 effects were positive across all sex and race/ethnicity subgroups of students. Because students from traditionally underserved groups including Black, Hispanic, and low-income students are more likely to enroll in developmental mathematics courses that have generated a disproportionate impact on them (Chen, 2016), these results are promising in that Quantway 1 can play a critical role in increasing the overall number of traditionally underserved students completing their mathematics requirements and hence in reducing outcome inequality across student subgroups. As consistent with this finding, we observed negative correlations between the Quantway 1 impacts and the mean outcomes of the matched comparison students across different levels of contexts, suggesting that Quantway 1 can contribute to advancing equitable outcomes (Raudenbush & Bloom, 2015).

Quantway 1 showed a positive effect in nearly all classrooms and colleges in the network, indicating that the program can work for varied faculty in different institutional settings. At the same time, we observed the significant variation in outcomes across faculty and colleges delineated in Figures 3 through 6. The goal of quality improvement is to reduce the variation between classrooms and colleges achieving positive results across diverse contexts. College 8, for example, significantly outperforms the other colleges in the network and maintains high performance across all the classrooms in the college, and this pattern persisted even in the following year when students enrolled in college mathematics courses. By contrast, College 3 consistently showed a wide range of variation among faculty members. Both colleges comprise a relatively large population of students who were placed into developmental mathematics courses. Future research should explore whether these colleges differ in how they enact the key design elements described above and study the various adaptations that these colleges made in response to their local context. In particular, College 8 may make a great case from which to learn how it effectively and reliably adapted Quantway 1 to local conditions. We can then spread its practices to other institutions, such as College 3, and facilitate network-wide progress through improvement science embedded in the NIC (Bryk et al., 2015). Discovering and sharing key practices across NIC colleges would enhance the network's ability to replicate Quantway 1's positive outcomes as it spreads to more diverse settings.

The results of Quantway as well as Statway (Yamada & Bryk, 2016) suggest that as portrayed in Figure 1, the Carnegie Math Pathways' comprehensive and systematic approach to tackling the typical barriers that developmental mathematics students face is key to its success. This holistic approach is indeed concordant with recent reviews of research in this area (Bailey et al., 2016; Hodara, 2013). Further empirical evidence may be needed to connect particular design elements to the positive effects of the Pathways. For now, we can conclude that the Pathways' multifaceted intervention packages are effective alternatives to the traditional developmental mathematics sequence, and accelerates the ability of a diverse range of students to complete their developmental mathematics requirements and achieve college mathematics credit in a variety of contexts. In particular, the Pathways' efficacy in advancing equity for historically underserved groups and across diverse contexts may have a profound implication to developmental mathematics reform. Without the Carnegie Math Pathways, many of those who successfully earned college mathematics credit would likely have needed an additional 1½ to 2 years to achieve the same outcome and might have given up their studies (Bailey et al., 2010). This would have resulted in negative ramifications for their educational, career, and life goals (Johnstone, 2013). The Pathways' approach to acceleration is particularly compelling for traditionally underserved students, who are more likely to be required to take and complete developmental mathematics sequences (Bailey et al., 2010; Chen, 2016). There are also large societal implications (Carnevale et al., 2013). Many community colleges are being pressed by local employers to produce more skilled workers. Students who leave without critical skills and credentials exacerbate existing labor shortages. It may be possible that if we multiplied Pathways success rates out for the 500,000 community college students who are annually placed into traditional developmental mathematics sequences (Bailey et al., 2010; Hoang et al., 2017), we could reduce that number by half. If we successfully improved the Pathways by reducing variation in performance through improvement science embedded in the NIC (Bryk et al., 2015), we might be able to save even more students' mathematical lives.

Finally, it is worth mentioning a future direction for further exploration. We are currently analyzing data obtained from the National Student Clearinghouse. Our particular interest is in 2-year and 4-year degree completion rates as well as transfer rates into 4-year colleges of the Pathways students as more distal outcomes. This analysis would further illuminate the extent and dimensions of the Pathways' long-term effectiveness against our goals of improving retention and completion.

Appendix

List of Participating Colleges

- Atlantic Cape Community College
- Borough of Manhattan Community College*
- Cuyahoga Community College*
- East Georgia State College*
- Madison College
- Marshall University
- Onondaga Community College*
- Ridgewater College*
- Rockland Community College*
- Sinclair Community College*
- South Georgia State College*
- University of North Georgia, Gainesville*
- University of Washington, Bothell
- Westchester Community College*

Note. Colleges with "*" provided data sufficient for conducting adequate propensity score matching and outcome analyses in this study.

Acknowledgments

Carnegie Math Pathways staff and networked improvement community (NIC) members made significant contributions to this research by implementing and adapting Quantway to their various contexts. In particular, we want to acknowledge the contributions of Anthony S. Bryk, Ann Edwards, Cinnamon Hillyard, Melrose Huang, Karon Klipple, and Suleyman Yesilyurt.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: 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.

Notes

1. The overarching design principles are (a) courses will focus on concepts over procedures, (b) instruction will make use of authentic contexts and real data, (c) struggling with problems—both large and small—is a core part of the instructional experience, (d) students will have access to appropriate technology to engage in practice outside of the classroom, and (e) all curricular materials (in-class, out-of-class, formative, and summative assessments) will be cohesive, articulate, and consistent in design.
2. These mathematical societies include the National Numeracy Network, American Mathematics Association of Two-Year Colleges, and the Mathematical Association of America.
3. Five institutions implemented Quantway 2, serving a total of 429 students over 3 years of implementation (Huang et al., 2016). Given its relatively small scale, we focused this efficacy study on Quantway 1.
4. One college was on a quarter system and implemented Quantway 1 for the first time in the winter of 2012. For the sake of simplicity, we included the results with the Spring 2012 cohort but conducted the propensity score matching for these students separately from the Spring 2012 cohort at this college.
5. We used HLM 7 for all hierarchical linear modeling (HLM) analyses in this study.
6. We initially included two covariates of college non-STEM (science, technology, engineering, and mathematics) courses (the number of courses attempted and the respective success rate). However, they involved collinearity with other covariates, and accordingly, the model did not converge. Thus, we excluded them from the propensity model.
7. A grade of C– or higher was used for a college that employed a ± grading scheme.
8. We assigned a value of 0 as a grade point to W (Withdrawal) and I (Incomplete) in this analysis to create a conservative metric while allowing us to maintain all matched students.
9. As described later, there were a smaller number of cohort groups in the analyses of college mathematics enrollment and grade point average (GPA) due to the limited availability of follow-up data.

10. For the success rate analysis, we also included in the Level 2 model TERM*jkl*, a dummy variable indicating whether the outcome for the matched comparison students was based on one semester (coded as 1) or the entire academic year (coded as 0). We also ran the same four-level model with individual propensity scores included in the Level 1 model and those cohort group variables added to the Level 2 model for the slope. The results from this model revealed no significant coefficients of these additional covariates and closely mirrored those from the simpler model. For ease of interpretation, we focus here on the results from the simpler model.
11. We also ran a series of random slope models. However, we observed high correlations involved in the slopes and the intercept, which suggested that a fixed slope model should be used. Results from those random slope models were very similar to those from the fixed slope model.
12. We used the sum of the product values for covariates requiring a set of dummy variables (e.g., cohort group, race/ethnicity).
13. HLM 7 generates 95% confidence intervals of odds ratios.
14. The obtained result does not necessarily mean there was no variation in Quantway 1 effect size among colleges. Data analyzed in this study were from a relatively small number of colleges (nine colleges), and it might be possible that we would detect significant, meaningful variation with more colleges.
15. We did not conduct this analysis for the Quantway 1 effect on college mathematics GPAs, because the GPAs of the Quantway 1 and matched comparison groups were practically comparable with each other.

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. doi:10.1002/pst.433

Bahr, P. R. (2008). Does mathematics remediation work? A comparative analysis of academic attainment among community college students. *Research in Higher Education*, *49*, 420-450. doi:10.1007/s11162-008-9089-4

Bailey, T., Bashford, J., Boatman, A., Squires, J., Weiss, M., Doyle, W., . . . Young, S. H. (2016). *Strategies for postsecondary students in developmental education—A practice guide for college and university administrators, advisors, and faculty*. Washington, DC: Institute of Education Sciences, What Works Clearinghouse.

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. doi:10.1016/j.econedurev.2009.09.002

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. doi:10.1111/j.1467-8624.2007.00995.x

Blair, R., & Getz, A. (2011, June 9). A brief history of the quantitative literacy movement. Retrieved from https://www.carnegiefoundation.org/blog/a-brief-history-of-the-quantitative-literacy-movement/

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.

Bueschel, A. C. (2004). The missing link: The role of community colleges in the transition between high school and college. In M. W. Kirst & A. Venezia (Eds.), *From high school to college: Improving opportunities for success in postsecondary education* (pp. 252-284). San Francisco, CA: Jossey-Bass.

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

Carnevale, A. P., Smith, N., & Strohl, J. (2013). *Recovery: Projections of jobs and education requirements through 2020*. Washington, DC: Center on Education and the Workforce, Georgetown University.

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

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

Edgecombe, N. (2011). *Accelerating the academic achievement of students referred to developmental education* (CCRC Working Paper No. 30). New York, NY: Teachers College, Community College Research Center, Columbia University.

Edwards, A., & Beattie, R. (2016). Promoting student learning and productive persistence in developmental mathematics. *NADE Digest*, *9*, 30-39.

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

Gillman, R. (2004). SIGMAA is Formed. *Focus, 24*(1), *5*.

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

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.

Handel, M. J. (2007, May-June). *A new survey of workplace skills, technology, and management practices (STAMP): Background and descriptive statistics*. Paper presented at the Workshop on Research Evidence Related to Future Skill Demands of the National Research Council of the National Academies, Washington, DC.

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. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (Vol. 1, pp. 371-404). Charlotte, NC: Information Age Publishing.

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.

Hoang, H., Huang, M., Sulcer, B., & Yesilyurt, S. (2017). *Carnegie math pathways 2015-2016 impact report: A five-year review*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Hodara, M. (2011). *Reforming mathematics classroom pedagogy: Evidence-based findings and recommendations for the developmental math classroom* (CCRC Working Paper No. 27). New York, NY: Teachers College, Community College Research Center, Columbia University.

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

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*, 945-960.

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. doi:10.3102/01623737027003205

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. doi:10.1198/016214506000000447

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.

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: Teachers College, Community College Research Center, Columbia University.

Johnstone, R. (2013). *Fiscal considerations of Statway® and Quantway®: We should be doing this anyway, but here's how it may help the bottom line*. National Center for Inquiry & Improvement. Retrieved from http://ncii-improve.com/wp-content/uploads/2017/06/NCII-Carnegie_SW-QW_Fiscal_Considerations_110713-Rob-NCII.pdf

Lin, D. Y., Psaty, B. M., & Kronmal, R. A. (1998). Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics*, *54*, 948-963.

Merseth, K. K. (2011). Update: Report on innovations in developmental mathematics—Moving mathematical graveyards. *Journal of Developmental Education*, *34*, 32-38.

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*, 21-55.

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

Neyman, J. (1990). On the application of probability theory to agricultural experiments: Essay on principles. Section 9. *Statistical Science*, *5*, 465-472.

Raudenbush, S. W., & Bloom, H. S. (2015). *Learning about and from variation in program impacts using multisite trials*. New York, NY: MDRC.

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

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

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.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*, 688-701.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, *6*, 34-58.

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

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*, 1-21.

Waycaster, P. (2001). Factors impacting success in community college developmental mathematics courses and subsequent courses. *Community College Journal of Research and Practice*, *25*, 403-416.

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. doi:10.1177/0091552116643162

Author Biographies

Hiroyuki Yamada is the director of analytics at the Carnegie Foundation for the Advancement of Teaching. He facilitates the development of an analytics agenda that supports the Foundation's networked improvement communities and leads the design of the analytic strategy and execution of that agenda. His primary interest revolves around practical applications of statistical and measurement models to inform actionable insight for decision making and improvement in education. He holds a Ph.D. in human development and education from UC Berkeley.

Angel X. Bohannon was a postbaccalaureate fellow at the Carnegie Foundation for the Advancement of Teaching, who is now a PhD Student at the Human Development and Social Policy program at Northwestern University. She is interested in research-practice partnerships as a mechanism for facilitating systemic change in both K-12 and community colleges. She holds a B.A. in Public Policy from the University of Chicago.

Alicia Grunow was a senior partner at the Carnegie Foundation for the Advancement of Teaching, who is now the cofounder of the Improvement Collective and a senior fellow at the Carnegie Foundation for the Advancement of Teaching. She specilizes in the science of improvement, helping teams, organizations, and networks to apply improvement methodologies to reach new levels of performance. She holds a Ph.D. from Stanford University.

Christopher A. Thorn was the director of knowledge management at the Carnegie Foundation for the Advancement of Teaching, who is now a Senior Research Associate at WestEd. His current work focuses on the application of improvement science methods to the challenge of scaling complex interventions. He completed his Doctor of Sociology at the University of Bielefeld.